Range-Free Passive Acoustic Localization

Aminreza Abrahami Saba [1], Hassan Abolhassani [2], Mohammad Ghodsi [3]

Computer Engineering Department, Sharif University of Technology

[1]*s_ebrahimi@ce.sharif.edu*
[2]*abolhassani@sharif.edu*
[3]*ghodsi@sharif.edu*

Abstract

We propose a range-free localization framework in which a network of randomly deployed acoustic sensors can passively use natural acoustic phenomena within its environment to localize itself. We introduce a novel approach for registration of sensors observations which takes advantage of a clustering technique on triplets of associated observations. A Bayesian filtering method is employed to incrementally improve system state estimation as more observations become available. To the best of our knowledge this is the first work done on range-free passive acoustic localization. Simulation experiments of the proposed algorithms are presented.

1. INTRODUCTION

An important first step in many applications of sensor networks is localization. Automatic localization in sensor networks is a well-studied problem and numerous techniques have been proposed in the literature to address it. These techniques fall under three main categories. Techniques in the first category equip each sensor with sufficient features (like GPS) so as to enable it to localize itself independently. These techniques are not feasible for large-scale sensor networks comprising hundreds of sensors. Furthermore, such techniques tend to have strict working conditions (for instance, a GPS node should have direct line-of-sight to at least 3 GPS satellites). The second group of techniques use inter-sensor range measurements to estimate the topology of the network. Such ranging measurements call for special hardware in each sensor node which will probably have no other use. Moreover, most accurate ranging techniques require direct line-of-sight between the two nodes which cannot be guaranteed considering the characteristics of many deployment environments. In [12] an active network of acoustic sensors and actuators has been presented that can accurately localize network nodes. Their technique depends upon range measurement among nodes using time-of-flight measurement of a sound signal. In [5] the authors have used inter-sensor range measurements to find the network topology using multi-dimensional scaling. [2], [11], [8] are other examples of works that have assumed presence of inter-sensor ranging hardware.

Techniques within the third category use no inter-sensor ranging but take advantage of a mobile which can be, in some way, sensed by the network nodes. These techniques assume that the mobile is somehow cooperative with the network or can localize itself independent of the network. The technique presented in [13] uses direct sensors-target range measurements via TDoA to localize sensors and track the target by iteratively accumulating measurements using a Bayesian filter. The solutions proposed in [9], [10] also use a self-localizing mobile which moves through the network. [3] uses an unlocalized mobile to localize the network by forming a set of geometrical constraints based on radio connectivity of mobiles and sensors.

In this paper we introduce RAPAL (RAnge-free Passive Acoustic Localization), a localization technique which uses no direct range measurements or cooperative mobiles. We show how to learn network topology by merely listening to near-field background acoustic phenomena within the environment. Each sensor node is assumed to have radio connectivity with a few nodes in its vicinity (ideally each node should have radio connectivity with at least all of its neighbors in the corresponding Delaunay triangulation of the sensor nodes). In [15] radio connectivity information of sensor nodes is used to coarsely estimate the network topology. RAPAL uses this technique to initialize its state variable and can then gradually increase the localization accuracy as it observes more and more acoustic events in the environment.

We code all unknown to-be-estimated parameters of the system (including sensors and targets positions, or any environmental parameters such as the speed of the propagation of sound) in a single probabilistic state vector. Our technique can robustly handle erroneous and spurious sensor observations. We put no restrictions on the position or movement of the sound sources. Every acoustic event in the environment that is heard by more than two sensors can be employed to increase network localization accuracy. One main difficulty in the way of such an approach is what we name it *out-of-order observation*. Since the speed of propagation of acoustic waves is considerably slow with respect to the normal operating areas that many sensor networks span, it is quite possible that the news of events from different sound sources reach a sensor out of the real order. We propose a solution for this poblem in section 2-A.

2. RAPAL

Localization in RAPAL is based on measurement of TDoA of acoustic events to different sensors. In this framework it is assumed that sensors are capable of identifying certain types of acoustic waves in the waves they observe. Generally the

framework can selectively use only those kinds of waves that are known to result in a higher accuracy in TDoA calculations (section 3). Several TDoA estimation techniques have been proposed in the literature, including cross correlation [6], and nonnegative deconvolution[7]. Another technique [14] is based on extracting characteristic points of a signal and assigning time-stamps to them. Thus the sensors either communicate recorded signals or the assigned timestamps. In either case the result can be virtually viewed as a set of timestamps assigned to sensors observations. Through the rest of this paper we use the notation ν_j^ψ to refer to the j-th observation of sensor ψ. These extracted timestamps are then passed to higher levels for further association and analysis. RAPAL algorithms are then responsible for correctly establishing the association among correlated sensors observations. Due to space limitation we do not describe further details about the signal processing algorithms and focus on the higher level parts of the RAPAL framework.

In our algorithm TDoA is calculated only for neighboring sensors, which are assumed to have direct radio connectivity. In this setting explicit time synchronization is only needed between neighboring nodes and is relatively easy to achieve.

Why acoustic sensors?: RAPAL has several interesting properties due to using acoustic sensors. In passive localization using environmental sound phenomena, the sensor network does not disturb its environment since it does not need to emit any waves. Furthermore, RAPAL localization and tracking algorithms are range-free. Ranging techniques usually pose some restrictions on the network deployment and operation which are not desirable. RAPAL is solely dependent on the time difference of observations of events by different sensors. Moreover, any targets of interest can be tracked passively using the sound they produce while they can not even detect the presence of the network. Moreover, apart from localization and Tracking, acoustic sensors are the natural sensing modality of many sensor networks applications, which makes acoustic localization and tracking highly suited for them.

### A. Observation Registeration

The order in which a sensor observes acoustic events in its environment is not necessarily the same order as they have actually happened. Moreover different sensors might hear different events in different orders. Thus, we cannot readily associate sensors observations with each other. For instance if sound sources τ_1 and τ_2 generate two acoustic event at times t_1 and t_2 respectively, then sensors ψ_1 and ψ_2 observe the events in different orders iff the following inequality holds,

$$J(\psi_1, \psi_1, \tau_1, \tau_2) < t_2 - t_1 < J(\psi_2, \psi_2, \tau_1, \tau_2) \quad (1)$$

where

$$J(\psi_1, \psi_2, x, y) = \zeta^{-1}(\|\psi_1 - x\| - \|\psi_2 - y\|),$$

and ζ is the speed of propagation of sound in the network environment. It can be shown that the rate of out of order observation by a single sensor is proportional to $\frac{\pi \lambda \mathcal{A}}{\zeta}$, where $\mathcal{A}$ is the area in which events are generated (region of interest),

Voronoi Tesselation of Sensors Map
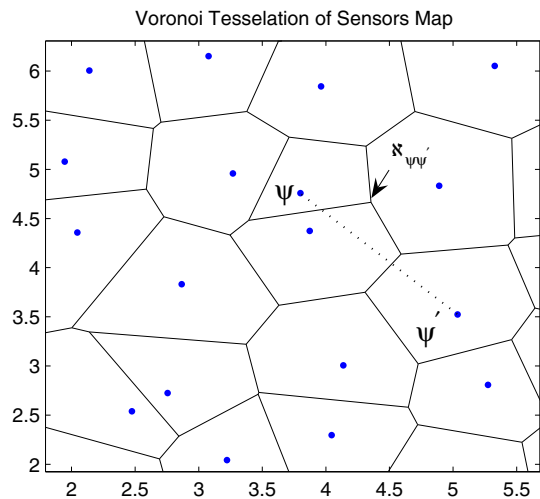
Fig. 1: Voronoi tessellation of a set of sensors. If sensor ψ assumes it is the first sensor to observe an event, then the event wave front can be neither closer to ψ' than $\|\psi'\aleph_{\psi\psi'} - \psi\aleph_{\psi\psi'}\|$ nor can it be farther than $\|\psi\psi'\|$.

and λ is the rate of event generation, assuming the events are generated by a spatial Poisson process. Although the probability of this happening is small in many cases, if not addressed, this phenomena can be intentionally exploited by an attacker to disrupt the localization and tracking algorithms of the network by a set of cooperative and synchronized sound sources. Because if out-of-order observations are not handled effectively, they appear as significant TDoA measurement errors (tens to hundreds of milliseconds) which will totally disrupt RAPAL's operation.

In the registration phase we want to recognize which observations at different sensors relate to one specific event. Let $\mathcal{V}_i$ denote the set of observations performed by ψ_i (including the null observation). In general the hypothesis space of the registration task is $\mathcal{V}_1 \times \ldots \times \mathcal{V}_p$ for each observed event. However, it is possible to limit the space as follows. If a sensor ψ assumes that it is the first sensor to observe an event, say observation ν_j^ψ, then based on its location and Voronoi cell in the corresponding Voronoi tessellation of the sensors, it can define a time window $\nu_j^\psi + \Delta_{\psi\psi'}$ within which, sensor ψ' is supposed to observe that same event (figure 1). Here $\Delta_{\psi\psi'} = [\Delta_{\psi\psi'}^{\min} \quad \Delta_{\psi\psi'}^{\max}]$. More generally, by relaxing the lower bound of this time window to zero, one can repeat the same statement for any observation of a sensor regardless of whether the sound source is in the Voronoi cell of the sensor. Let $\mathcal{L}_j^\psi(\psi')$, for $\psi \neq \psi'$, denote the set of all observations performed by the sensor ψ' within the $\nu_j^\psi + \Delta_{\psi\psi'}$ time window. Also $\mathcal{L}_j^\psi(\psi) \triangleq \{\nu_j^\psi\}$. Finally we define

$$\beth_j^\psi \triangleq \bigcup_{a < b} \mathcal{L}_j^\psi(\psi) \times \mathcal{L}_j^\psi(\psi') \times \mathcal{L}_j^\psi(\psi''), \quad (a,b) \in \mathbb{N}_p^2$$

$$\beth \triangleq \bigcup_{\nu_j^\psi} \beth_j^\psi, \quad (2)$$

where $\times$ denotes Cartesian product, and p is the number of

sensors in the network. Hence the interpretation of a triplet $\langle \nu_j^\psi, \nu_{j'}^{\psi'}, \nu_{j''}^{\psi''} \rangle$ in ∂ is that possibly there is an event that was heard by sensors ψ, ψ', ψ'' at the corresponding times denoted by the triplet. It will soon become clear why we use triplets of sensor observations instead of four or more sensors. Subsequently, we should propose a partitioning

$$\partial = \Xi \cup \bigcup_{i=1}^{m} \tilde{\partial}_{e_i} \tag{3}$$

such that all triplets of observations within a $\tilde{\partial}_{e_i}$ correspond to one specific event and Ξ is the set of yet unassigned triplets.

To elicit the desired partitioning of ∂, we first introduce a mapping of ∂ to a 2D space. A triplet $\langle \nu_j^\psi, \nu_{j'}^{\psi'}, \nu_{j''}^{\psi''} \rangle$ is mapped to point(s) x corresponding to the solutions of the following system of equation:

$$\begin{cases} J(\psi, \psi', x, x) = (\nu_j^\psi - \nu_{j'}^{\psi'}) \\ J(\psi', \psi'', x, x) = (\nu_{j'}^{\psi'} - \nu_{j''}^{\psi''}) \end{cases} \tag{4}$$

Each of the above equations corresponds to one branch of a hyperbolic curve and the solutions of this system of equations correspond to their intersection(s). The two hyperbolic curves can intersect each other in one or two points (they have a common focal point), where at most one of them is of interest to us. To find a unique intersection point we would have to amalgamate observations of at least 4 sensors, which give three hyperbolic curves. However, since sensors locations are approximate, the three hyperbolas may not have a common intersection point. Alternatively, we can use a region of interest criterion to choose between the two intersections of the two hyperbolas. However, we generally have more than three sensors and due to errors in sensors locations estimates, different sets of three sensors give different estimates for target location. As we will show none of these estimates can be trusted on its own. So the question is finding a procedure for proposing an initial estimate for the target location that converges to the true target location as the error in sensors locations estimates decreases. First we demonstrate the effect of errors in sensors locations estimates, on target localization accuracy based on only three sensors. Let $\Psi = [\psi_1, \ldots, \psi_6]$ be three sensors locations estimates, and let $\Delta = [\delta_1 \ \delta_2]$ be their respective time differences of observations of an event, and also let $\mathbf{g}(\Psi, \Delta) \triangleq [g_x(\Psi, \Delta) \ g_y(\Psi, \Delta)]$, be the solution of the corresponding system of equations 4. Here Ψ can be replaced by any set of three sensors locations estimates in the network, with their corresponding observation time differences as Δ. We want to analyze the sensitivity of $\mathbf{g}$ to error in Ψ for different Δ values.

$$\tilde{G}_k = \begin{bmatrix} \nabla_\Psi^k \mathbf{g}_x(\Psi, \Delta) \\ \nabla_\Psi^k \mathbf{g}_y(\Psi, \Delta) \end{bmatrix},$$
$$G_k = \left[\mathbf{trace}(\tilde{G}_k \tilde{G}_k^T) \right]^{\frac{1}{2}}.$$

Here $\nabla^k f$ denotes the vector of k-th order partial derivatives of f. G_1 and G_2 denote the first and second order sensitivity of $\mathbf{g}$ to Ψ. A high value of G_1 or G_2 indicates that $\mathbf{g}$

may not be stable for the given sensors and target location configuration, and can deviate from the actual target location significantly. Since the error in Ψ is assumed to be normal, then for regions of space where G_2 is small, $\mathbf{g}$ would be approximately normally distributed with mean $\mathbf{g}(\Psi_0, \Delta) + \epsilon_0$, where Ψ_0 denotes actual sensors locations. Here ϵ_0 depends on σ_ψ^2 (variance of Ψ) and the curvature of $\mathbf{g}$ around Ψ_0. This can be verified by considering the Taylor expansion of $\mathbf{g}$ around Ψ_0. Considering only sensor triplets for which $\mathbf{g}$ is smooth around Ψ_0 and Δ, we expect $\{\mathbf{g}(\Psi_i, \Delta)\}$ to form a relatively compact set of points around $\mathbf{g}(\Psi_0, \Delta)$. Figures 2 and 3 depict the bias and variance of $\mathbf{g}$ from the real target location for various, sensor configurations (assuming a normal error for sensors location estimates). Hotter colors indicate greater bias or variance. Note that with the increasing distance from the sensors, the bias and variance increase rapidly around the line connecting the sensors (figures 2 and 3). This is inline with the fact that G_1 and G_2 have significantly higher values near lines that connect any two of the sensors in a triplet (except the line segment that is between the sensors). G_1 and G_2 are monotonically increasing with respect to R, δ_1 and δ_2, where R is the distance from the centroid of the triangle formed by the sensors, and δ_1 and δ_2 are the differences in distances from the sensors.

As a result of the above discussions, for any specific network deployment topology, it should be checked if the region of interest is covered with enough number of sensor triplets with low bias and variance. In some sense the value of σ_Ψ^2 at the beginning of the network operation implicitly dictates the minimum network density. For instance, in case of a grid-like network (figure 4), where adjacent sensors in the same row or column have a distance of d, it can be easily verified that any event that resides within the network and has a detection range of $2d$ is covered by at least 13 to 24 sensor triplets. Thus, having a statistical knowledge of the type of acoustic events that take place in the target environment we can choose d appropriately.

Remember that at this stage we are not seeking to find the exact location of the target, but an initial guess which will be used to form the starting point for an optimization problem as will be described later. Thus to identify potential targets we search for compact sets of points, and take their means as the initial targets locations estimates. This can be seen as the mean of a number of independent normal random variables which ideally converges to the actual target location. More specifically, as $\sigma_\Psi^2 \to 0$, these estimates converge to the actual target locations.

Note that in the above clustering problem the following constraints should be applied. Any two points that their corresponding observation triplets contain two observations ν_j^ψ, ν_k^ψ such that $j \neq k$ then these two points connot reside in the same cluster. On the other hand if $j = k$ then the two points either reside in the same cluster or at least one of them is an outlier and cannot be contained in any cluster. Furthermore, if two points correspond to the same observation triplet then they cannot be in the same cluster. These constraints ensure that the

39
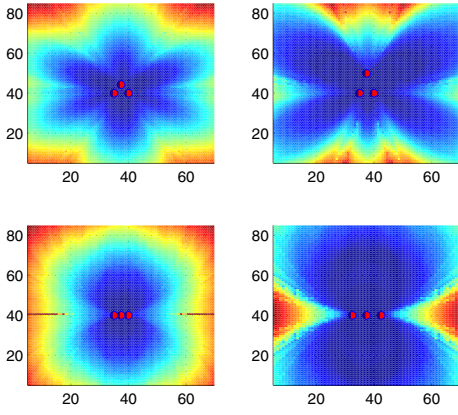
Fig. 2: Outline of bias maps for four sample three-sensor configurations. The map on the top left corresponds to three sensors that are 5 meters apart. In the top right map the top sensor has a distance of 10 meters from the connecting line of the other two sensors. In the bottom two maps, the sensors are collinear. In the bottom right map adjacent sensors have a distance of 5 meters. In the bottom left map this distance is 2.5 meters. All numbers in the above diagrams are in meters.
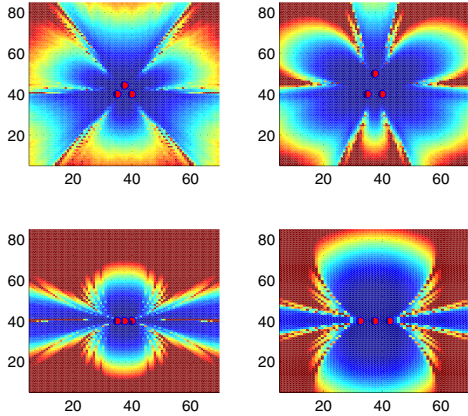

Fig. 3: Estimation variance maps of the corresponding sensor configurations in figure 2. All numbers in the above diagrams are in meters.

proposed partitioning is consistent as defined in equation 3.

Targets that do not reside in the smoothness region of enough number of sensor triplets might get missed in this initial stage. However as more and more observations become available, σ_Ψ^2 decreases and the network accuracy and coverage increases gradually.

In summary we should perform clustering on this set of points and associate each cluster with one of the targets. If all events are assumed to be global, then the number of clusters is already known to be the number of observations performed by a single sensor. Otherwise, determining the number of clusters is not trivial. In this work we used a grid-based clustering algorithm which is specifically adaptable for distributed memory computation. After the clusters have been elicited, we find an LSE matching between (a subset
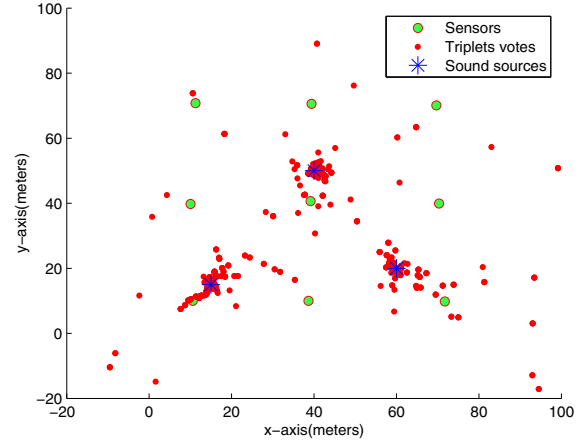
Fig. 4: Triplets generated from plausible combinations of sensors observations are mapped to a set of points in the space representing the network environment. These points form a set of clusters which together point out the location of sound sources.

of) these cluster centers and the predicted location of the previously discovered targets, that is $E[\Gamma_{t+1}^c | \Gamma_t^c]$ (section 2-C). The centers that remain unmatched are introduced as newly discovered targets, and their initial position is set to the corresponding cluster centers. Finally, $\tilde{\partial}_i$ is defined to comprise all observations that belong to a triplet in the i-th cluster (equation 3).

B. Algorithm Outline and Distribution Strategy

In a wide-area deployment of this sensor network, most events may not be heard by more than only a relatively small fraction of the sensors. Thus to facilitate parallelization of the algorithm We divided the sensor network into connected subnetworks where adjacent subnetworks have at least three sensors in common. Each subnetwork performs RAPAL algorithm independently and finally the subnets are aligned using conventional coordinate registration techniques. Furthermore, the data registration steps of the algorithm can be performed in a distributed manner. To this end, sensor ψ follows the following reactive rules in its operation:

1) Upon observing an event e_t, send a packet $\langle (\psi, e_t), ., . \rangle$ to all direct neighbors of ψ. Wait $\mathcal{T}$ seconds and then activate step 4. $\mathcal{T}$ is dependant on the network diameter.

2) Upon receiving a packet $\langle (\psi', e_{t'}), ., ., . \rangle$, wait until time $t' + \Delta_{\psi'\psi}^{\max}$, and then generate packets $\langle (\psi', e_{t'}), (\psi, e_t), . \rangle$ for any e_t that $0 < t - t' < \Delta_{\psi'\psi}^{\max}$. Send all generated packets to all direct neighbors excluding ψ'. If there is no such t, drop the received packet.

3) Upon receiving a packet $\langle (\psi', e_{t'}), (\psi'', e_{t''}), . \rangle$, wait until time $t' + \Delta_{\psi''\psi}^{\max}$, and then generate packets $\langle (\psi', e_{t'}), (\psi'', e_{t''}), (\psi, e_t) \rangle$ for any t such that $0 < t - t' < \Delta_{\psi''\psi}^{max}$. If there is no such e_t, drop the received packet. For any generated triplet calculate the intersection points of the corresponding hyperbolas and send each point to the sensor whose Voronoi cell contain the point.

40

4) Search for dense clusters of points in the points that have been received. If dense clusters are detected near the boundary between the Voronoi cells of two sensors, the detected clusters will be moved to the sensor that contains the larger part. After ℓ clusters have been detected solve the optimization problem 6 and then expunge the corresponding clusters.

In rules 1 and 2, sensors communicate their observations only with their direct neighbors (in the corresponding Delaunay triangulation). This is because we expect TDoA measurement error to be lower for sensor pairs that are closer to each other (section 3). Furthermore assuming that these neighboring sensors have direct radio communication simplifies the system time synchronization.

Calculation of $\Delta^{\max}$ depends on the actual sensors locations, and since they are not known exactly, to ensure that the time window covers the corresponding observation, we replace it with $[0 \quad \Delta^{\max} + \delta_e]$. Thus, according to Chebyshev inequality, with a probability of at least $1 - \frac{2\sigma_\Psi^2 + \zeta\sigma_T^2}{\zeta\delta_e^2}$ the corresponding time window covers the relevant observation.

C. Bayesian Filtering

In this section we describe how sensors observations are used to increase the accuracy of system state estimation through a Bayesian filter [13], [4]. Due to space limitation we have to omit the details of the calculations. Let $\mathbf{X}_t = [\Psi_t, \Gamma_t, \zeta_t]$ denote the state variable of the system at t-th step, where Ψ_t and Γ_t denote the sensors and targets locations, and ζ_t is the speed of propagation of sound in the network environment. Also, let $\Phi_t \triangleq p(X_t|Y_0,\ldots,Y_t)$ be the probability density function of the system state after having observed t batches of observations.

$$\Phi_t \propto p(Y_t|X_t) \int \Phi_{t-1}p(X_t|X_{t-1})dX_{t-1}$$

To efficiently estimate system state over time, we resort to a second-order approximation of Φ_t (based on Taylor expansion), which involves a point X_t^* and the Hessian matrix H at this point. We take X_t^* to be the most probable state of the system, given the observations and the current system state. At the t-th time step after receiving Y_t we perform a local search to find the X_t^* that maximizes $\Phi_t(X_t)$ Thus, the problem boils down to the following optimization problem,

$$\begin{aligned} X_t^* &= \arg\max_{X_t} \Phi_t(X_t) \\ &= \arg\max_{X_t} p(Y_t|X_t)p(X_t|Y_0,\ldots,Y_{t-1}) \end{aligned} \quad (5)$$

The two probability densities on the right-hand side of equation 5 denote the *observation* and *dynamics* models of the system respectively. Given the system state variable, sensor observations are independent. Thus

$$\log p(Y_t|X_t) = \sum_{\tilde{\mathfrak{d}}_e} \sum_{\nu_j^{\psi_i} \in \tilde{\mathfrak{d}}_e} \log p(\nu_j^{\psi_i}|\tau_e)$$

defining $p(\nu_j^{\psi_i}|\tau_e)$ directly in terms of observations and sensor positions is not straightforward. To circumvent this problem

we introduce an intermediate variable t_e as the ML estimator of the time at which the event e has occurred. t_e is defined as:

$$t_e(X_t,\tilde{\mathfrak{d}}_e) = \arg\min_{t_e} \sum_{\nu_j^{\psi_i} \in \tilde{\mathfrak{d}}_e} \left[\nu_j^{\psi_i} - t_e - \frac{\|\tau_e - \psi_i\|_2}{\zeta}\right]^2.$$

The observation model can finally be written as:

$$p(\nu_j^{\psi_i}|\tau_e) = t_e(X_t,\tilde{\mathfrak{d}}_e) + \frac{\|\tau_e - \psi_i\|_2}{\zeta} + \gamma_i,$$

where γ_i is a zero-mean Gaussian with variance σ_i^2. The dynamics model reflects our knowledge of how the system changes over time ($E[X_t|X_{t-1}]$). In this problem, it is assumed that sensors are stationary ($E[S_t|S_{t-1}] = S_{t-1}$). Also we generally do not expect ζ to change wildly and its dynamics can be modeled as $E[\zeta_t|\zeta_{t-1}] = \zeta_{t-1}$, with a small variance. Furthermore, we preserve those targets that have been matched with a cluster center and marginalize the rest of them. Unmatched cluster centers will also be added as new targets. We can simply model the targets as a set of drifting points with high variance or alternatively we can estimate the heading and velocity of moving targets by measuring the relative Doppler shift of the received acoustic waves among different sensors. Finally substituting the observation and the dynamics model into equation 5 results in the following optimization problem:

$$\begin{aligned} X_t^* = \arg\min_{X_t,t_l} \frac{1}{\sigma^2} \sum_l \sum_{\nu_j^i \in \tilde{\mathfrak{d}}_l} (t_l + \frac{\|\tau_l - \psi_i\|_2}{\zeta_t} - \nu_j^i)^2 \\ + (X_t - \mu)^T \Sigma (X_t - \mu), \end{aligned} \quad (6)$$

Here, Σ and μ are the predicted inverse covariance and mean of the system state, and can be directly calculated from the estimation of the previous step by marginalizing out targets and sensors that are not present in this step of the algorithm [1].

3. SIMULATION RESULTS

In this section we present a summary of the simulation results that demonstrate the correctness and feasibility of our algorithms. It is assumed that the network is coarsely initialized. That is,

$$X_0 = \begin{bmatrix} \Psi_A + \epsilon_\Psi \\ 0 \\ \zeta_A + \epsilon_\zeta \end{bmatrix},$$

where Ψ_A and ζ_A indicate actual sensors locations and actual speed of sound in the network environment. We model the initial error in sensors locations estimates ϵ_Ψ with a Gaussian noise with variance σ_Ψ^2. The error in measurement of the time difference of arrival is also modeled by a Gaussian noise with variance σ_T^2. The simulated network consists of 16 sensor nodes as depicted in figure 6. $'+'$ signs indicate true sensor locations and $'\times'$ signs indicate the final convergence values of the localization algorithm. The two anchors are depicted as black disks (due to the initial estimations, 2 anchors are sufficient instead of 3). The optimization is performed in batches of 10 events. In our simulations the error in sensors

		σ_T^2			
		0.001	0.0007	0.0005	0.0003
	4.0	1.076	0.811	0.782	0.461
σ_ψ^2	8.0	1.205	0.912	0.827	0.701
	10.0	1.602	1.388	1.106	0.832
	12.0	1.910	1.487	1.213	0.940

Fig. 5: SSE of sensors locations estimates for the 16-sensor network (figure 6) after observing 4 batches of events. TDoA and sensors locations are measured in seconds and meters respectively.
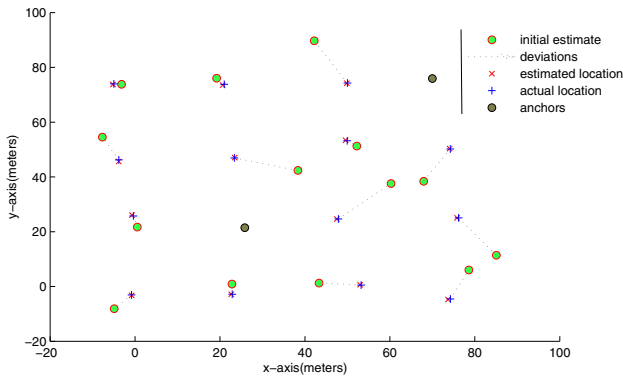
Fig. 6: Initial and actual sensors locations and the corresponding estimates after application of the first batch of 10 acoustic events.

locations estimates reduced to less than 0.1 meters with at most 4 batches. Figure 5 presents the sensors locations errors after 4 batches for various values of e_ψ^2 and σ_T^2.

As expected, simulation results indicate that TDoA measurement error has a considerable effect on the localization accuracy. Our preliminary experiments suggest that the accuracy of this measurement has a strong dependence on the bandwidth and center frequency of the observed sound wave. Generally, signals with higher bandwidth and lower frequency lead to more accurate TDoA measurements. Moreover, TDoA measurement error is directly proportional to the distance between sensors. Thus it can be helpful to use only those acoustic events for which TDoA measurement error is expected to be lower, and use only neighboring sensors for this purpose. This can be implemented as a preprocessing filtering step on the sensors. Furthermore, the simulations indicate that adaptation of the speed of sound ζ is rather difficult in presence of high TDoA noise. Thus a relatively close initialization of ζ can lead to significantly better approximations. This can be achieved by a few sensors that can measure relevant parameters (temperature and humidity) in various locations in the field to estimate the speed of sound.

It is known that multipath effect and reverberant environments can significantly contribute to the TDoA measurement error. In [7] it is shown that nonnegative deconvolution can be applied to TDoA measurement in a reverberant environment effectively.

4. CONCLUSION

In this paper we introduced a framework for passive acoustic localization and tracking using only acoustic sensors. An observation clustering phase is used for observation registration, and the results are used to sequentially update the system state variable, by means of a Bayesian filtering technique. The state variable of the system may contain sensors and targets localization parameters, as well as any other unknown parameters such as the speed of propagation of sound in the environment. This framework can passively use natural environmental acoustic phenomena to localize sensors. Experimental results of the proposed algorithms demonstrate the feasibility and correctness of the system.

REFERENCES

[1] David Roxbee Cox and Nanny Wermuth. *Multivariate Dependencies: Models, Analysis and Interpretation*. Chapman & Hall, 1996.

[2] L. Doherty, L. El Ghaoui, and K. S. J. Pister. Convex position estimation in wireless sensor networks. In *Infocom'01*, 2001.

[3] A. Galstyan, B. Krishnamachari, K. Lerman, and S. Pattem. Distributed online localization in sensor networks using a moving target. *Information Processing in Sensor Networks, 2004. IPSN 2004. Third International Symposium on*, pages 61–70, April 2004.

[4] Mohinder S. Grewal and Angus P. Andrews. *Kalman Filtering: Theory and Practice Using MATLAB*. John Wiley and Sons, Inc., 2001.

[5] X. Ji and H. Zha. Sensor positioning in wireless ad hoc networks using multidimensional scaling. In *Infocom'04*, 2004.

[6] H. Krim and M. Viberg. Two decades of array signal processing research: the parametric approach. *Signal Processing Magazine, IEEE*, 13(4):67–94, Jul 1996.

[7] Yuanqing Lin, Daniel D Lee, and Lawrence K Saul. Nonnegative deconvolution for time of arrival estimation. 2004.

[8] David Moore, John Leonard, Daniela Rus, and Seth Teller. Robust distributed network localization with noisy range measurements. In *SenSys '04: Proceedings of the 2nd international conference on Embedded networked sensor systems*, pages 50–61, New York, NY, USA, 2004. ACM.

[9] E. Olson, J. Leonard, and S. Teller. Robust range-only beacon localization. *Autonomous Underwater Vehicles, 2004 IEEE/OES*, pages 66–75, June 2004.

[10] P.N. Pathirana, N. Bulusu, A.V. Savkin, and S. Jha. Node localization using mobile robots in delay-tolerant sensor networks. *Mobile Computing, IEEE Transactions on*, 4(3):285–296, May-June 2005.

[11] Nissanka B. Priyantha, Hari Balakrishnan, Erik Demaine, and Seth Teller. Anchor-free distributed localization in sensor networks. Number TR-892, April 2003.

[12] Vikas C. Raykar, Igor Kozintsev, and Rainer Lienhart. Position calibration of audio sensors and actuators in a distributed computing platform. In *MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia*, pages 572–581, New York, NY, USA, 2003. ACM.

[13] Christopher J. Taylor. Simultaneous localization and tracking in wireless ad-hoc sensor networks. M.S. thesis, MIT university, 2006.

[14] Nishant Unnikrishnan, Probir Kumar Ray, Ajay Mahajan, and Tsuchin Chu. Accuracy considerations in a 3-d ultrasonic positioning system based on the difference in time of flights. *Journal of Dynamic Systems, Measurement, and Control*, 24(4):688–693, December 2002.

[15] Shang Yi, Ruml Wheeler, Zhang Ying, and Fromherz Markus. Localization from connectivity in sensor networks. *IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS*, 15:961–974, 2004.