Nondeterministic Sound Source Localization with Smartphones in Crowdsensing

Duc V. Le, Jacob W. Kamminga, Hans Scholten, and Paul J.M. Havinga
Pervasive Systems Group, University of Twente, Enschede, The Netherlands.
Email: {v.d.le, j.w.kamminga, hans.scholten, p.j.m.havinga}@utwente.nl

***Abstract*—The proliferation of smartphones nowadays has enabled many crowd assisted applications including audio-based sensing. In such applications, detected sound sources are meaningless without location information. However, it is challenging to localize sound sources accurately in a crowd using only microphones integrated in smartphones without existing infrastructures, such as dedicated microphone sensor systems. The main reason is that a smartphone is a nondeterministic platform that produces large and unpredictable variance in data measurements. Most existing localization methods are deterministic algorithms that are ill suited or cannot be applied to sound source localization using only smartphones. In this paper, we propose a distributed localization scheme using nondeterministic algorithms. We use the multiple possible outcomes of nondeterministic algorithms to weed out the effect of outliers in data measurements and improve the accuracy of sound localization. We then proposed to optimize the cost function using least absolute deviations rather than ordinary least squares to lessen the influence of the outliers. To evaluate our proposal, we conduct a testbed experiment with a set of 16 Android devices and 9 sound sources. The experiment results show that our nondeterministic localization algorithm achieves a root mean square error (RMSE) of 1.19 m, which is close to the Cramer-Rao bound (0.8 m). Meanwhile, the best RMSE of compared deterministic algorithms is 2.64 m.**

I. Introduction

Sound localization has been used for decades using a number of microphones at known locations and measuring the time difference of arrival (TDOA). Although sound localization is widely applied in many applications such as sniper localization, intruder detection, vehicle tracking, event detection, and rescue operations [1], the problem of using only on-board microphones of smartphones, without support from a dedicated sensing systems has received comparatively less attention, especially with Android smartphone-based platforms that share more than 80% smartphone market in 2015. The main challenge is that Android devices have noticeable uncertainties in time synchronization and signal processing latency. For example, the standard deviation of the TDOA we measured on a set of 16 Nexus-7 tablets is approximately 15 ms. In other words, the TDOA ranging has an error of 15×0.34029 m/ms $= 5.1$ m, where 0.34029 m/ms is the speed of sound in air. Meanwhile, that of iPhone smartphones is roughly 1 ms or 0.34 m. In summary, the nondeterminism challenges smartphone-based sound localization systems.

In this paper we study the nondeterministic characteristic of smartphones and solve the sound localization problem using only a set of smartphones and their TDOA measurements, without the help of an extra infrastructure or dedicated sensing systems. In particular, we propose a Distributed variant of the RANdom SAmple Consensus (DRANSAC) scheme to deal with the effects of nondeterministic TDOA measurements on localization accuracy. Our testbed experiment results show that, when using the Levenberg-Marquardt method (LM) [2] in the DRANSAC scheme to solve the least absolute deviations (LAD) cost function [3], the nondetermnistic algorithm significantly improves the root mean square error (RMSE), two times better the deterministic version of the algorithm (1.19 m vs. 2.64 m). An application example is event detection within crowds where people often carrying their smartphones and there is no dedicated sound localization systems deployed.

As a nondeterministic platform, smartphones produce considerable and unpredictable uncertainties in the TDOA measurements, mostly outliers. In general, a random sample consensus (RANSAC) scheme [4] can be used to weed out outliers in erroneously filtered estimated parameters, which are unknown locations of sounds in our problem. Such conventional RANSAC schemes require pre-calibration to optimize RANSAC parameters, such as the number of iterations, the threshold of the number of inliers, and the threshold of the distances between points and the fitting line. Although the calibration is straightforwards, it is ill suited to dynamic environments. Therefore, we take advantage of the proliferation of smartphones to propose a distributed variant of RANSAC as most people always carry at least one smartphone wherever they go. This DRANSAC-like scheme distributes iterations to neighboring smartphones to perform decentralized optimization simultaneously using a deterministic fitting model. This approach eliminates the pre-calibration phase, and allows the scheme to run faster since it use a deterministic model to fit estimated outcomes rather than the RANSAC-based fitting. In addition, we propose to formulate the localization problem using the LAD cost function to lessen the influence of outliers in the TDOA measurements on the estimates.

The paper is organised as follows. After describing related work in Section II, we formulate the problem in Section III. Section IV describes nondeterministic factors that effect the TDOA measurements measured by smartphones. Our proposed DRANSAC scheme for sound localization is presented in Section V and evaluated in Section VI. Finally, we conclude this work with Section VII.

II. Related Work

Range-based localization with known locations of sensors can be applied on either the received signal strength [5], the time of arrival (TOA) [6], the TDOA, or a hybrid [7]. Of such ranging measurements, the TDOA is the most suited and common for natural sound source localization, in which the time of sound emission is unknown. Since the rise of smartphones, researchers are interested in using available microphones on smartphones to perform sound localization. However, smartphones are a nondeterministic platform that produces large and unpredictable outliers in measurements. In a recent work Shang *et al.* [8] demonstrate a method that localizes a sound source with 4 Nexus-One smartphones running Android 2.3.4. The localization error is within ± 15 cm in the x-direction and ± 80 cm in the y-direction when the synchronization and start-time offsets are removed. However, [8] is not evaluated in crowd scenarios, which includes abundant smartphones. Moreover, they solve the problem by using a deterministic approach based on intersections of the TDOA hyperbolic curves [9]. Theoretically proven, this estimator cannot outperform the maximum likelihood estimation (MLE) optimized by the LM method that we used to compare to our proposal in our testbed experiment.

Very recently, Burgess et al. [10] constructed a RANSAC-like scheme [4] to simultaneously solve the calibration problem and remove severe anomalies, which is a common problem in the TOA localization. Their scheme is centralized and specificly designed to use their minimal solver for the TOA calibration problem to localize sounds that are far away from smartphones, albeit their RANSAC-like scheme is not their main contribution. In addition, the real experiment in [10] was conducted with a very low-noise setup using dedicated microphones and speakers. Therefore, their system can achieve a RMSE of 2.35 cm. To the best of our knowledge, there is no existing work that has performed natural-sound localization in a noisy environment using only nondeterministic smartphones such as Android devices with a market share of 81% in 2015 carrying forward until 2019. In addition, smartphones subscriptions are increasing tremendously, probably 1436.5 millions smartphones will be shipped worldwide by the end of 2015 [11]. To this end, we propose the distributed RANSAC that can overcome nondeterministic effects by using smartphone collaboration. Last but not least, we also propose to formulate the localization problem using the LAD to avoid emphasizing anomalies in measurements.

III. Problem Formulation

In general, we consider a network comprising a set $\boldsymbol{S} = \{\boldsymbol{x}(k) \ : \ k = 1, \ldots, n\}$ of n acoustic emitters with unknown location information, named source nodes, and a set $\boldsymbol{A} = \{\boldsymbol{x}(k) \ : \ k = n+1, \ldots, n+m\}$ of m acoustic sensor devices such as smartphones with known location information, named anchor nodes. To idealise the localization problem, we assume that all the source nodes and anchor nodes are stationary on a 2-dimensional plane for a few seconds that is long enough to perform the distributed localization with smarphones. Localization with mobility and in higher space is straightforward. We also assume that the size of each node is small enough to be treated as a point on the localization map, with has the coordinate $\boldsymbol{x} = (x, y)^T \in \mathbb{R}^2$. Each source node generates a limited-power acoustic signal that can be used to estimate the distance to near receivers. Let $\boldsymbol{N}(k)$ denote the set of anchor nodes in range of the k^{th} node. Clearly, $\boldsymbol{x}_k \notin \boldsymbol{N}(k)$ and $\boldsymbol{N}(k) \subset \boldsymbol{A}$.

The localization problem is to estimate the vector positions of the source nodes, the unknown parameter $\Theta = \{\theta(k) = (\hat{x}_k, \hat{y}_k)^T, \ k = 1, \ldots, n\}$, given the vector positions of the anchor nodes in $\boldsymbol{N}(k)$ and the TDOA measurements $\{\delta_{i,j}(k) : i, \ j \in \boldsymbol{N}(k), \ j \neq i\}$. In a 2-dimensional plane we have $\delta_{i,j}(k) = c(t_j(k) - t_i(k))$, where c is the speed of sound (≈ 0.34029 m/ms), $t_i(k)$ and $t_j(k)$ are the TDOA values of the sound source k when the signal arrives anchor nodes i and j, respectively.

Without loss of generality we assume that all measurements $\{\delta_{i,j}(k) : \ i, \ j \in \boldsymbol{N}(k), \ j \neq i\}$ between a blind node $\boldsymbol{x}(k)$ and its neighbouring anchor nodes in $\boldsymbol{N}(k)$ are available. We also assume that the distribution of $\delta_{i,j}(k)$ is Gaussian. In addition, we assume that $\{\delta_{i,j}(k)\}$ are statistically independent. This assumption somewhat oversimplifies the practical environment but it is necessary for analysis. Finally, we assume that there are an adequate number of anchor nodes, for example, at least three anchor nodes in any set of $\boldsymbol{N}(k)$ in 2-dimensional space.

In a real-world application, a source node can be either natural or synthesised. Some kinds of source nodes might have microphones to receive the acoustic signal as well, for example, smartphones that emit sounds. Therefore, $\boldsymbol{N}(k)$ might consist of not only anchor nodes but also source nodes that can receive the acoustic signal from the *k*th node. Since most of the modern devices possess accurate positioning systems such as the acoustic localization system proposed by Liu *et al.* [12] could achieve an accuracy of 23 cm, and our aim is sound source localization in crowds, we will focus on localizing natural sound sources in this work.

IV. Nondeterministic Platforms

Unlike dedicated sensing devices that have been used for decades in sound source localization, off-the-shelf smartphones are nondeterministic since they are in situ designed for communication, not for sensing. Especially, smartphones with the Android operating systems are well-known for their tedious inaccuracy in audio applications due to a large and unpredictable round-trip audio latency.

A. Sensing Latency

Fig. 1 illustrates error in distance measurement based on acoustic sensors. For indeterministic sensing platforms, the delay between a sound wave arriving at the microphone and having the digital sound information ready in a buffer, can be hundreds of milliseconds. In other words, the estimated distance from the sound source to the destination contains

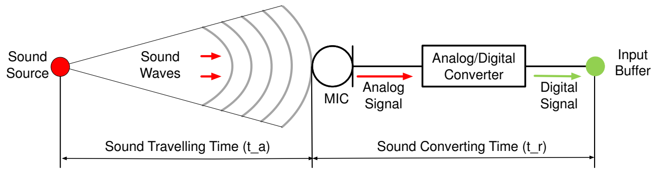


Fig. 1. Sensing latency in measuring travel time of sounds. That converting time is usually unknown and unstable in generic sensors devices results in unacceptable error in the measurements.

an error of dozens of meters. This is unacceptable for most applications such as acoustic-based localisation.

As we assumed that we only can get timestamps when the data are loaded into the output buffer and when the data are available in the input buffer, the latency we can measure, denoted as τ_s, includes the signal travelling latency t_a and the receiving latency t_r, $\tau_s = t_a + t_r$.

B. Clock Synchronization Errror

When a global positioning system (GPS) radio is connected to multiple satellites the device can acquire very accurate time synchronisation with the satellites. This accurate time t_g is communicated through so called the National Marine Electronics Association (NMEA) messages and has a resolution of nano seconds.

The NMEA messages are communicated to the Android framework from the GPS chip through several layers, this causes an offset between the timestamp stored inside the message and the moment the message received in the Android framework. This offset also varies and introduces jitter in the offset. The user application, that handles the received NMEA message in a callback, also suffers from latency. The user application handles the callback after a small delay, as Android is not a real-time OS.

The Android framework provides a time-stamp t_s along with the received NMEA message based on the system clock. This time-stamp is the device's system time at the moment the Android framework received the message. The experimental application calculates the offset between the provided time-stamp in the callback and the GPS-time, $\tau_c = t_s - t_g$. This offset is the parsed time from the NMEA message minus the time-stamp that is provided in the callback.

In a very similar way, we studied that clock synchronization using the Network Time Protocol (NTP) also faces an offset problem, the offset between the system clock and NTP server clock.

C. Empirical Evaluation

To evaluate the theoretical analysis, we developed an Android application that can measure the clock synchronization as well as the audio latency. The results of experiments with a set of 16 Nexus-7 tablets show that the variance caused by clocks even after being calibrated with GPS is 68.98 ms as shown in Fig. 2(a). If we consider the effect of audio latency, the variance of TDOA ranging is even larger. Fig. 2(b) shows the error variance effected by both clock synchronization (τ_c) and sensing latency (τ_s), which is up to 213.16 ms. In other words, the error in TDOA ranging can have a standard deviation of 0.34029 m/ms$\times 213.16^{1/2} = 4.9$ meters, where 0.34029 m/ms is the speed of sound. Such error is significantly larger than that of dedicated TDOA ranging devices, which is typically only several centimeters. Fortunately, from the histograms plotted in Fig. 2, we can observe that the errors of audio latency and clock synchronization have the distributions that are quite close to a normal distribution. Therefore, our assumption that the distribution of $\delta_{i,j}(k)$ is Gaussian still holds.

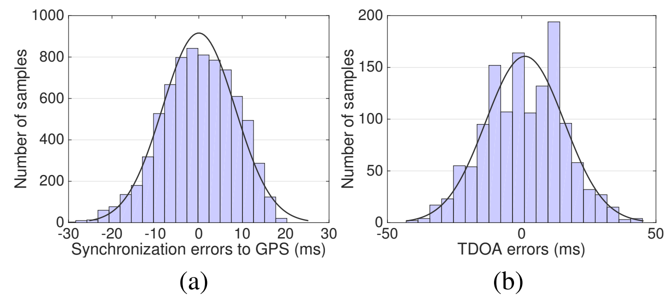


Fig. 2. Distribution of error: (a) clocks synchronized with GPS, (b) TDOA error effected by both clock synchornization and audio latency.

V. Nondeterministic Localization

In this section, we first introduce the popular approach for sound localization, which is linearizing the problem as a least-squares problem. However, the least-squares approach faces a big problem with outliers due to nondeterministic sensing platforms. Therefore, we propose to solve the localization problem by using nondeterministic approaches based on a least absolute deviation problem. Finally we present our proposed distributed nondeterministic scheme using RANSAC-like scheme with the LAD cost function and an area-fitting model.

A. Time Difference of Arrival Variables

The TDOA is used in the case where the time of sound emission is inavailable. The TDOA measurements can be obtained by computing the difference between the TOA of source node k and its anchor nodes as follows:

$$\boldsymbol{\delta}(k) = \begin{pmatrix} \delta_{2,1}(k) \\ \delta_{3,1}(k) \\ \vdots \\ \delta_{|\boldsymbol{N}(k)|,1}(k) \end{pmatrix}, \tag{1}$$

where $\delta_{i,j}(k)$ is the TDOA measurement defined in Section III. Since the TDOA measurement $\delta_{i,j}(k)$ is computed by multiplying the time-of-flight and the speed of sound, $\boldsymbol{\delta}(k)$ is a vector of Euclidean distances, which has the length depending on how many anchor nodes in range. Note that the TDOA vectors $\{\boldsymbol{\delta}(k)\}$ may have different lengths.

TABLE I
COMPARING STATISTICAL OPTIMALITY CRITERIA

Criteria	Least Squares	Least Absolute Deviations
Robust	✗	✓
Stable	✓	✗
Multiple outcomes	✗	✓

B. Least Squares Problem

Typically, a localization problem can be formulated as a least-squares problem so that it can be optimized by using a least-squares optimization, such as the LM method [2] or Trust-Region-Reflective Algorithm [13]. Let $J(\boldsymbol{x}(k))$ denote the Jacobian (cost) function of the location estimation for source node k , we have

$$J(\boldsymbol{x}(k)) = \sum_{i=2}^{|\boldsymbol{N}(k)|} \Big(\|\boldsymbol{x}(k)-\boldsymbol{x}(i)\|-\|\boldsymbol{x}(k)-\boldsymbol{x}(1)\|-\delta_{i,j}(k)\Big)^2, \tag{2}$$

where $\|.\|$ is the Euclidean norm and $|\boldsymbol{N}(k)|$ is the size of $\boldsymbol{N}(k)$. Recall that $|\boldsymbol{N}(k)| \geq 3$.

C. Least Absolute Deviations

As the ordinary least squares (OLS) problem formulation amplifies the outliers of measurement by squaring the residuals, we propose to use the LAD instead. The Jacobian function then becomes

$$J(\boldsymbol{x}(k)) = \sum_{i=2}^{|\boldsymbol{N}(k)|} \Big|\|\boldsymbol{x}(k)-\boldsymbol{x}(i)\|-\|\boldsymbol{x}(k)-\boldsymbol{x}(1)\|-\delta_{i,j}(k)\Big|, \tag{3}$$

where $|.|$ is the absolute value.

Table I may help to understand why we choose LAD as the criterion for our localization optimization. The LAD method is more robust compared to the least squares method as LAD is resistant to outliers in the data, which are very common when measured with nondeterministic smartphone-based platforms. In addition, the instability property of the LAD method makes it self-adaptive to the change of the random parameter in nondeterministic approaches, for instance, the random subset of anchors discussed in Section V-D is used to iteratively estimate the position of source nodes. A small change in the configuration of input data may produce a significant difference in the optimization results. Finally, the LAD is also a better choice since it can provide multiple outcomes that offer an opportunity to select the best solution.

D. Distributed Nondeterministic Estimation

Since it would be too costly to obtain high localization accuracy by an exact solution, we propose to use nondeterministic approaches to find an approximation to a solution. In general, a nondeterministic algorithm is an algorithm that may provide different results on different runs, even for the same given input. To solve the sound source localization problem, we address a probabilistic algorithm of which behaviors rely on a random number generator. The random number can be the number of devices selected to perform distributed localization. Every estimated location the nondeterministic algorithm computes is valid, regardless of which set of devices the algorithm chose while running.

In particular, we propose to optimise the LAD problem described in Section V-C by using an iterative algorithm such as the LM method on a random subset. The bottom line of the idea is to minimize the effect of measurement outliers caused by the nondeterministic behaviors. The run of the algorithm with a subset that contains some anomalies will result in a large residual, which can be filtered out later by some outlier detection, such as Online Frequency Likelihood Estimation (FLEAD) [14], [15].

The localization algorithm is designed in a distributed and cooperative manner. In other words, for each source node k, the localization computation tasks are split and assigned for selected anchor nodes, which are smartphones in $\boldsymbol{N}(k)$. Then such anchors nodes combined their estimated models to enhance the estimated location of the source node. More details of the algorithm are presented in Algorithm 1. For implementation clarity, the algorithm is expressed with regard to anchor node i and source node k.

Algorithm 1 <Cooperative Nondeterministic Localization>

1: INPUT: $\boldsymbol{N}(k)$, $\boldsymbol{\delta}(k), \mathcal{A}$
2: OUTPUT: $\hat{\boldsymbol{x}}(k)$
3: Initialize subset size $p_i \sim \mathcal{U}(3, |\boldsymbol{N}(k)|)$, $p_i \in \mathbb{N}$
4: Compute number of run loops $l_i \leftarrow f\left(p_i, |\boldsymbol{N}(k)|\right)$, $l_i \in \mathbb{N}$
5: **for** $j = 1 : l_i$ **do**
6: Select a random subset $\boldsymbol{C}_j \leftarrow C\left(|\boldsymbol{N}(k)|, p_i\right)$
7: Minimize the Jacobian function $J\left(\boldsymbol{x}(k)\right)$ given $\boldsymbol{C}_j$
8: **if** $\hat{\boldsymbol{x}}(k) \notin \mathcal{A}$ **then**
9: Discard new estimated position $\hat{\boldsymbol{x}}(k)$
10: **else**
11: Add new estimated position $\hat{\boldsymbol{X}}_i(k) \leftarrow \hat{\boldsymbol{x}}(k)$
12: **end if**
13: **end for**
14: **for** $j = 1 : |\boldsymbol{N}(k)|$ **do**
15: Augment estimated positions, $\hat{\boldsymbol{X}}(k) \leftarrow \hat{\boldsymbol{X}}_j(k)$
16: **end for**
17: Estimate the source location $\hat{\boldsymbol{x}}(k) \leftarrow \hat{\boldsymbol{X}}(k)$
<end>

To run Algorithm 1 on anchor node i, it requires that smartphone i has to connect to other smartphones in neighbourhood $\boldsymbol{N}(k)$ to obtain the TDOA measurements $\boldsymbol{\delta}(k)$, which are correlated to a same sound source k, and estimate the area $\mathcal{A}$ bounded by anchors in $\boldsymbol{N}(k)$. The random size of the subset p_i, which will be used to localize source node k, is drawn from a discrete uniform distribution with lower and upper endpoints specified by 3 and $|\boldsymbol{N}(k)|$, respectively. Note that the minimum number of anchor nodes to perform $2D$ localization is 3. Line 3 in Algorithm 1 determines the number of loops l_i to perform location estimation, $l_i = f\left(p_i, |\boldsymbol{N}(k)|\right)$ given by (6). For each loop, a random subset $\boldsymbol{C}_j$ is selected without replacement from all possible combinations of anchor nodes in $\boldsymbol{N}(k)$ taken p_i at a time. Without replacement means that if a subset is already selected, it will not be selected again in the next loop. Given the TDOA measurements $\boldsymbol{\delta}(k)$ and the anchor nodes in $\boldsymbol{C}_j(k)$, the coordinates of source node k are estimated by applying an estimator such as the LM on the non-linear cost function $J(\boldsymbol{x}(k))$ described in (2) or (3). If an estimated position $\hat{\boldsymbol{x}}(k)$ does not fall in area $\mathcal{A}$, it will be discarded. Note that other approaches use a RANSAC-based outlier detection instead of the area-based constraints like us. Combining the valid estimated coordinates of source node k provided by other smartphones in the group, we obtain set $\hat{\boldsymbol{X}}(k)$ of which elements are considered as samples of a distribution. Finally, the expected value of the coordinate distribution is computed through lines 12 and 13, which is also the final estimated location of source node $|\boldsymbol{N}(k)|$.

In particular, the amount of loops l_i is computed based on the subset size p_i and the neighbor-set size. A small subset will result in a higher variation of the estimated locations, and thus needs more loops to determine the inliers given by:

$$l_i = \left\lfloor \beta \frac{|\boldsymbol{N}(k)|}{p_i} \right\rceil, \tag{4}$$

where $\lfloor . \rceil$ is the nearest integer and β is the *loop factor*. The loop factor is the minimum number of loops that are required for a single anchor node (smartphone) to run all possible subsets of the smallest combination, when $p_i = |\boldsymbol{N}(k)| - 1$. We have,

$$\begin{aligned} \min_{p_i} C^{p_i}_{|\boldsymbol{N}(k)|} &= C^{|\boldsymbol{N}(k)|-1}_{|\boldsymbol{N}(k)|} = \frac{|\boldsymbol{N}(k)|!}{(|\boldsymbol{N}(k)|-1)!\ 1!} \\ &= |\boldsymbol{N}(k)|\ \forall p_i \in [3, |\boldsymbol{N}(k)|]\,. \end{aligned} \tag{5}$$

Therefore, the number of loops can be rewritten as follows:

$$l_i = \left\lfloor \frac{|\boldsymbol{N}(k)|(|\boldsymbol{N}(k)|-1)}{p_i} \right\rceil. \tag{6}$$

Equation 6 tells that in case of smallest combination, $p_i = |\boldsymbol{N}(k)| - 1$, a single anchor node needs to run all the possible subsets, $l_i = |\boldsymbol{N}(k)|$.

VI. Testbed Experiment

In this section, we discuss a testbed that was conducted to evaluate our proposed cooperative nondeterministic approach for sound source localization. We demonstrate our scheme with one of the best MLE methods, the LM method [2]. The LM method is used to solve non-linear least squares problems by interpolating between the Gauss-Newton algorithm and the method of gradient descent. If an iteration gives insufficient reduction in the residual, the damping factor can be increased to bring the LM closer to the gradient descent, whereas if an iteration reduces the residual enormously, the damping factor can be decreased to bring the LM closer to the Gauss-Newton algorithm. The results shows that nondeterministic approaches provide localization accuracy that is significantly better than the standard variants and very close to the Cramer-Rao Lower bound (CRB), especially when there are many smartphones participating in localization. The CRB is a benchmark that has been popular to evaluate localization algorithms [16]. It provides a lower bound of estimators given measurements containing Gaussian noises. The CRB can be computed from an inverse Fisher Information Matrix in information theory given the Gaussian parameters of the noises.

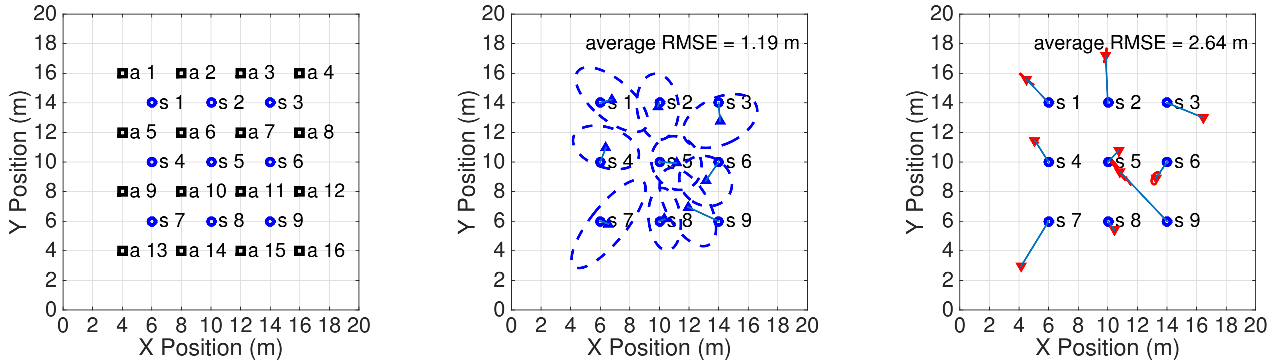
A. Experiment Setup

In order to conduct the experiment, we developed a localization application and installed it on 16 Nexus-7 tablets. In particular, there are 16 anchor nodes marked as "□", Nexus-7 tablets installed with Android 4.4.4 (KitKat), which were deployed over an outdoor area with a pairwise distance of 4 m as illustrated in Fig. 3(a). From our experience working with smartphone sensor networks, this experiment setup is very similar to daily circumstances such as in a class room, in a conference, and on a walking-street. Since we address the environmental sounds within a crowd, sound sources were sequently generated at 9 positions marked as "○" in Fig. 3(a), five successive sounds per each position. The aim of conducting multiple measurements is to increase possible position estimates. Although the positions of the tablets are known and fixed in our testbed, we placed them next to the absolute reference points with a random error of 20 cm. This random placement intends to emulate the random errors of the device positions obtained by an off-the-shelf systems such as the acoustic localization system proposed by Liu *et al.* [12].

B. Experimental Results

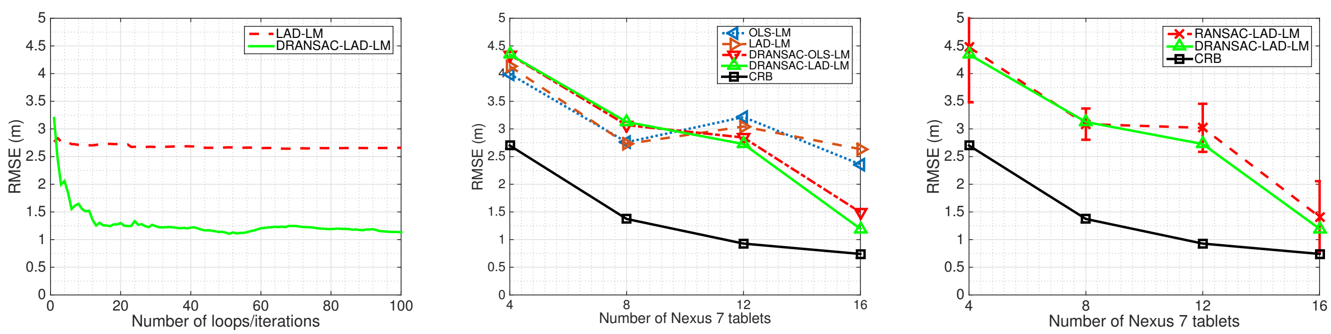
Fig. 3(b) shows sound sources and their means of estimated locations for with DRANSAC-LAD-LM using 16 Nexus-7 tablets, which are marked as "○" and "△", respectively. The dash eclipses illustrate the bounds of possible outcomes that vary around the mean of estimated locations. The nondeterministic algorithm has a RMSE of 1.19 m, which is amazingly good since the standard deviation of TDOA measurements is approximately 5 m. Especially, the localization error can achieve very good at some locations such as an error of 28 cm for node number 2. It is consistent with our theoretical analysis, DRANSAC-like localization scheme efficiently deals with outliers of measurements when combining LAD and LM. In particular, our algorithm iteratively filters out outliers by assuming that there exists subsets of inliers from measurements. These inliers can be used to estimate the location of sounds. Sound node number 9 has the worst error since it produces a lot of outliers in our measurement data. Conducting more measurements would help improving the accuracy.

On the other hand, Fig. 3(c) shows sound sources and their mean-estimates of sound locations with the deterministic approach LAD-LM using 16 Nexus-7 tablets, which are marked as "○" and "▽", respectively. As being consistent with theoretical analysis, the deterministic algorithm mostly returns a constant estimated location for every source, of which RMSE is 2.64 m. Note that both DRANSAC-LAD-LM and LAD-LM use a same measurement inputs, of which outliers are removed based on the mean and standard deviation of the TDOA measurements. That is to say DRANSAC-LAD-LM



(a) (b) (c)

Fig. 3. Testbed experiment: **(a)** 16 Nexus-7 tablets and 9 sound sources marked as "□" and "○", respectively; **(b)** estimated locations provided by the nondeterministic algorithm marked as "△"; **(c)** estimated locations provided by the deterministic algorithm marked as "▽". Dashed ellipses represent bounds of possible location estimates.



(a) (b) (c)

Fig. 4. Testbed experiment: **(a)** Convergence of nondeterministic DRANSAC-LAD-LM vs. deterministic LAD-LM; **(b)** RMSE when varying the total number of Nexus-7 tablets; **(c)** Comparing RANSAC-like schemes with 100 runs. Distributed variant provides smaller std. of final estimates.

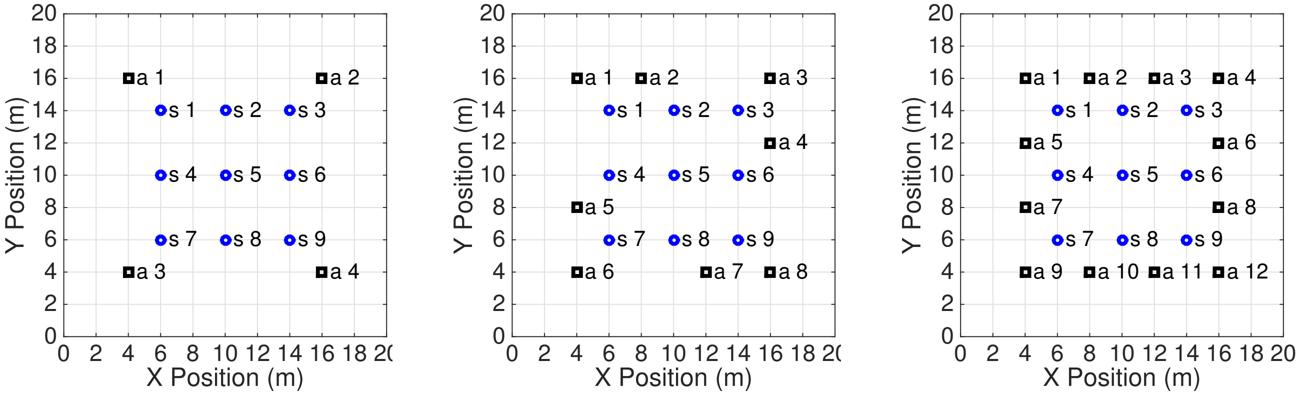
can remove further outliers from the TDOA measurements through estimates whereas LAD-LM cannot. As a set of pairwise measurements may contain chiefly outliers due to a channel shadowing, an outlier detection algorithm will remove inliers, but outliers. This explains why LAD-LM still has the poor accuracy even if applying outlier detection on the inputs. However, the effects of outliers can be revealed after using them to estimate the unknown location. The estimated location would be unrealistic and easier to be filtered with a constraint, for example, with the area formed by the locations of anchor nodes.

It is interesting to study the computation load of nondeterministic localization through how fast a nondeterministic localization converges to a stable accuracy. Fig. 4(a) shows the convergence speeds of LAD-LM (deterministic localization algorithm) and DRANSAC-LAD-LM (nondeterministic localization algorithm). Surprisingly the nondeterministic algorithm converges very quickly to its stable estimates (RMSE$\approx$ 1.2 m) just after around 20 loops. Note that the deterministic version also requires similar number of loops to converge to its best estimates (RMSE$\approx$ 2.7 m).

Varying the total number of anchor nodes (Nexus-7 tablets), we investigate the performance of the LM technique when being used in different localization schemes. Overall, the RMSE of the algorithms decreases when increasing the number of the devices as shown in Fig. 4(b). In particular, the LM using the DRANSAC-like scheme shows its advantage when having sufficient anchor number ($>$ 12 anchors) that would enable efficiently iterating the subsets of anchors to remove outliers. Fig. 4(b) also shows that our proposed DRANSAC-LAD-LM can achieve closely to the CRB when using 16 Nexus-7 tablets.

In addition, we compare the conventional RANSAC-like scheme to our distributed version that uses the Levenberg-Marquardt method to optimize the LAD cost function. The results plotted in Fig. 4(c) show that our distributed version is more reliable than the conventional one. The reason is that DRANSAC uses the deterministic model fitting, based on estimated areas and mean value. Conversely, the conventional RANSAC uses line fitting on a subset of estimated positions, which is nondeterministic. Although we did fine tune parameters of RANSAC line fitting, the RMSEs of RANSAC are still higher than our DRANSAC, which does not require tuning parameters.

Last but not least, we emphasise that results used to plot graphs in Fig. 4(b) and 4(c) are optimally obtained from all combinations from the collection of 16 Nexus-7 tablets'



(a) Optimal 4-anchor placement. (b) Optimal 8-anchor placement. (c) Optimal 12-anchor placement.

Fig. 5. Optimal placement of anchor subsets obtained by trying all possible combinations with the testbed experiment. Nexus-7 tablets and 9 sound sources marked as "□" and "○", respectively.

positions. The corresponding optimal combinations of four, eight, and twelve Nexus-7 tablets are illustrated in Fig. 5(a), 5(b), and 5(c), respectively. Fig. 5 indicates that anchors should be placed symmetrically and towards the bound of the areas of interests in order to obtain the highest localization accuracy.

VII. Conclusion

In this paper we have pointed out that localizing natural sound sources using smartphones without support infrastructures is a daunting challenge because smartphones are nondeterministic and produce large and unpredictable errors in the TDOA measurements. As most conventional least-squares approaches are ill suited to deal with large outliers in the TDOA measurements, we proposed the DRANSAC scheme, a distributed nondeterministic localization based on the random sample consensus. In addition, we proposed using least absolute deviations rather than least squares as a cost function to avoid emphasising the outliers. The experiment results with 16 Nexus-7 tablets show that our nondeterministic algorithm gives a RMSE of 1.19 m, which is close to the Cramer-Rao bound. The RMSE of DRANSAC-LAD-LM is two times better than using its deterministic version, which was considered as an optimal estimator for non-linear problems by most existing work.

Acknowledgment

This work is supported by the SenSafety project in the Dutch Commit program, www.sensafety.nl.

References

[1] P. Transfeld, U. Martens, H. Binder, T. Schypior, and T. Fingscheidt, "Acoustic event source localization for surveillance in reverberant environments supported by an event onset detection," in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 2629–2633.

[2] J. J. Moré, "The levenberg-marquardt algorithm: implementation and theory," in *Numerical analysis*. Springer, 1978, pp. 105–116.

[3] P. Bloomfield and W. Steiger, "Least absolute deviations, theory, applications," *Algorithms*, 1983.

[4] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[5] V.-D. Le, V.-H. Dang, S. Lee, and S.-H. Lee, "Distributed localization in wireless sensor networks based on force-vectors," in *Intelligent Sensors, Sensor Networks and Information Processing, 2008. ISSNIP 2008. International Conference on*. IEEE, 2008, pp. 31–36.

[6] W. van Kleunen, N. Meratnia, and P. J. Havinga, "als-coop-loc: cooperative combined localization and time-synchronization in underwater acoustic networks," in *Proceedings of the 4th ACM SIGBED International Workshop on Design, Modeling, and Evaluation of Cyber-Physical Systems*. ACM, 2014, pp. 23–26.

[7] W. van Kleunen, K. C. Blom, N. Meratnia, A. B. Kokkeler, P. J. Havinga, G. J. Smit *et al.*, "Underwater localization by combining time-of-flight and direction-of-arrival," in *OCEANS 2014-TAIPEI*. IEEE, 2014, pp. 1–6.

[8] Y. Shang, W. Zeng, D. K. Ho, D. Wang, Q. Wang, Y. Wang, T. Zhuang, A. Lobzhanidze, and L. Rui, "Nest: Networked smartphones for target localization," in *Consumer Communications and Networking Conference (CCNC), 2012 IEEE*. IEEE, 2012, pp. 732–736.

[9] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *Signal Processing, IEEE Transactions on*, vol. 42, no. 8, pp. 1905–1915, 1994.

[10] S. Burgess, Y. Kuang, and K. Åström, "Toa sensor network self-calibration for receiver and transmitter spaces with difference in dimension," *Signal Processing*, vol. 107, pp. 33–42, 2015.

[11] K. Nagamine, "Idc worldwide quarterly mobile phone tracker," http://www.idc.com/getdoc.jsp?containerId=prUS25860315, August 2015.

[12] K. Liu, X. Liu, L. Xie, and X. Li, "Towards accurate acoustic localization on a smartphone," in *INFOCOM, 2013 Proceedings IEEE*, April 2013, pp. 495–499.

[13] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM Journal on optimization*, vol. 6, no. 2, pp. 418–445, 1996.

[14] V. D. Le, H. Scholten, and P. Havinga, "Flead: Online frequency likelihood estimation anomaly detection for mobile sensing," in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM, 2013, pp. 1159–1166.

[15] V.-D. Le, H. Scholten, and P. M. Havinga, "Online change detection for energy-efficient mobile crowdsensing," in *Mobile Web Information Systems*. Springer, 2014, pp. 1–16.

[16] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero III, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *Signal Processing Magazine, IEEE*, vol. 22, no. 4, pp. 54–69, 2005.