High-Accuracy TDOA-Based Localization without Time Synchronization

Bin Xu, *Member, IEEE*, Guodong Sun, *Member, IEEE*,
Ran Yu, *Student Member, IEEE*, and Zheng Yang, *Student Member, IEEE*

Abstract—Localization is of great importance in mobile and wireless network applications. Time Difference of Arrival (TDOA) is one of the widely used localization schemes, in which the target (source) emits a signal and a number of anchors (receivers) record the arriving time of the source signal. By calculating the time difference of different receivers, the location of the target is estimated. In such a scheme, receivers must be precisely time synchronized. But time synchronization adds computational cost, and brings errors which may lower localization accuracy. Previous studies have shown that existing time synchronization approaches using low-cost devices are insufficiently accurate, or even infeasible under high requirement for accuracy. In our scheme (called Whistle), several asynchronous receivers record a target signal and a successive signal that is generated artificially. By two-signal sensing and sample counting techniques, time synchronization requirement can be removed, while high time resolution can be achieved. This design fundamentally changes TDOA in the sense of releasing the synchronization requirement and avoiding many sources of errors caused by time synchronization. We implement Whistle on commercial off-the-shelf (COTS) cell phones with acoustic signal and perform simulations with UWB signal. Especially we use Whistle to localize nodes of large-scale wireless networks, and also achieve desirable results. The extensive real-world experiments and simulations show that Whistle can be widely used with good accuracy.

Index Terms—Localization, TDOA, time synchronization, wireless application

✦

1 INTRODUCTION

THE proliferation of wireless and mobile devices has fostered the demand for context-aware applications, among which location is often the most important. One of most commonly used localization is the TDOA-based system, where the target emits a signal and a set of anchors (receivers) with known location record the arriving time of the signal. Calculating the time differences of signal's arrival over anchors will localize the target. Another TDOA scheme uses a combination of signals of different propagation speeds. Taking Cricket [33] as a typical example, a sender emits ultrasound and RF signals simultaneously, while a receiver records the arrival time of them. The difference of two time stamps can be further used to calculate the sender-receiver distance. Different from the multimodule signal TDOA, in this study, we consider time difference as the difference of the arrival times of a same signal at different anchors [13], [18], [20], [39], [42]. In fact, for both TDOA systems, the time synchronization among anchors is highly needed.

Many synchronization polices for low-cost wireless devices are designed, such as RBS [12], TPSN [21], FTSP [29], GTSP [38], providing a common timescale for local clocks of nodes in a network for many applications. The time synchronization is generally obtained based on a set of beacon message exchanges. Unfortunately, the message carrying the time stamp often suffers from several sorts of delay before it can be confirmed by other nodes. Sivriava and Yener [36] decomposes the time uncertainties into four components:

- *Send time*. Time spent to generate a beacon message at the sender.
- *Access time*. Delay at the medium access control (MAC) layer before actual transmission.
- *Propagation time*. Time spent for a message to be transmitted in the air from a sender's network interface to a receiver's.
- *Receive time*. Time needed for a receiver to receive and decode the message, and report it to upper layer applications.

Some synchronization solutions can remove one or two above error sources, but not all of them [21], [38]. What's worse, even a solution that can correctly handle the four errors is still not perfect due to the clock skew and drifting. Thus, synchronization is often not a one-time task and needed to be carried out frequently during the lifetime of nodes, thereby introducing substantial costs. So, the insufficiency of existing synchronization solutions makes high-accuracy localization extremely challenging, especially for COTS devices.

In this paper, we propose Whistle, a novel TDOA localization without time synchronization, which achieves high accuracy by two techniques: Two-signal sensing and sample counting. First, anchors record both the source signal and a successive signal emitted by a Whistle anchor (called two-signal sensing). After the recording stage, each

---
- *B. Xu and R. Yu are with the Department of Computer Science, Tsinghua University, 10-206, The East Main Building, Beijing 100084, China. E-mail: xubin@tsinghua.edu.cn, yuran@keg.cs.tsinghua.edu.cn.*
- *G. Sun is with the School of Information, Beijing Forestry University, 412, The Main Building, Beijing 100083, China. E-mail: gdsun.thu@gmail.com.*
- *Z. Yang is with the School of Software, Tsinghua University, 10-205, The East Main Building, Beijing 100084, China. E-mail: yangzh@cse.ust.hk.*

anchor can count the number of samples between these two signals and then derive the elapsed time between them (called sample counting). Then, each anchor reports its measurement to a back-end server, which will calculate the TDOA values between distinct anchors. Finally, Whistle can locate the sound target by solving the set of equations of TDOA measurements.

Compared with other TDOA-based systems, which often use customized hardware and software to reduce the error in synchronization and TDOA measurement, Whistle achieves good accuracy with only low cost and easy deployment. Totally, the main contributions of this work are as follows: 1) Synchronization free. Whistle changes the scheme of TDOA fundamentally, removing the synchronization requirement, and avoiding many sources of inaccuracy found in other TDOA approaches. 2) High time resolution and accuracy. Using 44.1-kHz sampling rate, Whistle achieves an a time granularity of 0.023 ms. Such a sound recording rate is supported by the microphones of most COTS devices, such as mobile phones, PDAs, and so on. The mean location error is $10 \sim 20$ centimeters in a $9 \times 9 \times 4 \, \mathrm{m}^3$ three-dimensional space, which satisfies the requirement of a large number of applications. The simulations with UWB signal also get good accuracy. 3) Low costs, easy deployment, and widespread use. We implement Whistle on COTS smart phones with Windows Mobile platform, and so Whistle is of low cost and easy deployment. The simulations reveal that the scheme of Whistle can also be used in large-scale wireless network localization.

The rest of the paper is organized as follows: In Section 2, we review the significant researches related to our work. In Section 3, the design of Whistle is presented with more details and analyses. In Section 4, we introduce the architecture of our Whistle prototype system and its basic working process. Section 5 involves extensive testbed examination of Whistle in terms of localization accuracy. We also conduct extensive simulations to investigate what performance Whistle can achieve in localization with RF signals, and in node localization of large-scale wireless networks (see the supplementary file, which can be found on the Computer Society Digital Library at http://doi. ieeecomputersociety.org/10.1109/TPDS.2012.248). Section 6 concludes this paper.

2 RELATED WORK

Many range-based localization algorithms have been proposed recently. Existing works basically use three types of measurements of acoustic signals: TDOA [33], [5], [28], [40], [19], direction of arrival (DOA) [25], and received signal energy [34], [35], [6], [30], [4].

Cricket [33] is a TDOA-based localization system. It uses concurrent RF and ultrasonic signal to measure TDOA between two nodes referring to their own CPU time stamps. Mahajan and Walworth [28] use TDOA measurements to solve a linear equations system similar to ours. Since their TDOAs are measured by CPU time stamps, clock synchronization is needed. The work in [40] employs eight microphones to obtain the TDOA information, which is further used to estimate the bearings of sound sources. Motivated by animal monitoring in field, the works [3], [22]

implement and evaluate a DOA-based sound localization system called ENSBox [22] and employ it to localize marmot alarm calls. An energy-based localization system is proposed in [6], in which the positions of speakers are estimated using an ad hoc microphone array and the maximum likelihood estimation method. In [30], a set of methods based on energy measurement are developed; particularly, a weighted direct least-squares formulation is presented to improve the performance.

BeepBeep [32] is an acoustic-based ranging system with high accuracy. It measures the distance of two cell phones using only their microphones, speakers, and WiFi connection, without leveraging any pre-planned infrastructure. In BeepBeep, two cell phones emit the same sound signal one after one, and then each phone computes the elapsed time between the two time-of-arrivals (ETOA). With the two ETOAs, the distance between the two phones can be obtained. Both the two nodes of BeepBeep must actively participate the ranging process, so the scheme of BeepBeep can be used for TOA localization, rather than for TDOA inherently because the target (or event) cannot often cooperate with anchors.

Though range-based methods can get high accuracy, they often need expensive equipments and have strict connectivity requirements. Some range-free algorithms [9], [23], [14], [16], [15] have been proposed for localization. DV-hop [9] uses the average distance of each hop and hop counts to calculate the real distance between two nodes. APIT [23] makes tests to know whether the unknown node is inside the triangle formed by different anchor nodes, and obtains the intersected region of all the triangles covering the unknown node, then sets the centroid of that region as the result. Li and Liu [26] reveal that these range-free schemes fail in anisotropic WSNs with possible holes, and propose the Rendered Path (REP) protocol.

Several techniques have been proposed for solving nonlinear TDOA equations. Fang [13] reduces the computation to the solution of a quadratic or a quartic equation, but his method cannot make use of extra measurements from extra receivers to improve position accuracy. More general methods with extra receivers can be found in [20], [37]. They provide closed-form solutions, but their estimators are neither unbiased nor optimal. Abel [24] proposes a divide and conquer (DAC) method that can achieve optimum performance and unbiased estimator when the data vector is appropriately subdivided, but DAC requires quite large Fisher information and is difficult to be implemented. The Taylor-series method [39], [18] can get high accuracy at reasonable noise levels, but it is computationally intensive because of its iterative course and finding a proper initial point to avoid the convergence problems is not easy. The method proposed by Chan and Ho [42] is noniterative and gives an explicit solution, and attains the Cramer-Rao lower bound near the small error region. This method also has a higher noise threshold than DAC. In short, the method balances computational complexity and accuracy, and is thus adopted in our work. As the emitter is usually far away from the receivers, it is natural for [10], [11] to use the hyperbolic asymptotes to approximate the TDOA hyperbolae. This method simplifies the problem but gives a trouble that how to pick $N - 1$ bearings when there are N receivers.
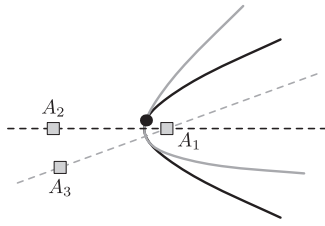
Fig. 1. Illustration of TDOA-based localization, where the target is the intersection point of two hyperbolas (the dark one is formed by anchors A_1 and A_2, and the gray one by A_1 and A_3.).

Doğancay [10] designs a wonderful idea of clustering using kernel density estimation, which greatly reduces computational complexity; however, this kind of methods cannot be used to three-dimensional or near-field localization.

3 MAIN DESIGN

In TDOA-based localization, a TDOA measurement gives the time difference of arrival of a signal emitted from a target to two distinct anchors. With multiple TDOAs as well as locations of anchors, we can deduce the location of target. The critical task in TDOA-based localization is to accurately measure the TDOA between any pair of different anchors. In general, it needs at least three (four) anchors to localize a target in two-dimensional (three-dimensional) space. Fig. 1 shows a two-dimensional example of localization, where a TDOA measurement together with locations of two anchors define a branch of a hyperbola, whose foci are the locations of both anchors.

3.1 System Overview

In the design of Whistle, the target should be able to emit some type of signal to demonstrate its existence. Sound signals are suitable for a large set of devices, including mobile phones, PDAs, and sensor motes. In this study, we choose a sound source as the target. In fact, the techniques developed here can be applied for other scenarios. We denote by S an acoustic signal that is able to emit an omnidirectional sound signal lasting for a specified period. In Whistle, we assume S can be mathematically described in advance. Technically, Whistle does not rely on any specific form of sound signals and consequently can be used for more general purpose. Besides the acoustic source S, Whistle involves a few anchors (also called receivers) that are placed with known locations and receive the acoustic signal from S. Each anchor has a basic set of hardware, including a speaker, a microphone, and wireless connector (such as Bluetooth or WiFi).

The basic idea of Whistle is as follows: For an anchor, it needs to record not only the source signal but another signal emitted by some Whistle anchor. According to the sound record, the anchor can locally detect the interval of the arrival time of these two signals, without referring to the clock of any other anchor. Whistle uses such time interval measured locally to calculate TDOA between two anchors, still without needing any time synchronization. The essential contribution of Whistle is removing the time synchronization required by conventional TDOA-based method.

It is noticeable that in TDOA-based system, the TDOA equations, often nonlinear, cannot be solved directly.
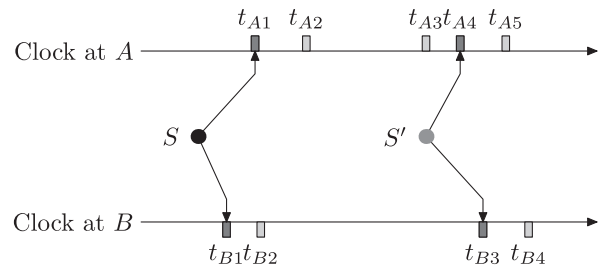

Fig. 2. The illustration of TDOA and TD2S measurements. The dark box represents the time of signals arriving at microphone, the light one the time of signals arriving at (emitting from) CPU.

Researchers [42], [13], [24] have developed several estimators to achieve better accuracy in solving TDOA equations . In this study, we adopt Chan's method [42] because of its accuracy and the desirable computation complexity. So Whistle concentrates on how to accurately measure the TDOA values between any pair of anchors and how to effectively handle with the uncertainties in processing those TDOA measurements.

3.2 Measuring TDOA by TD2S

Here, we describe how to accurately measure TDOA values between two asynchronous anchors. Fig. 2 shows a typical time sequence of two Whistle anchors A and B. Let t_{A1} and t_{B1} denote the times of Ss arriving at the microphones of A and B, respectively. Clearly, the TDOA between A and B can be measured with $(t_{B1} - t_{A1})$ if two anchors could be perfectly time synchronized. In reality, however, A's CPU can identify the arrival of S just at t_{A2}, not at t_{A1}. The same latency occurs at anchor B. Many experimental studies demonstrate that the software latency, such as $(t_{A2} - t_{A1})$ or $(t_{B2} - t_{B1})$, is inevitable and unpredictable [32]. Therefore, we still cannot exactly express the TDOA value by $(t_{B2} - t_{A2})$. In summary, the uncertainties and errors caused by software and hardware make it impossible to implement perfect time synchronization, thereby disabling $(t_{B1} - t_{A1})$ or $(t_{B2} - t_{A2})$ to be an effective estimate of the TDOA between anchors A and B.

To eliminate the uncertainties and errors in time synchronization, we introduce another acoustic signal S' (called the *system sound*) emitted by one Whistle anchor, which is called the *base anchor* (other anchors are called the *nonbase anchor*). With the help of the system sound S', we develop a method to measure $(t_{B1} - t_{A1})$. As shown in Fig. 2, after a specified time interval, A, as the base anchor, emits S' from its speaker at time t_{A3} according to its own clock. Next, we show the calculation of the TDOA between two anchors according to the arrival times of S and S'.

As shown in Fig. 2, the TDOA value between receivers A and B, denoted by T_{AB}, is expressed with $(t_{B1} - t_{A1})$ if both anchors are time synchronized. Now, suppose A is the base anchor. A emits system signal S' from speaker at time t_{A3}, and then S' arrives at the microphones of A and B at t_{A4} and t_{B3}, respectively. A and B's CPUs detect S' at t_{A5} and t_{B4}, respectively. We define the TD2S value of an anchor as the time interval between arrivals of S and S' at this anchor, and denote by T_{A2S} the TD2S value measured by an anchor A. Obviously, we have $T_{A2S} = (t_{A4} - t_{A1})$ and $T_{B2S} = (t_{B3} - t_{B1})$. Note that the two TD2S values of A and
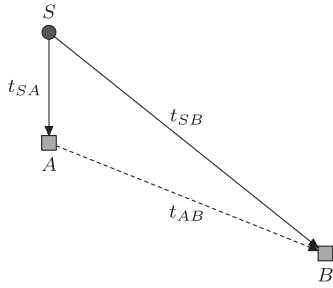
Fig. 3. An illustration of transmissions of two signals.

Fig. 4. The illustration of arrivals of S and S'.

B can be calculated using only their own local clock. Equation (1) presents a mathematical manipulation to express the TDOA between A and B using two TD2S values measured by A and B, respectively,

$$
\begin{aligned}
T_{AB} &= t_{B1} - t_{A1} \\
&= (t_{B3} - t_{A3}) - (t_{B3} - t_{B1}) + (t_{A3} - t_{A1}) \\
&= (t_{B3} - t_{A3}) - (t_{B3} - t_{B1}) \\
&\quad + (t_{A4} - t_{A1}) - (t_{A4} - t_{A3}) \\
&= \frac{d_{AB}}{v} - (t_{B3} - t_{B1}) + (t_{A4} - t_{A1}) - \frac{d_{AA}}{v} \\
&= \frac{1}{v}(d_{AB} - d_{AA}) - (T_{B2S} - T_{A2S}),
\end{aligned} \tag{1}
$$

where d_{AB} is the distance from A's speaker to B's microphone, and d_{AA} the distance from A's speaker to A's microphone. The two distance variables and the speed v are all constants. Clearly, we only need to calculate T_{A2S} and T_{B2S} to calculate the TDOA value T_{AB}. Moreover, the TD2S value of a receiver can be obtained by its local clock, without any synchronization with others. Thus, we use (1) to eliminate the requirement for time synchronization in measuring TDOA between two anchors.

Before introducing the method of precisely measuring TD2S values, we explain why the TD2S value is existing—why is the arrival of S' at a nonbase receiver always later than the arrival of S at this receiver? Suppose that the target and all anchors are placed in a line-of-sight environment. And suppose the base anchor A sends S' with the time delay of Δ_t after it receives S. We use Fig. 3 to prove the existence of TD2S. We assume that the propagation times of S arriving at A and B are t_{SA} and t_{SB}, respectively, and that the propagation time of S' from A to B is t_{AB}. From Fig. 3, we can see the three propagation times construct a triangle in the line-of-sight environment, which means $(t_{SA} + t_{AB}) \geq t_{SB}$. Since the time Δ_t, the delay at the system anchor A, is always greater than zero, we have $(t_{SA} + t_{AB}) > t_{SB}$, i.e., for any nonbase anchor, it always receives S' later than S. Thus, we can measure the TD2S value at any nonbase anchor, according to its local clock.

3.3   Measuring TD2S

In Fig. 2, because latencies are unpredictable, $(t_{A4} - t_{A1})$ is not always equal to $(t_{A5} - t_{A2})$. For a receiver, the TD2S measured by the difference of its two CPU time stamps is not promising, as traditional synchronizing methods attempt to do. Recall that when an anchor is in the recording state, it samples the acoustic signal at a fixed frequency (fHz). In
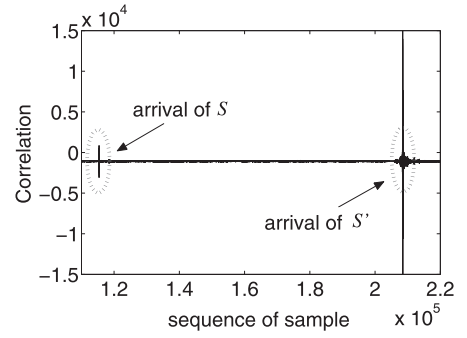
other words, every $1/f$ second, the anchor senses and translates sound into real or complex digits by its A/D converter. For example, while the signal S arrives at the microphone of anchor X, it is recording the ith sample. After a while, X's microphone detects the signal S' when its sample counting goes to j. Now, we can express T_{X2S}, the TD2S value of X, by $(j - i)/f$. Thus, we eliminate the uncertainties that are inherent for the time-stamp-based synchronization A higher sampling rate results in a higher accuracy. In this paper, the sampling rate is 44.1 KHz, which is supported by most COTS devices. Another critical challenge for Whistle is precisely detecting both arrivals of S and S' for calculation of TD2S as shown in Fig. 4.

We assume that both S and S' use the same form of sound signal. If the signal has a good autocorrelation property, it is easy to accurately determine the peak relevant to the arrival of S by correlating it with the known signal. Specifically, the two maximal peaks returned by the cross correlation of time domain can represent the arriving times of S and S'. In practice, the highest peak cannot always represent the arrival of signal because of multipath effects and other uncertainties. Empirically, we determine the arrival of signal by choosing the earliest *sharp peak* in a shadow window of ω_0 sampling points on the right side of the highest peak. Here, we weigh the sharp peak with two parameters: The height and the average slope of point p, which is denoted by $L(p)$ and calculated by $(Y_p - Y_v) / (X_p - X_v)$, where (X_p, Y_p) is the coordinate of peak p in the time-domain data of cross correlation and (X_v, Y_v) is the nearest valley point before p. Only the peak will be chosen whose height is larger than $Y^{max} \times TH_Y$ and whose slope is larger than $L^{max} \times TH_P$. In Whistle, we empirically set $\omega_0 = 5,000$, $TH_Y = 0.5$, and $TH_P = 0.5$.

Generally, we can use any signal with good autocorrelation property as the reference signal. Considering the properties of the built-in speaker and microphone of a cell phone, in this paper, we select the linear chirp signal (sounds like a whistle) with the frequency changing from 2 to 6 KHz as S and S', and the duration of the chirp signal is set with 50 ms.

We have shown a method using TD2S to calculate TDOA. But, sometimes, we may obtain wrong TD2S from defective recording possibly caused by obstacles between S and anchors or background noises. Moreover, because we do not know when and where the source signal S is emitted, it is very difficult to give a range that the exact value of TD2S should reside in. In this paper, we use a majority

decision-based policy to detect wrong TD2S values. We first present an important observation as follows:

Suppose that the sound source M emits S at t, and that the base anchor N emits S' at $t'(t' > t)$. Nonbase anchors A and B both record the two signals. Thus, the two TD2S values of A and B must satisfy the following relation:

$$
\begin{aligned}
|\Delta_{TD}| &= |T_{A2S} - T_{B2S}| \\
&= \left| \left(\left(t' + \frac{d_{AN}}{v} \right) - \left(t + \frac{d_{AM}}{v} \right) \right) \right. \\
&\quad \left. - \left(\left(t' + \frac{d_{BN}}{v} \right) - \left(t + \frac{d_{BM}}{v} \right) \right) \right| \\
&= \frac{|(d_{AN} - d_{BN}) - (d_{AM} - d_{BM})|}{v} \\
&\leq \frac{|d_{AN} - d_{BN}| + |d_{AM} - d_{BM}|}{v} \leq \frac{2d_{AB}}{v}.
\end{aligned}
\tag{2}
$$

Inequality (2) demonstrates that the absolute difference of two TD2S values measured by two anchors must not be greater than the euclidean distance between these two anchors multiplied by $2/v$, where v is the speed of sound. If the maximum distance between any pair of anchors is D, then it holds that for any two anchors A and B, the absolute difference between their TD2S values, $|\Delta_{AB}|$, is always less than or equal to $2D/v$. In fact, D is easy to be obtained because the locations of all anchors are known in advance. We introduce how to detect outliers of TD2S as follows:

We first form a set P of all TD2S values returned by a localization and sort P in an increasing order. Then, we use a window of size $2D/v$ to cover as many elements of P as possible. Clearly, the difference of any two elements of P is not greater than $2D/v$. In the experiments of testbed, we find that most of TD2S values of P can be covered by such window. According to the majority decision principle, we treat those TD2S values outside P to be outliers. Note that if such method of pruning out outliers is used, we possibly need to introduce redundant anchors or repetition of localization to guarantee there are sufficient TD2S information to calculate TDOA values required by localizing target. In fact, the experiments illustrates that two or three redundant anchors can provide sufficient validated TD2S values, while increasing the location accuracy significantly.

4 SYSTEM IMPLEMENTATION

We first introduce the architecture of Whistle in both hardware and software and then give a walk-through example about how Whistle works.

4.1 System Architecture

In the Whistle prototype, we use eight mobile phones, consisting of six Dopod P800, one O2 XDA Atom, and one SonyEricsson XPeria. Each of them has a pair of built-in speaker and microphone, and a WiFi module. In the system, the O2 phone serves as the base anchor emitting S', the Sony phone as the target emitting S, and all other Dopod phones as nonbase anchors. The signals S and S' are both designated to the linear chirp signal lasting 50 ms.

We organize these eight mobile phones, a WiFi AP (TP-LINK WR340), and a laptop (Lenovo T60) into a star-like network, where all phones connect directly to the AP via
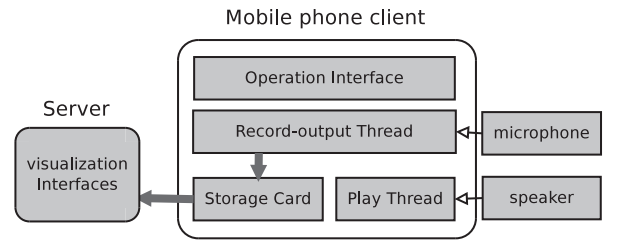

Fig. 5. Software architecture of Whistle.

WiFi interface and the AP relays received data (sound records in wav format) to the laptop for location computation. In fact, a high-performance mobile phone can replace both the laptop and the AP, serving as a hub connecting the other phones.

Fig. 5 shows the software architecture of Whistle, which consists of two components: The server and the client. The server runs on the laptop, providing a localization GUI, by which the user can configure the mobile phone connection and issue to them the command of localization. The client codes run on each anchor phone, using its microphone to record sound with the sampling rate of 44.1 KHz. A mobile phone client is composed of an operation interface, a record-output thread, and a playing thread. The data recorded by microphone are written into phone's storage card as a .wav file and then sent to the server via wireless channel.

4.2 A Walk-Through Example

In this section, we use a walk-through instance to show the working process of Whistle in experiments.

First, we connect all mobile phones to the laptop in a star-like formation. After successfully connecting those devices, mobile phones send to the laptop their basic information, covering their IP, location (excluding S's location), ID, and so on. These information can be displayed and managed by the server. Second, being initialized, the server starts a task of localization by sending to anchors a command and determines whether all anchors are ready to record by counting the ACK messages replied from these anchors. If yes, then the server automatically triggers the target phone and the base anchor phone to let them emit S and S'. Third, the base anchor phone will reply an ACK to the server if it finishes emitting signal S'. Upon receiving this ACK, the server informs all nonbase anchor phones to stop recording and send back their sound records. At last, the server receives and stores all the sound data, and calculates the location of S as follows:

1. Correlate the sound data from each nonbase anchor phone, with a template signal that is the same as the source signal S.
2. Find the arrival points of S and S' in the data resulted from correlation and then calculate the TD2S value of each nonbase anchor, using the method described in Section 3.3.
3. Prune out the TD2S outliers according to method of detecting outliers in Section 3.3.
4. Calculate the TDOA value between each nonbase anchor and the base anchor, and then use these TDOA values to compute the location of S by Chan's method of solving equations [42].

(a)The 3D view of the testbed   (b)The top view   (c)The right view
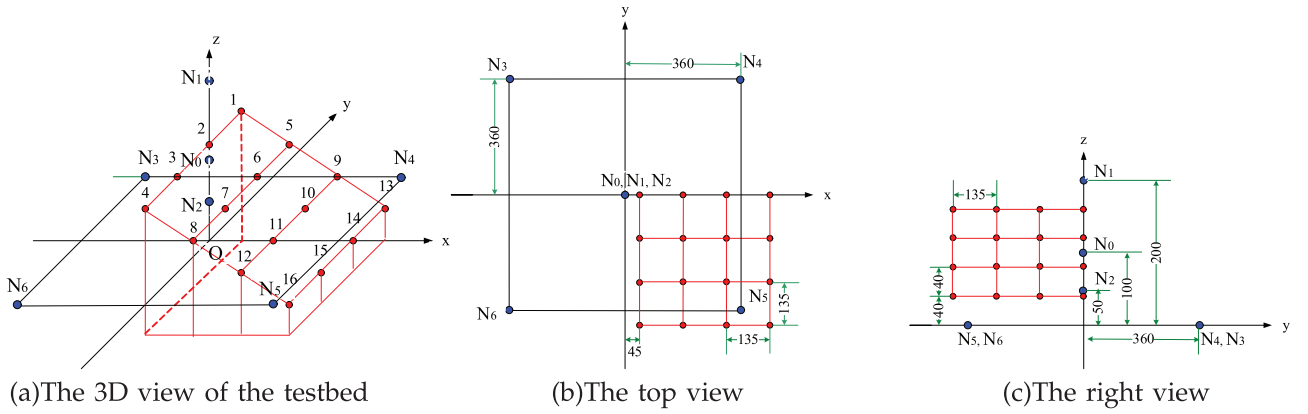
Fig. 6. Deployment of anchors and acoustic targets in three-dimensional experiments (unit: cm). The blue point is the location-fixed anchor and the red one the target examined.

5. After obtaining the location of S, the server will display this location on its UI—an entire localization course ends.

5 EXPERIMENTAL EVALUATION

Through extensive testbed experiments, we examine the performance of Whistle in terms of localization accuracy. Except for some negligible factors, like the variation of speed of sound, there are still other three factors that may weigh against localization accuracy. The first is the signal-to-noise ratio (SNR). Environmental noises are always recorded. If the energy of the transmit signals is limited or the frequency of noise is close to the signals, it will be difficult to detect the arrival time of the signals. The second is the multipath effect. Due to reflection, an acoustic signal may reach to the receiver via different paths. Though we take the earliest peak instead of the highest peak to deal with multipath effects, sometimes the right time point cannot be captured easily. Equation solving is the third factor. Chan's method performs well when the TDOA measurement errors are small [41].

Considering these possible sources of errors, we conduct testbed experiments in different practical scenarios to examine the effect of practical environmental factors on the accuracy of Whistle. Additionally, we perform extensive simulations to examine the feasibility of Whistle working with RF signals and the feasibility of Whistle for node localization in large-scale wireless network. The simulation results and analysis are involved in the supplementary file, available online.

5.1 Settings of Test Environment

Whistle is evaluated in two-dimensional and three-dimensional scenarios. In each scenario, according to the error sources mentioned above, we intentionally conduct the experiments in the following environments. Case 1 (Outdoor, quiet): Whistle is deployed outside a large gymnasium in our campus. Case 2 (Outdoor, noisy): the deployment is identical to Case 1, but a speaker is used to play music loudly for building noisy atmosphere. Case 3 (Indoor, quiet): Whistle runs within a quiet hall of size about $9 \, \mathrm{m} \times 9 \, \mathrm{m} \times 4 \, \mathrm{m}$. Afterward, we describe Case 1 by *Normal*, Case 2 by *Noisy*, and Case 3 by *Inside*, respectively. We place

the sound source at 16 different points in each environment, and collect 10 repeated samples of measuring data at each point. Since temperature changes with time increasing, we record temperature when S changes its position. Formally, the model of sound speed in use is $v = 331.3 + 0.6 \times \theta \, (\mathrm{m/s})$ [1], where θ is the air temperature in Celsius.

For two-dimensional or three-dimensional target localization of all environments, we have the same deployment of anchors and sources. Since the sound signal attenuates rapidly when transmitting in the air, our localization is distance restricted—so we confine the Whistle prototype within an $9 \times 9 \times 4 \, \mathrm{m}^3$ space. If a speaker of higher power is used as sound source, then Whistle will be able to work under a larger space. The three-dimensional and the two-dimensional deployments are shown in Figs. 6 and 7, respectively. Note that we only examine sound sources placed within a quarter of the experimental space because of the spatial symmetry.

5.2 Outlier Identification

In the experiment, we find some localization results with extremely large errors; those errors are from obviously failed experiments that cannot produce enough eligible TD2S or TDOA values. These results significantly lower the localization accuracy. In our accuracy evaluation, we identify the results with the distance error greater than one meter as outliers and prune out them before evaluating the accuracy, much the way BeepBeep [32] does.
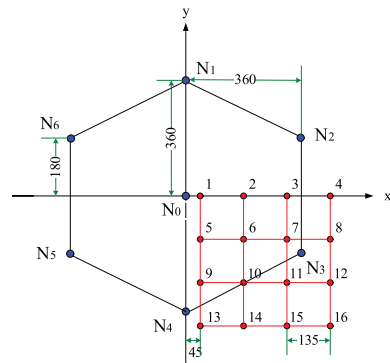

Fig. 7. Deployment of anchors and acoustic targets in 2D case (unit:cm). The blue point is the anchor, and the red one is the target.
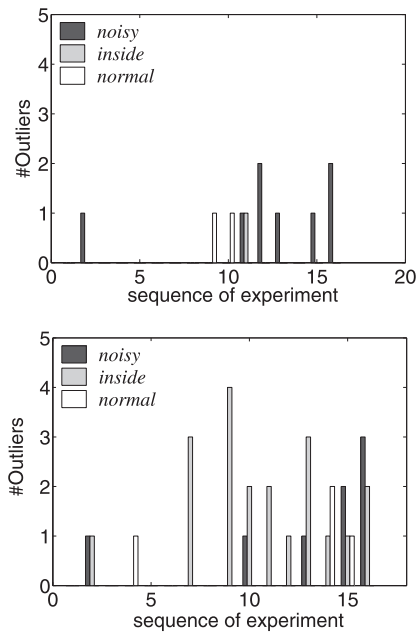
Fig. 8. The number of outliers over 10 repeated localizing operations. The top and bottom figures represent the results for two-dimensional and three-dimensional deployments, respectively.

Fig. 9. The distribution of TDOA errors (ms) for two-dimensional and three-dimensional cases.

In fact, the threshold of 1 meter for outlier identification reflects the bound for the capacity of Whistle. Since the target sound signal transmitted by the mobile phone platform is very low in power, it attenuates so much when arriving at anchor's microphone that it cannot be detected with desirable accuracy. We find, in three-dimensional, *Normal* experiments, that when the source is placed at least 11.2 meters (or more) away from any one of anchors, almost all locations returned are more than 1 meters in error. So we empirically set with 1 meter the threshold for outlier identification—which reflects the maximum capacity of Whistle in a general scenario.

Fig. 8 shows the number of outliers of the 10 repeated operations for each source. It is obvious that there are only very few outlier for most configurations. The outlier percentage for the three-dimensional, *Inside* localization, slightly higher than that forx the two-dimensional localization, is still not beyond 12 percent. The outliers from within the three-dimensional inside case indicates that the multipath effect, which is significant in indoor environment, impacts the accuracy of localization based on the acoustic signals.

Note that here, the outliers are analyzed in an offline manner; in other words, if a location result is recognized as an outlier after being compared with the ground truth of the target, the localization task will be looked as a failure. In practice, the real location of target is unknown in priori, but we may turn to some methods to verify whether a localizing operation fails or not. Our experiments, for instance, show that there exists a close relation between the validity of localization and the distribution (cluster) of the TD2S measurements, i.e., these TD2S values and their distribution feature(s) may indicate a successful or failed localization to be obtained. In this paper, we do not present any formal analysis about the TD2S distribution, because the real environment, as we know, significantly affects the propagation of sound in a prediction-hard way;
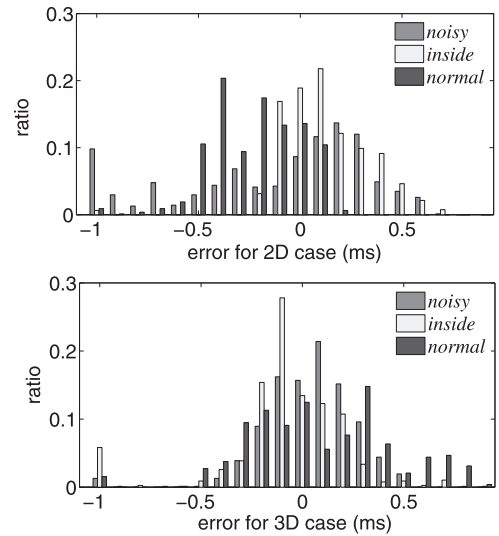
and it may be our future work to improve Whistle for some specific scenarios.

In the Fig. 8, we can see that generally, the two-dimensional experiment have much less outliers than the three-dimensional one (even when both are carried out in the *Indoor* case). The reason behind relies on that redundant nodes benefit the localization accuracy. The two-dimensional case has one more redundant node than the three-dimensional case, so it is easier for two-dimensional to avoid outliers. It is also observed that the noise as well as the multipath effect can result in more outliers. Especially, the multipath effect of three-dimensional case gives strong impacts on the localization than that of two-dimensional case does.

5.3 Accuracy of Localization

To evaluate the accuracy Whistle can achieve, we use an intuitionistic but effective method [32] that examines the TDOA error and the *distance* error of each localizing result, according to the comparison with the ground truth. The distance error is defined as the euclidean distance between the location of the exact target and the location returned by Whistle.

We compare the TDOAs calculated by Whistle with ones calculated according to the ground truth, and plot the distribution of TDOA errors in Fig. 9. We can see that for two-dimensional case with different configurations, more than 85 percent of TDOA errors are less than 0.4 ms, and for three-dimensional case, more than 90 percent TDOA errors are less than 0.4 ms.

Fig. 10 plots the localization accuracy of Whistle. We can see that most of the errors are below 30 cm. Though there is no obvious trend in the Fig. 10, we still find that from point $4m + 1$ to $4m + 4 (0 \leq m \leq 3)$, errors increase mostly. Because in our deployment, among the four points, point $4m + 1$ is the nearest to the center while point $4\,m + 4$ is the farthest (see Figs. 6 and 7), we find that our method performs not so well when S is far from most receivers. Considering that TDOA method is essentially to get the intersection point of a set of hyperbolas or hyperboloids,
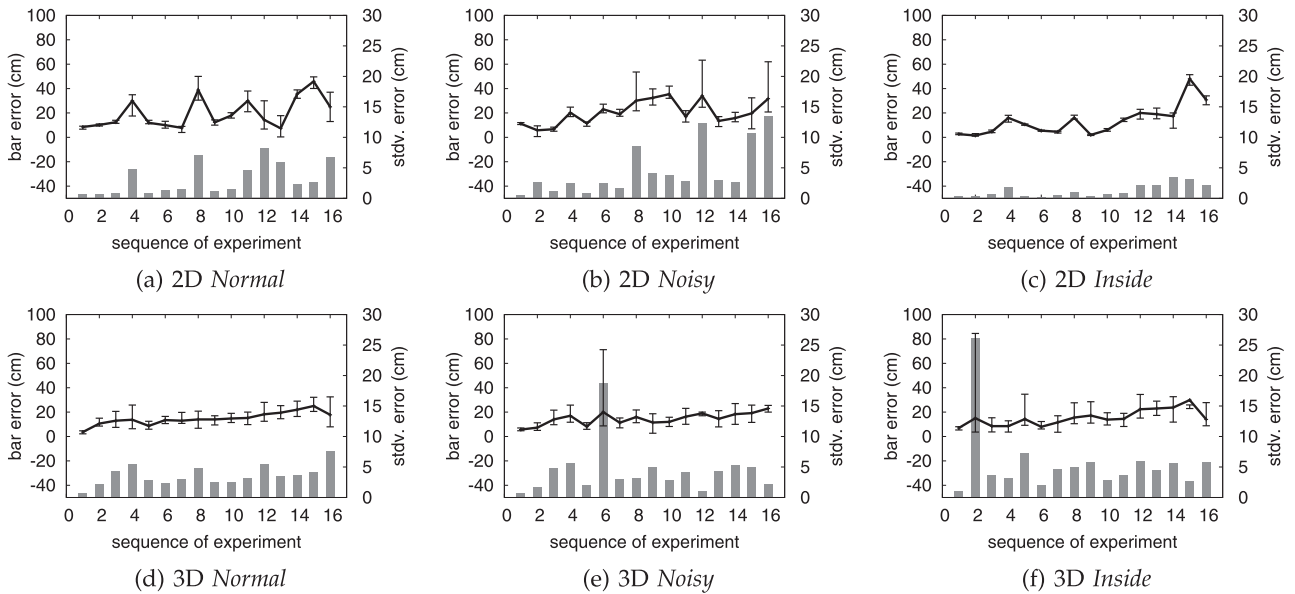
Fig. 10. The accuracy of localization for different scenarios. The bar-error curve (to left y-axis) plots the maximum, mean, and minimum errors over all repeated runs, and the box (right y-axis) plots the standard variance (stdv.) of errors over all repeated runs.

when the sound source is far away, included angle between two hyperbolas would be so small that a little movement of one hyperbola cause significant positional change of intersection point, our conclusion is reasonable.

Overall, Whistle can achieve relatively high accuracy on time measurement and correspondingly on localization, even though the background noise or multipath effect is very significant. Noise and multipath increase the number of outliers, but do not lower the localization accuracy directly. More respects of Whistle's accuracy for two-dimensional and three-dimensional environments are shown in Fig. 11 and Table 1, from which can be seen that Whistle achieves high accuracy with different configurations.
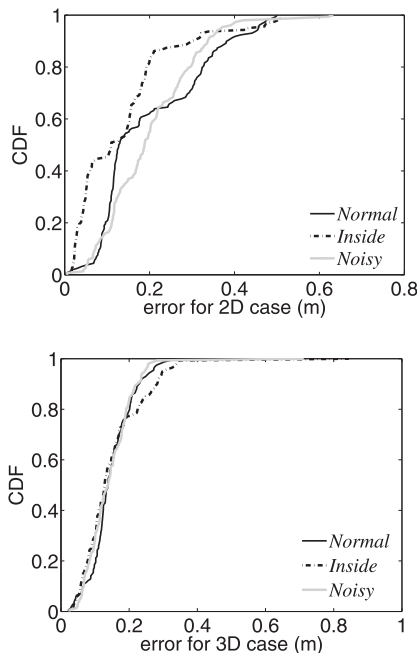
We can see that results in *Normal* environments are not always the best. Though noise and multipath increase the number of outliers, they do not directly enlarge errors of right results. The consequences of three environments are quite similar. We think that our system resists multipath because of the efficient method for finding the first peak, and resists noises because of the good autocorrelation property of chirp signal.

5.4 Discussions about Implementation

By correlating the recording data with template signal, we can calculate the TD2S and then the TDOA between two anchors. Here, a practical problem is: Where should we perform the calculation of TD2S and TDOA? At individual anchor, or at the server? If using the former plan, each anchor needs to locally process the sound file. For a COTS phone (for instance, Dopod P800 with a TI processor of 200 MHz and Windows Mobile 5.0 installed), it will take about 80 seconds to calculate the TD2S and TDOA

Fig. 11. The CDFs of distance error under two deployments.

TABLE 1
Other Statistics of Whistle's Accuracy

Metrics	Case	*Normal*	*Noisy*	*Inside*
mean error (cm)	2D	20.0	20.3	13.7
	3D	14.7	14.4	15.2
50% error (cm)	2D	12.7	18.8	10.9
	3D	13.7	13.8	12.8
90% error (cm)	2D	38.1	33.7	29.9
	3D	22.8	21.9	28.2
std. error (cm)	2D	12.8	10.9	12.1
	3D	6.2	7.4	9.9
outlier (%)	2D	1.25	5	0.6
	3D	2.5	5	12.5

measurements from a 5-second sound record containing 220 thousand sampling points (44.1 KHz). In this way, however, the anchor needs extremely less wireless bandwidth to transmit to the server the obtained TD2S and TDOA data that are of size tens of bytes. If we use the latter plan, each anchor only needs to transmit all its sound data in compressed formation to the server, where TD2S and TDOA are calculated and then localization can be deduced. In practice, it will take a few milliseconds for a COTS phone in 54 Mbps WLAN to transmit to the server a compressed sound file of size 25 Kb. Clearly, the calculation at anchor is a potential bottleneck of real-time performance, and hence, Whistle puts all computation burden (i.e., CPU clock) to the server. If the anchor in use is sufficiently high in computation performance, or if the communication bandwidth is very limited, the first plan may be preferable. In other words, where to take calculation of TD2S and TDOA values needs comprehensively considering the communication bandwidth and the computation capability of anchors.

6 CONCLUSIONS

In this paper, we propose a TDOA-based acoustic source localization framework, called Whistle, which needs no synchronization and achieves high accuracy with COTS devices for both two-dimensional and three-dimensional cases. We also investigate the feasibilities of Whistle for UWB system and for large-scale network localization by extensive simulations. The simulation results, shown in the supplementary file, available online, demonstrate the desirable performance of Whistle. In the future, we will pay particular attention on both extending Whistle to be an energy efficient, scalable localization scheme for multihop wireless network as well as UWB-based deployment, and performing the theoretical error analysis under diverse and practical settings. Additionally, we will examine the Whistle under other smart-phone platforms with high-performance processors.

ACKNOWLEDGMENTS

REFERENCES

[1] B. Smith, R. Peters, and S. Owen, *Acoustics and Noise Control,* second ed. Longman, 1996.

[2] http://www.ubisense.net/en/, 2013.

[3] A.M. Ali, T.C. Collier, and L. Girod, "An Empirical Study of Collaborative Acoustic Source Localization," *Proc. Sixth Int'l Conf. Information Processing in Sensor Networks (IPSN),* pp. 41-51, 2007.

[4] D. Ampeliotis and K. Berberidis, "Linear Least Squares Based Acoustic Source Localization Utilizing Energy Measurement," *Proc. IEEE Fifth Sensor Array Multichannel Processing Workshop (SAMSP),* 2008.

[5] M. Pollefeys and D. Nister, "Direct Computation of Sound and Microphone Locations from Time-Difference-of-Arrival Data," *Proc. IEEE Int'l Conf. Acoustics Speech and Signal Processing (ICASSP),* pp. 2445-2448, 2008.

[6] M. Chen, Z. Liu, L. He, P. Chou, and Z. Zhang, "Energy-Based Position Estimation of Microphones and Speakers for Ad Hoc Microphone Arrays," *Proc. IEEE Workshop Applications of Signal Processing to Audio and Acoustics,* 2007.

[7] W.C. Chung and D. Ha, "An Accurate Ultra Wideband Ranging for Precision Asset Location," *Proc. IEEE Conf. Ultra Wideband Systems and Technologies,* pp. 389-393, Nov. 2003.

[8] B. Denis, J. Keignart, and N. Daniele, "Impact of Nlos Propagation upon Ranging Precision in UWB Systems," *Proc. IEEE Conf. Ultra Wideband Systems and Technologies,* pp. 379-383, Nov. 2003.

[9] D. Niculescu and B. Nath, "Ad Hoc Positioning System (aps)," *Proc. IEEE GlobeCom,* Nov. 2001.

[10] K. Dogancay, "Emitter Localization Using Clustering-Based Bearings Association," *IEEE Trans. Aerospace and Electronic Systems,* vol. 41, no. 2, pp. 525-536, Apr. 2005.

[11] S.P. Drake and K. Dogancay, "Geolocation by Time Difference of Arrival Using Hyperbolic Asymptotes," *Proc. IEEE Int'l Conf. Acoustics Speech and Signal Processing (ICASSP),* 2004.

[12] J. Elson, L. Girod, and D. Estrin, "Fine-Grained Network Time Synchronization Using Reference Broadcasts," *Proc. Fifth Symp. Operating Systems Design and Implementation (OSDI),* 2002.

[13] B.T. Fang., "Simple Solutions for Hyperbolic and Related Position Fixes," *IEEE Trans. Aerospace and Electronic Systems,* vol. 26, no. 5, pp. 748-753, Sept. 1990.

[14] H. Chen, Q. Shi, R. Tan, and H.V. Poor, "Mobile Element Assisted Cooperative Localization for Wireless Sensor Networks with Obstacles," *IEEE Trans. Wireless Comm.,* vol. 9, no. 3, pp. 956-963, Mar. 2010.

[15] B. Liu, H. Chen, Z. Zhong, and H.V. Poor, "Asymmetrical Round Trip Based Synchronization-Free Localization in Large-Scale Underwater Sensor Networks," *IEEE Trans. Wireless Comm.,* vol. 9, no. 11, pp. 3532-3542, Nov. 2010.

[16] Q. Shi, S. He, H. Chen, and L. Jiang, "Distributed Wireless Sensor Network Localization via Sequential Greedy Optimization Algorithm," *IEEE Trans. Signal Processing,* vol. 58, no. 6, pp. 3328-3340, June 2010.

[17] J. Foerster, "Channel Modeling Sub-Committee Report (Final)," Technical Report P802.15-02/490rl-SG3a, IEEE P802.15 Working Group for Wireless Personal Area Networks (WPANs), Feb. 2003.

[18] W.H. Foy, "Position-Location Solutions by Taylor-Series Estimation," *IEEE Trans. Aerospace and Electronic Systems,* vol. AES-12, no. 2, pp. 187-194, Mar. 1976.

[19] K. Frampton, "Acoustic Self-Localization in a Distributed Sensor Network," *IEEE Sensor J.,* vol. 6, no. 1, pp. 166-172, Feb. 2006.

[20] B. Friedlander.Mo, "A Passive Localization Algorithm and Its Accuracy Analysis," *IEEE J. Oceanic Eng.,* vol. OE-12, no. 1, pp. 234-245, Jan. 1987.

[21] S. Ganeriwal, R. Kumar, and M.B. Srivastava, "Timing-Sync Protocol for Sensor Networks," *Proc. First Int'l Conf. Embedded Networked Sensor Systems (SenSys),* pp. 138-149, 2003.

[22] L. Girod, M. Lukac, V. Trifa, and D. Estrin, "The Design and Implementation of a Self-Calibrating Distributed Acoustic Sensing Platform," *Proc. Fourth Int'l Conf. Embedded Networked Sensor Systems (SenSys),* pp. 71-85, 2006.

[23] T. He, C. Huang, B. Blum, J. Stankovic, and T. Abdelzaher, "Range-Free Localization Schemes for Large Scale Sensor Networks," *Proc. ACM Mobicom,* pp. 81-95, 2003.

[24] J.S. Abel, "A Divide and Conquer Approach to Least-Squares Estimation," *IEEE Trans. Aerospace and Electronic Systems,* vol. 26, no. 2, pp. 423-427, Mar. 1990.

[25] A. Ledeczi, G. Kiss, B. Feher, P. Volgyesi, and G. Balogh, "Acoustic Source Localization in Fusing Sparse Direction of Arrival Estimates," *Proc. Int'l Workshop Intelligent Solutions Embedded Systems (ISES),* 2006.

[26] M. Li and Y. Liu, "Rendered Path: Range-Free Localization in Anisotropic Sensor Networks With Holes," *IEEE/ACM Trans. Networking,* vol. 18, no. 1, pp. 320-332, Feb. 2010.

[27] Z.N. Low et al., "Pulse Detection Algorithm for Line-Of-Sight (Los) UWB Ranging Applications," *IEEE Antennas and Wireless Propagation Letters,* vol. 4, pp. 63-67, 2005.

[28] A. Mahajan and M. Walworth, "3-D Position Sensing Using the Differences in the Time-of-Flights from a Wave Source to Various Receivers," *IEEE Trans. Robotics and Automation,* vol. 17, no. 1, pp. 91-95, Feb. 2001.

[29] M. Maroti, B. Kusy, G. Simon, and A. Ledeczi, "The Flooding Time Synchronization Protocol," *Proc. Second Int'l Conf. Embedded Networked Sensor Systems (SenSys),* pp. 39-49, 2004.

[30] C. Meesookho, U. Mitra, and S. Narayanan, "On Energy-Based Acoustic Source Localization for Sensor Network," *IEEE Trans. Signal Processing,* vol. 56, no. 1, pp. 365-377, Jan. 2008.

[31] M. Pendergrass, "Empirically Based Statistical Ultra-Wideband Channel Model," technical report, IEEE P802.15-02/240-SG3a.

[32] C. Peng, G. Shen, Y. Zhang, Y. Li, and K. Tan, "Beepbeep: A High Accuracy Acoustic Ranging System Using Cots Mobile Devices," *Proc. Fifth Int'l Conf. Embedded Networked Sensor Systems (SenSys),* pp. 1-12, Nov. 2007.

[33] N.B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The Cricket Location-Support System," *Proc. ACM Mobicom,* pp. 32-43, 2000.

[34] X. Sheng and Y. Hu, "Energy Based Acoustic Source Localization," *Proc. Second Int'l Conf. Information Processing in Sensor Networks (IPSN),* 2003.

[35] X. Sheng and Y. Hu, "Maximum Likelihood Multiple-Source Localization Using Acoustic Energy Measurements with Wireless Sensor Networks," *IEEE Trans. Signal and Processing,* vol. 53, no. 1, pp. 44-54, Jan. 2005.

[36] F. Sivrikaya and B. Yener, "Time Synchronization in Sensor Networks: A Survey," *IEEE Network,* vol. 18, no. 4, pp. 45-50, July/Aug. 2004.

[37] J.O. Smith and J.S. Abel, "The Spherical Interpolation Method of Source Localization," *IEEE J. Oceanic Eng.,* vol. OE-12, no. 1, pp. 246-252, Jan. 1987.

[38] P. Sommer and R. Wattenhofer, "Gradient Clock Synchronization in Wireless Sensor Networks," *Proc. Int'l Conf. Information Processing in Sensor Networks (IPSN '09),* pp. 37-48, 2009.

[39] D.J. Torrieri, "Statistical Theory of Passive Location Systems." *IEEE Trans. Aerospace and Electronic Systems,* vol. AES-20, no. 2, pp. 183-198, Mar. 1984.

[40] J.-M. Valin, F. Michaud, J. Rouat, and D. Letoumeau, "Robust Sound Source Localization Using a Microphone Array in a Mobile Robot," *Proc. IEEE/RSJ Int'l Conf. Intelligent Robots Systems (IROS),* 2003.

[41] B. Xu, R. Yu, G. Sun, and Z. Yang, "Whistle: Synchronization-Free Tdoa for Localization," *Proc. 31st Int'l Conf. Distributed Computing Systems (ICDCS '11),* June 2011.

[42] Y.T. Chan and K.C. Ho, "A Simple and Efficient Estimator for Hyperbolic Location," *IEEE Trans. Signal Processing,* vol. 42, no. 8, pp. 1905-1915, Aug. 1994.

Bin Xu received the PhD degree from the Department of Computer Science and Technology of Tsinghua University, Beijing, China, in 2006. He is currently an associate professor of computer science in Tsinghua University. His research interests mainly include sensor networks, mobile computing, and service composition. He is a member of the IEEE and the ACM.

Guodong Sun received the PhD degree of computer science from Harbin Institute of Technology, Harbin, China, in 2009. From 2009 to June 2012, he was an assistant researcher of computer science at Tsinghua University, Beijing, China. Since June 2012, he has been with Beijing Forestry University Information School, Beijing, China. His research interests include wireless sensor networks and mobile computing. He is a member of the IEEE.

Ran Yu received the BEng degree in computer science from Tsinghua University, Beijing, China, in 2004. Since 2008, he has been a postgraduate student with the Department of Computer Science and Technology. His research interests include wireless sensor networks and mobile computing. He is a student member of the IEEE.

Zheng Yang received the BEng degree in computer science from Tsinghua University, Beijing, China, in 2006, and is currently working toward the PhD degree with the Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong. His primary research interests include wireless ad hoc/sensor networks and pervasive computing. He is a student member of the IEEE and a member of ACM.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.