# ARABIS: an Asynchronous Acoustic Indoor Positioning System for Mobile Devices

# ARABIS: AN ASYNCHRONOUS ACOUSTIC INDOOR POSITIONING SYSTEM FOR MOBILE DEVICES

BY

YU-TING WANG, M.Sc., Electrical Engineering

National Taiwan University, Taipei, Taiwan

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

Master of Applied Science (2017) McMaster University

(Electrical & Computer Engineering) Hamilton, Ontario, Canada

| | |
|---|---|
| TITLE: | ARABIS: an Asynchronous Acoustic Indoor Positioning System for Mobile Devices |
| AUTHOR: | Yu-Ting Wang<br>M.Sc., Electrical Engineering<br>National Taiwan University, Taipei, Taiwan |
| SUPERVISOR: | Dr. Rong Zheng and Dr. Dongmei Zhao |
| NUMBER OF PAGES: | xii, 54 |

# Abstract

Acoustic ranging based indoor positioning solutions have the advantage of higher ranging accuracy and better compatibility with commercial-off-the-self consumer devices. However, similar to other time-domain based approaches using Time-of-Arrival (ToA) and Time-Difference-of-Arrival (TDoA), they suffer from performance degradation in the presence of multi-path propagation and low received signal-to-noise ratio (SNR) in indoor environments. In this thesis, we present ARABIS, a robust and low-cost acoustic indoor positioning system (IPS) for mobile devices using an asynchronous TDoA-based solution. We develop a low-cost acoustic board custom-designed to support large operational ranges and extensibility. To mitigate the effects of low SNR and multi-path propagation, we devise robust algorithms that iteratively remove possible outliers by taking advantage of redundant TDoA estimates. Experiments have been carried out in two testbeds of sizes 10.67m $\times$ 7.76m and 15m $\times$ 15m, one in an academic building and one in a convention center. The proposed system achieves average and 95% quantile localization errors of 7.4cm and 16.0cm in the first testbed with 8 anchor nodes and average and 95% quantile localization errors of 20.4cm and 40.0cm in the second testbed with 4 anchor nodes only.

# Acknowledgements

The completion of this degree is very rewarding for me because this work not only advances a new approach to deploy indoor localization system but also accomplishes the whole system design. As part of the work in this thesis, I participated in 2017 Microsoft Indoor Localization Competition with Jun Li and Dr. Rong Zheng. It was one of the most stressful periods in my graduate studies in McMaster but I never regret it. Without the competition, we will not devote all research resources to implementing a system in real-world environment and optimizing the performance from an engineering point of view. Luckily, I came up with new algorithms in the end of this process and was still able to continue making research progress.

I would like to thank my supervisor Dr. Rong Zheng for all the training she provided. She is extremely dedicated to research and always brings up new research ideas. Whenever students want to discuss research, she is always there to help. We also have opportunities to work on various platforms because of her generous funding for the hardware. Our group participation in the competition was an awesome idea suggested by her and it turned out to be an unforgettable event in my student life.

Advice given by my co-supervisor Dr. Dongmei Zhao has always been a great help. Whenever I felt lost, she always gave me a different perspective and cheered me up. She also helped improve my technical writing skills. My first TA experience,

working for her, was a real pleasure because she gave me all the freedom in designing the tutorial contents.

It is my honor to express the thankfulness to Dr. Terry Todd and Dr. Douglas Down. They have been in my supervisory committee for two years and have offered valuable suggestions and discussion.

I would like to offer my special thanks to Jun Li, my best partner in WiSeR Group. He is an outstanding hardware designer with a strong sense of responsibility. I cannot remember how many times we stayed up all night together to test the hardware and deployment of our localization system. I was so touched especially when he still had other projects and courses to work on. I have no words to express my appreciation to him.

I am grateful for all kinds of assistance provided by my lab members Qiang Xu, Ala Shaabana, and Zhe Gong. They offered me crucial supports with strong technical skills all the time.

Special thanks to my first Canadian friend Nehad Hirmiz. He has been giving me suggestions and helps in all aspects from life to academic and career since we met in 2014.

Finally and most importantly, my sincere gratitude goes to my family for all the support they have provided me throughout the years. Without them, I would not be able to come this far.

# Abbreviations

| | |
|---|---|
| **COTS** | Commercial Off-The-Shelf |
| **LBS** | Location-Based Service |
| **GPS** | Global Positioning System |
| **IPS** | Indoor Positioning System |
| **RSSI** | Received Signal Strength Indication |
| **BLE** | Bluetooth Low Energy |
| **PDR** | Pedestrian Dead Reckoning |
| **IMU** | Inertial Measurement Unit |
| **UWB** | Ultra Wide Band |
| **ToF** | Time-of-Flight |
| **ToA** | Time-of-Arrival |
| **TDoA** | Time-Difference-of-Arrival |
| **AoA** | Angle-of-Arrival |
| **DToA** | Differential-Time-of-Arrival |
| **RF** | Radio Frequency |
| **SNR** | Signal-to-Noise Ratio |
| **LOS** | Line-of-Sight |

| | |
|---|---|
| **NLOS** | Non-Line-of-Sight |
| **NLLS** | Non-Linear Least Squares |
| **ID** | Identifier |
| **CRC** | Cyclic Redundancy Check |
| **TDMA** | Time Division Multiple Access |
| **CDMA** | Code Division Multiple Access |
| **FEC** | Forward Error Correction |
| **MAC** | Media Access Control |
| **EKF** | Extended Kalman Filter |
| **NTP** | Network Time Protocol |
| **WLAN** | Wireless Local Area Network |
| **SPL** | Sound Pressure Level |
| **Hz** | Hertz |
| **dB** | decibel |
| **OS** | Operating System |
| **RTOS** | Real-Time Operating System |
| **MCU** | Microcontroller |
| **ISR** | Interrupt Service Routine |
| **AWGN** | Additive White Gaussian Noise |
| **BER** | Bit Error Rate |
| **CDF** | Cumulative Distribution Function |
| **NaN** | Not a Number |

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Location-based services (LBS) have experienced substantial growth in the last decade with the proliferation of smart devices. A recent survey by Pew Research Center's Internet Project found that 90% of US adult smartphone owners aged eighteen and older use their phones to get directions or other information based on their current location [2]. Service providers can benefit from users' location information in facilitating precision advertising, personalized recommendation, resource tracking, proximity notification, etc.

To provide LBS, location awareness is an essential step. The Global Positioning System (GPS) complimented by cellular and WiFi access map based methods can provide accurate and robust location information in most outdoor environments. In contrast, despite the fact that people spend the majority of their time indoors, indoor positioning systems (IPS) only have limited success due to low accuracy. GPS signals are highly attenuated indoors rendering GPS-based position estimates inaccurate. The main challenges in IPS arise from the lack of pervasive infrastructural support, low accuracy, and the desire to keep user devices as simple as possible.

Various approaches have been considered in IPS solutions. Fingerprinting, an RSSI-based technique, is widely used because it avoids hardware deployment cost and effort by relying on existing network infrastructure such as WiFi, Bluetooth Low Energy (BLE), and Zigbee. This approach can achieve close to 1m accuracy [3]. The main disadvantage is that the fingerprint changes with the environment and needs an update to the fingerprint database, which requires labor-intensive site-survey. Pedestrian Dead Reckoning (PDR), which is infrastructure-free, utilizes IMU sensors on smartphones or wearable devices and can achieve meter-level localization accuracy as well [4]. However, regular absolute position fix must be provided in PDR to avoid error accumulation in the long-term. Time-of-arrival (ToA) or time-difference-of-arrival (TDoA) estimates from radio-frequency (RF) transmitters (e.g., access points, ultra-wide band beacon nodes) address the above shortcomings by directly determining the ranges or pseudo-ranges from anchor nodes to target devices and inferring locations of the latter [5, 6, 7]. In contrast to RF-based solutions, acoustic indoor localization has the advantage of less stringent requirements on timing accuracy and the potential to work with commercial off-the-shelf (COTS) mobile phone devices irrespective of hardware capabilities, as all mobile phones are equipped with at least one speaker and one microphone. Moreover, acoustic-based ToA or TDoA systems increase localization accuracy to decimeter-level [8]. Recently, several acoustic IPS solutions have been developed for mobile phones, including those operating in audible frequency ranges [9, 10, 11], and those in the human inaudible range but within the hearing range of the majority of smart phone devices [1, 12, 13, 14, 15].

We develop ARABIS [14, 16], a robust and low-cost acoustic IPS for mobile devices. The system adopts fixed acoustic anchor nodes that transmit acoustic signals in

the human inaudible range but decodable by smart phone devices. Target mobile devices are acoustically passive. They listen to beacons emitted from anchor nodes and transmit the timestamps of the decoded acoustic beacons to a location server using radio waves. The location server communicates with anchor nodes to compute the locations of the target devices. The localization algorithm is done by an asynchronous TDoA-based solution that uses asynchronous beacons, where anchors can transmit in an uncoordinated matter. This design is advantageous in several aspects. First, it scales well even with a large number of mobile devices despite the limited acoustic bandwidth available on mobile phones in the inaudible range (e.g., 17kHz – 22kHz) since both the locations of anchor nodes and the frequencies of beacon transmissions are fixed. Second, transmissions of acoustic signals consume much more power than RF wireless communication. Thus, our design can reduce the additional power consumption on battery-powered mobile devices compared to those that require active acoustic transmissions from user devices [10, 17, 18, 19, 11]. Third, location updates can be done faster compared to a purely acoustic-based solution due to the higher data rate over RF wireless communication.

However, a number of challenges exist for acoustic IPS to be feasible for large scale deployments. First, tight clock synchronization is required among anchor nodes in a TDoA-based localization system. Second, the speaker transmission power and directionality, and microphone facing of anchor nodes and target nodes, limit the applicability of the approach in large and complex indoor environments. Furthermore, when operating in large indoor areas, the effects of multi-path propagation and low received signal-to-noise ratio (SNR) are not negligible. Non-Line-of-Sight (NLOS)

components may lead to larger range estimations, while low SNR results in mis-detection of transmitted waveforms. Though both are well known problems in time-domain localization solutions, asynchronous beacons pose additional challenges as the pair-wise measurements between anchor nodes are also subject to both effects.

The proposed solution approach builds and improves upon many existing works. Notably, in terms of ranging estimates among anchor nodes, we borrow the ideas in [10] by taking advantage of the physical separation between the acoustic transmitter (namely, the speaker) and the acoustic receiver (namely, the microphone) to allow concurrent transmission and reception of acoustic signals, also known as full-duplex communication. The sample counting technique in [10] is applied to mitigate the uncertainty in recorded time in user-space processes on mobile devices. The waveform design in the proposed solution is inspired by the work [1] and is further improved based on our own simulation and experimental study.

Our main contributions are three-fold. First, we develope a TDoA estimation algorithm with asynchronous acoustic beacons. To the best of our knowledge, this is the first work that eliminates the need for clock synchronization among anchor nodes in TDoA-based acoustic localization with passive target devices. Second, we develop an extensible acoustic board featuring on-board power amplifier, microphone and speaker. The components are chosen to have an operational range of at least 11 meters when two anchors are facing 45 degrees away from each other and at least 50 meters when they face directly to each other. The range can be further extended by connecting external speakers. Third, to mitigate the effects of low SNR and multi-path propagation, we devise robust algorithms that iteratively remove possible outliers by taking advantage of redundant TDoA estimates. Experiments have been

carried out in two testbeds of sizes 10.67m $\times$ 7.76m and 15m $\times$ 15m, one in an academic building on McMaster University campus and one in a convention center in Pittsburgh where the 2017 Microsoft Indoor Localization Competition was held. The proposed system achieves average and 95% quantile localization errors of 7.4cm and 16.0cm in the first testbed with 8 anchor nodes, and average and 95% quantile localization errors of 20.4cm and 40.0cm in the second testbed with 4 anchor nodes only.

The rest of this thesis is organized as follows. We first discuss related work in Chapter 2. The proposed solution approach is presented in Chapter 3. In Chapter 4 we present the detailed implementation including hardware design, acoustic beacon design, and beacon scheduling. Simulation, testbed setup, and experimental evaluations are presented in Chapter 5. Finally, we conclude the thesis in Chapter 6.

# Chapter 2

# Related Work

Indoor localization has received much attention in recent years. Existing approaches can be roughly broken down into two categories, infrastructure-less and infrastructure-based. Infrastructure-less approaches leverage existing infrastructure (e.g., WiFi, cellular) or do not require any infrastructure at all. Infrastructure-based approaches require new infrastructure supports or augmenting existing infrastructure to realize indoor positioning. In both categories, various sensing modalities and signal sources can be utilized. Comprehensive surveys of related work in this field can be found in [20, 21, 22]. In this chapter, we focus on infrastructure-based indoor localization using acoustic signals. There are two types of devices in such localization systems: target nodes and anchor nodes. The former are user devices to be localized, and the latter transmit (receive) acoustic signals, which are usually called "beacons", to (from) target nodes as the signal source of the localization.

The majority of acoustic-based localization systems use ranging techniques including Time-of-Flight (ToF), Time-of-Arrival (ToA), Time-Difference-of-Arrival (TDoA),

etc. ToA-based systems [9, 23] measure the time of flight (ToF) of the signal between a target node and anchor nodes. In order to get the time of flight, stringent time synchronization among anchor nodes and target nodes is needed. Also, the start time of both the beacon transmission and reception in the medium is required. With one additional anchor node, target devices in TDoA-based systems [1, 13, 24] can be located without having to synchronize with anchor nodes or having knowledge of the transmission schedule of anchor nodes. This is easier to deploy than ToA-based systems. However, synchronization among the anchor nodes is still needed. In asynchronous TDoA systems [11, 14, 16], or synchronization-free TDoA systems, no synchronization is required even between anchor nodes. In general, this is achieved by exchanging the beacon timestamps information between anchor nodes and infering the relative transmission time offset. As part of our contribution, ARABIS falls into this category. To summarize, the main difference between ToA, TDoA, and asynchronous TDoA systems lies in whether synchronization is required between the target and the anchor nodes. In the following, we discuss some recent related works of acoustic localization systems.

BeepBeep [10] measures the pair-wise differential time of arrival (DToA) to achieve high-accuracy ranging between two mobile devices without infrastructure support and device synchronization. Three techniques including two-way sensing, self-recording, and sample counting, are proposed to achieve the goal. In Beepbeep, two devices take turns transmitting linear-chirp modulated signals and record signals both from itself and from the other device. With the pulse compression property of chirp signals, devices are able to detect the signal arrival time with tight timing resolution. Next, each device will count the number of sound samples between their beacons, and divide

it by the sampling rate to get the elapsed time between the two ToA events. After exchanging the elapsed time information with each other, the differential of these two elapsed times represents the sum of the time of flight of the two beacons and hence the two-way distance between the two devices. It achieves about 1cm and 2cm average ranging accuracy with less than 2cm standard deviation for typical indoor and noisy outdoor environments, respectively. Although BeepBeep is not a localization system, those three techniques are widely used in other acoustic localization systems.

Whistle [11] proposes a synchronization-free TDoA for localization. It does not require periodic beacon messages from anchor nodes. However, target devices must actively transmit acoustic source signals. The first anchor node which receives a source signal has to transmit a successive signal and all the other anchor nodes record both the source signal and the successive signal. Thus, acoustic transmissions of both anchor nodes and target devices are required. Their 95% quantile localization error in a 9m $\times$ 9m space is 43.2cm.

Lazik et al. [1] propose a TDoA-based acoustic localization system. All anchor nodes are connected together to a multi-channel amplifier like a public announcement infrastructure. It adopts inaudible frequency range and fading between chirp symbols so humans cannot perceive any beacon transmission. Target nodes only passively receive beacons from anchor nodes. To decode concurrent transmitted beacons, Code Division Multiple Access (CDMA) with Forward Error Correction (FEC) is used for media access control (MAC). They are able to localize 95% of the test points to 16.4cm accuracy in a small 5m $\times$ 5m room.

Lazik et al. [13] uses IEEE 802.15.4 radios to distribute clock synchronization from a subset of infrastructure anchor nodes connected to a GPS source. Since

anchor nodes are fed by GPS signals, mobile devices as target nodes can synchronize their continuously free-running audio sampling subsystem with global time. Once synchronized, the system can use TDMA to multiplex the beacon transmission and perform localization using ToA directly instead of TDoA. The latter reduces the number of anchors needed for covering an indoor space and improves localization performance in the face of obstructions. They also design an omni-directional horn to improve the signal quality. In a 4.5m $\times$ 5.5m area with line-of-sight (LOS) signals, the 95% quantile localization error achieved is 10.9cm.

In ALPS [24], anchor nodes are still time synchronized using IEEE 802.15.4 radios with reference to a master anchor node. However, BLE is used to replace GPS in [13]. Anchor nodes continuously broadcast BLE advertisement packets with the shortest advertisement interval (20ms). Each advertisement packet contains a counter value $\tau_{TX}$ indicating the time offset from the broadcast of the BLE advertisement packet to the beginning of the TDMA cycle. After receiving an acoustic beacon, a target node can be synchronized to the TDMA cycle by timestamping the BLE advertisement reception as $\tau_{RX}$ and subtracting the received counter value $\tau_{TX}$ from $\tau_{RX}$. To start using the localization system, ALPS requires users to hold their phones close to one of the beacons at zero-distance for 10 seconds to synchronize with anchor nodes. As a result, ToA localization can be used and the number of anchors required reduce. If target nodes are only synchronized to the TDMA schedule, TDoA localization can still be used. After the previous target nodes synchronization, user-assisted mapping is provided in the system to automatically determine the anchor positions, which are necessary in the localization process. BLE and acoustic propagation characteristics are also collected in the initialization process and will be used later in localization to

classify LOS and NLOS signals. Finally, ALPS use IMU and EKF for tracking moving target devices. Evaluation shows that ALPS can estimate 3D anchor position with a Euclidean distance error of 16.1cm, and can generate maps with room measurements with a 2D Euclidean distance error of 19.8cm. Experiments in six different environments show that the system can identify NLOS signals with over 80% accuracy and track a user's location to within less than 100cm.

In terms of operating acoustic frequency, we can divide related work into three categories. The first category operates in audible frequency ranges using generic devices equipped with a speaker and a microphone. For example, BeepBeep [10] and Whistle [11] choose the frequency between 2–6kHz, where the frequency response is the best on mobile phone speakers and microphones as their primary usage is for voice communication and music playing. The authors of [9] also choose audible sound for similar reasons. The second category employs custom-built ultrasonic devices [17, 18, 19, 25, 26]. Such dedicated systems can achieve high location accuracy but require end-users to be equipped with specialized hardware. With the popularity of smartphones over the past decade, the third category takes advantage of different hearing ranges between a typical human being and high-end smartphone microphones, and utilizes the frequency range above 16kHz or higher to modulate acoustic signals [12, 23, 1, 13, 24, 14, 15, 16]. This can mitigate unwelcome acoustic pollution in the first category of approaches and eliminate the need for special end-user devices in the second category of approaches. Our proposed solution falls in the third category with the additional goal of removing the need for time synchronization.

Depending on whether target devices need to actively transmit acoustic signals for localization, we can further categorize infrastructure-based solutions into two classes:

active and passive systems. Active transmission of acoustic signals [17, 18, 19, 9, 11] from target devices poses difficulties of media access control as the number of users increases in the system. In contrast, in passive systems [25, 26, 23, 1, 13, 24], target devices only need to receive and decode acoustic signals from infrastructure nodes resulting in better scalability, less energy consumption on user devices, and better privacy preservation. Our proposed solution employs the passive approach and thus has similar benefits.

In summary, we list the comparison of selected related works in Table 2.1.

Table 2.1: Comparison of selected related works.

| **Related Work** | **Approach** | **Synchronization** | **User Devices** | **Frequency** |
|---|---|---|---|---|
| Lopes et al. [9] | ToA | Yes | Active | 4.01kHz |
| Whistle [11] | TDoA | No | Active | 2–6kHz |
| Lazik et al. [1] | TDoA | Yes | Passive | 19–23kHz |
| Guoguo [23] | ToA | Yes | Passive | 17–20kHz |
| Lazik et al. [13] | TDoA/ToA | Yes | Passive | 18–24kHz |
| ALPS [24] | TDoA/ToA | Yes | Passive | 20–21.5kHz |
| ARABIS [14, 16] | TDoA | No | Passive | 17.5–21.5kHz |

# Chapter 3

# Solution Approach

## 3.1 System Architecture

There are three types of devices in the system (see Figure 3.1), namely, a location server, anchor nodes, and target mobile devices. On the infrastructure side, multiple acoustic anchors are deployed at fixed locations in an indoor space, each equipped with a microphone and a speaker. The microphone and the speaker on an anchor are spatially separated at a known distance. Locations of all anchors are known. Periodically, the anchors transmit acoustic beacons subject to random backoffs to minimize the chance of collision. Each beacon message includes the identifier (*id*) of the transmitting anchor node and a sequence number (*seqno*) by designs described in Section 4.2. An anchor node decodes the acoustic beacons it receives including ones from itself. For each successful decoded beacon message, a local timestamp (*ts*) is recorded at the time the preamble portion of the message is detected.

Target mobile devices passively listen to acoustic beacons and record the *ids* and *seqnos* from the decoded beacon messages and the timestamps of the preambles. Both

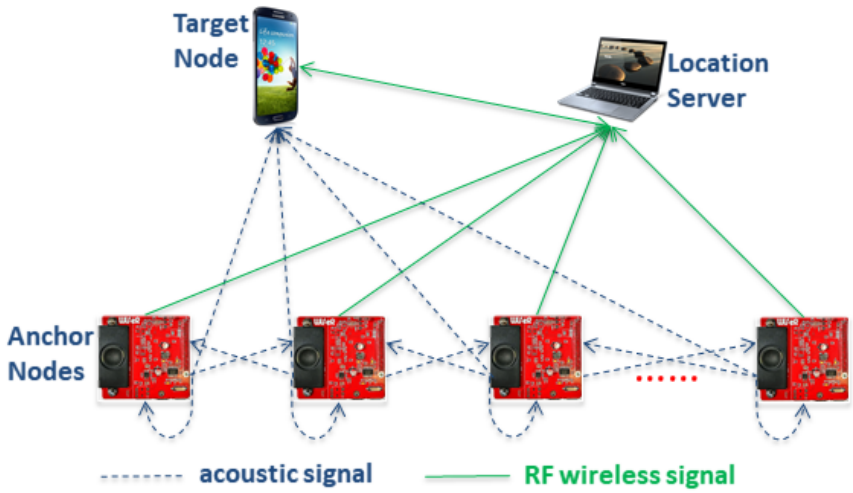


Figure 3.1: System architecture.

anchor nodes and mobile devices can communicate with the location server wirelessly (e.g., via WiFi, bluetooth or Zigbee). Upon receiving sufficient $\langle id, ts, seqno \rangle$ tuples from both sets of devices, the location server will compute the locations of the anchor nodes and the mobile devices using the algorithms detailed in Section 3.3. Without a location server, the location computation can be moved to one of the anchor nodes or each target node. However, extra processing delay may occur and existing tasks may be slowed down. Besides, as will become clear in Section 3.3, part of the location computation is the same for all target nodes. Therefore, centralized computing with a location server is preferred if user location privacy is not concerned.

## 3.2 TDoA Estimation from Asynchronous Beacons

If beacons are transmitted at exactly the same time, TDoA can be directly measured at a target device as the time elapsed between the reception of consecutive beacons. Alternatively, if all anchor nodes have synchronized clocks [9, 1, 12, 13, 15], one can

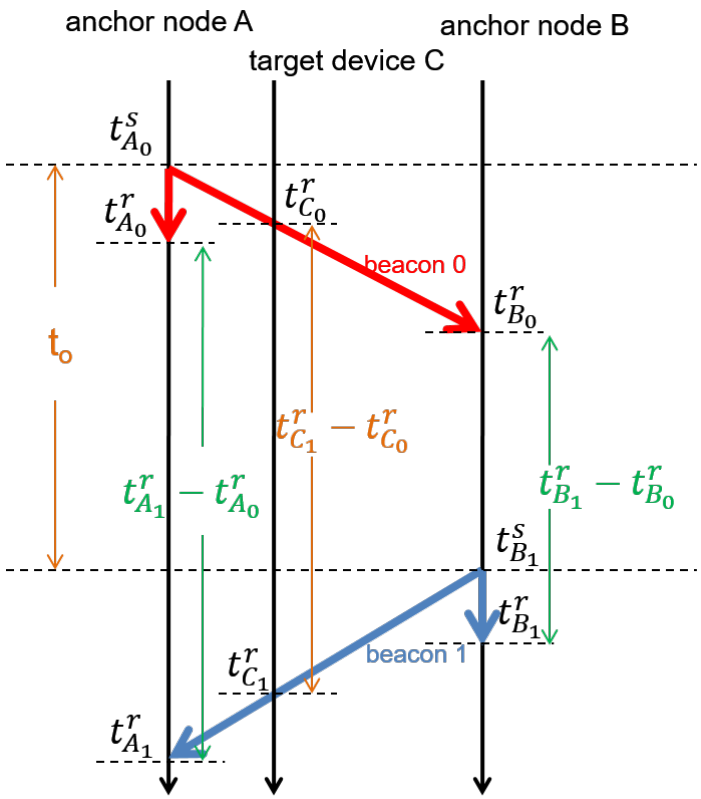


Figure 3.2: Determination of TDoA from a pair of anchor nodes transmitting asynchronously

easily determine the differences in beacon transmission times from their local clocks. In ARABIS, however, we rely entirely on asynchronous beacons. The key idea is to utilize the full-duplex acoustic communication capability of anchor nodes, where a node can receive and decode its own transmitted acoustic signal.

TDoA estimation from asynchronous beacons is best understood through an example as illustrated in Figure 3.2, where there are two anchor nodes A and B and a target device C. At time $t^s_{A_0}$ node A transmits a beacon, which is received at time $t^r_{A_0}$, $t^r_{B_0}$, and $t^r_{C_0}$ at nodes A, B, and C, respectively. All timestamps are based on local clocks. At time $t^s_{B_1}$, anchor B transmits a beacon message, which is received at time $t^r_{A_1}$, $t^r_{B_1}$, and $t^r_{C_1}$ at nodes A, B, and C, respectively. Note that $t^s_{A_0}$ and $t^s_{B_1}$ are not known due to uncertain delays in the acoustic interfaces.

To compute TDoA, we need to know the interval $t_o$ between the transmissions of the two beacons in a common reference time. Denote by $d_{AA}$ and $d_{BB}$, respectively, the distances from A's speaker to its microphone, and from B's speaker to its microphone. Similarly, we denote $d_{AB}$ and $d_{BA}$, respectively, as the distance from A's speaker to B's microphone and the distance from B's speaker to A's microphone. Let $c$ be the speed of sound in the medium. We can see that at node A,

$$\begin{aligned} t_o &= (t^r_{A_1} - \tfrac{d_{BA}}{c}) - t^s_{A_0} \\ &= (t^r_{A_1} - \tfrac{d_{BA}}{c}) - (t^r_{A_0} - \tfrac{d_{AA}}{c}) \\ &= (t^r_{A_1} - t^r_{A_0}) + \tfrac{d_{AA}}{c} - \tfrac{d_{BA}}{c} \end{aligned} \tag{3.1}$$

and, similarly at node B,

$$\begin{aligned} t_o &= t^s_{B_1} - (t^r_{B_0} - \tfrac{d_{AB}}{c}) \\ &= (t^r_{B_1} - \tfrac{d_{BB}}{c}) - (t^r_{B_0} - \tfrac{d_{AB}}{c}) \\ &= (t^r_{B_1} - t^r_{B_0}) - \tfrac{d_{BB}}{c} + \tfrac{d_{AB}}{c}. \end{aligned} \tag{3.2}$$

Taking the average of $t_o$ in (3.1) and (3.2), we have

$$t_o = \frac{(t^r_{B_1} - t^r_{B_0}) + (t^r_{A_1} - t^r_{A_0})}{2} + \frac{(d_{AA} - d_{BB}) + (d_{AB} - d_{BA})}{2c}, \tag{3.3}$$

which can be further simplified as follows under the assumptions that $d_{AB} = d_{BA}$ and $d_{AA} = d_{BB}$,

$$t_o \approx \frac{(t^r_{B_1} - t^r_{B_0}) + (t^r_{A_1} - t^r_{A_0})}{2} \tag{3.4}$$

A couple of comments are in order. First, it is reasonable to assume $d_{AA} = d_{BB}$ as long as the anchor nodes are of the same model. On the other hand, the equality $d_{AB} = d_{BA}$ does not always hold. However, it is feasible to show that $|d_{AB} - d_{BA}| \leq d_{AA} + d_{BB}$. Second, the right hand of (3.4) only requires local information that can be directly measured at the respective anchor nodes.

Once the transmission time offset $t_o$ is determined, the TDoA of the beacons from A and B on node C can be easily computed as

$$TDoA_{A_0B_1} = t^r_{C_1} - t^r_{C_0} - t_o \tag{3.5}$$

We also note that the distance between node A and node B could be approximated as (3.6) [10]. Such ranging measurement between anchors is used later in Section 3.3

to detect outlier time offset.

$$c \times \frac{(t^r_{B_1} - t^r_{B_0}) - (t^r_{A_1} - t^r_{A_0})}{2} + d_{AA} + d_{BB}. \tag{3.6}$$

## 3.3 Robust TDoA Trilateration

The main challenge in TDoA trilateration is the degradation of timing accuracy due to low received SNR and/or existence of multiple paths in signal propagation. The problem is compounded by the fact that the time offset approximation in (3.4) requires the estimation of arrival times on the anchor nodes as well. Consider a pair of anchor nodes $A$ and $B$ and target node $C$, as evident from Figure 3.2, the pair-wise TDoA error as the result of beacon 0 from node $A$ and beacon 1 from node $B$ arises from two sources[1], namely, i) inaccurate timestamps $t^r_{B_0}$ or $t^r_{A_1}$, and ii) inaccurate timestamps $t^r_{C_0}$ or $t^r_{C_1}$.

Our proposed robust TDoA trilateration scheme is motivated by two observations. First, the bi-directional acoustic channels between two anchor nodes tend to be symmetric. As a result, inaccurate timestamps at the anchors typically imply larger pair-wise ranging errors between the anchors. Second, under proper deployment, the majority of anchor nodes that are in the range of a target device should have LOS paths. Therefore, for both sources of errors, if we can remove outlier TDoA estimates, trilateration using the remaining TDoA estimates would be more accurate. Next, we present the detailed procedure for outlier removal.

[1]We assume the timestamps of the local transmissions $t^r_{A_0}$ and $t^r_{B_1}$ are accurate due to high SNR and strong LOS component

### 3.3.1 Detection of Outlier Time Offset

For symmetric acoustic channels, $t^r_{B_0}$ and $t^r_{A_1}$ would err on the same side. From (3.6), we see that if $t^r_{B_0}$ and $t^r_{A_1}$ change by the same amount, the distance estimate between $A$ and $B$ would differ by twice the amount divided by the speed of sound in the medium. Since anchor locations are known, by comparing the ground truth distance and the estimated distance, we can eliminate time offset estimates from the pairs with large ranging errors. For subsequent discussion, we only consider anchor pairs with valid time offset estimates.

For $N$ beacons from different anchor nodes, there are $N$ sets of time offset estimate pairs. Each set consists of $N-1$ pairs of time offset estimates and takes a unique anchor node as the reference node. If there are still multiple sets available after removing all outlier time offsets, we select the best set, which has the smallest average of pairwise ranging errors.

The pseudocode of detecting outlier time offset is given in Algorithm 1. In Line 14, $numAnchorsReq$ is 3(4) for 2D(3D) localization. Besides, $NaN$ element does not count in $length()$ in Line 14 and $average()$ in Line 15.

### 3.3.2 Iterative Outlier TDoA Removal

Under the assumption that the majority of anchors have strong direct LOS paths to the target device, we aim to eliminate timestamps from "bad" anchor nodes. Such outlier detection (also anomaly detection) is achieved by identifying observations which do not conform to an expected pattern or other observations in a dataset. The procedures are described below.

Now we are in the position of deriving the location of the target device from the

**Algorithm 1** Detection of Outlier Time Offset

**Input:** timestamps received at anchor nodes
**Output:** the best set of time offset estimate pairs

```
procedure TimeOffsetEstimates({t^r_{i_j}})        ▷ ∀i, j ∈ the same set of anchors specified
    avgErrD ← ∞
    r ← NaN                                        ▷ best reference anchor ID
    for all i ∈ anchorIDs do                       ▷ as time reference
        for all j ∈ anchorIDs do
            D̂_ij ← Eq (3.6)                        ▷ ranging
            t̂o_ij ← Eq (3.4)                       ▷ time offset
            errD_ij ← abs(D̂_ij − D_ij)             ▷ D_ij is ground truth
            if errD_ij > rangingErrThr then
                errD_ij ← NaN
                t̂o_ij ← NaN
            end if
        end for
        if length({errD_ij}) ≥ numAnchorsReq then
            if average({errD_ij} < avgErrD then
                r ← i
            end if
        end if
    end for
                                                   ▷ return a set of time offset estimate pairs
    if r ≠ NaN then
        return {t̂o_rk}                             ▷ {k} ← {j} \ {NaN entries}
    else
        return ∅
    end if
end procedure
```

TDoA estimates. Many algorithms exist in the literature for TDoA trilateration [27, 28, 29, 30]. We formulate it as a non-linear least squares (NLLS) problem so that multiple measurements can be utilized. Let $A_1, A_2, \ldots, A_m$ be the set of anchor nodes whose beacons are received by the target device. Let $TDoA_{ij}$ be the TDoA estimate between anchors $A_i$ and $A_j$. For ease of presentation, we only consider one beacon from each anchor node, and all the time offsets are valid. The unknown target location $x$ can be determined by solving the following non-linear optimization problem:

$$\min_{x} \sum_{i,j=1,i\neq j}^{m} \{c \times TDoA_{ij} - (dist(A_i, x) - dist(A_j, x))\}^2, \tag{3.7}$$

where $dist(\cdot)$ is the Euclidean distance. The iterative Gauss-Newton algorithm [31] can be used to solve this problem.

Let $\hat{x}$ be the estimated location from (3.7). If $x$ is known, $dist(A_i, x) - dist(A_j, x)$ can be computed from the known locations of anchors $A_i, A_j$ and $x$. If the measured $TDoA_{ij}$ differs from this quantity by more than a threshold value, this implies that at least one of the beacon timestamps is erroneous. Not knowing $x$, we can use $\hat{x}$ as its approximation. Enumerating through all valid pairs, we count the number of times an anchor node contributes to "erroneous" TDoA estimation and remove the anchor node with the highest count from the list. This procedure repeats with the remaining anchors until only 3 (4) anchors remain for 2D (3D) localization or all TDoA estimates have small errors. Clearly, such a procedure is not guaranteed to find all outliers. However, our experimental study shows that it can indeed improve the localization accuracy.

The pseudocode of the iterative outlier TDoA removal is given in Algorithm 2.

*Permute* in Line 7 operates in this way:

$$TDoA_{ij} \leftarrow TDoA_{rj} - TDoA_{ri}, \forall i, j \in \{k\} \tag{3.8}$$

In Line 18, $d(\cdot)$ is the Euclidean distance.

**Algorithm 2** Iterative Outlier TDoA Removal

**Input:** (1) timestamps received at a target node (2) corresponding time offset estimate pairs from Algorithm 1 **Output:** a location estimate of the target

```
1: procedure LocationEstimate({t^r_{C_i}}, {t̂o_{rk}})  ▷ C is target, ∀ i ∈ A, where A
   is the set of anchors from which beacons are received at C
2:     if {t̂o_{rk}} = ∅ then
3:         return NaN
4:     end if
5:     B ← {k}
6:     {TDoA_{rk}} ← Eq (3.5)                                  ▷ r ∈ B, ∀k ∈ B
7:     {TDoA_{ij}} ← Permute({TDoA_{rk}})                      ▷ ∀i, j ∈ B
8:     {ddoa_{ij}} ← {c × TDoA_{ij}}
9:     isAllPairsGood ← False
10:    repeat
11:        if |B| < numAnchorsReq then
12:            return NaN
13:        end if
14:        x̂ ← Eq (3.7)
15:        for all i ∈ B do
16:            numBadPair_i ← 0
17:            for all j ∈ B, j ≠ i do
18:                errddoa_{ij} ← ddoa_{ij} − (d(i, x) − d(j, x))
19:                if errddoa_{ij} > ddoaErrThr then
20:                    numBadPair_i ← numBadPair_i + 1
21:                end if
22:            end for
23:        end for
24:        numMax ← 0 ; idx ← NaN
25:        for all i ∈ B do
26:            if numBadPair_i > numMax then
27:                numMax ← numBadPair_i ; idx ← i
28:            end if
29:        end for
30:        if idx ≠ NaN then
31:            B ← B \ idx
32:        else
33:            isAllPairsGood ← True
34:        end if
35:    until isAllPairsGood
36:    return x̂                                                ▷ return a location estimate
37: end procedure
```

# Chapter 4

# Implementation

In this chapter, we first introduce the hardware design of anchor nodes. Next, we discuss the acoustic beacon design including beacon message format, modulation, demodulation, etc. Finally, a beacon scheduler for multiple anchor nodes is presented.

## 4.1 Hardware Design

In ARABIS, target devices can be any smart phone equipped with a microphone that can receive acoustic beacons in the range of 17.5kHz – 21.5kHz. Anchor nodes are purposefully designed. To allow fast prototyping, we develop an acoustic board that can be attached to a Raspberry Pi 3 (RPI3), a single board computer that supports Linux systems (Figure 4.1). The hardware diagram of the acoustic board is shown in Figure 4.2. It consists of an audio CODEC IC (WM8731), voltage reference, MIC amplifier circuit, speaker amplifier circuit, microphone (INMP411) and speaker (ASE06008MR-LW150-R). The board performs two functions, 1) broadcasting acoustic beacon signals, and 2) recording beacon signals both from itself and from other


Figure 4.1: ARABIS Acoustic Board

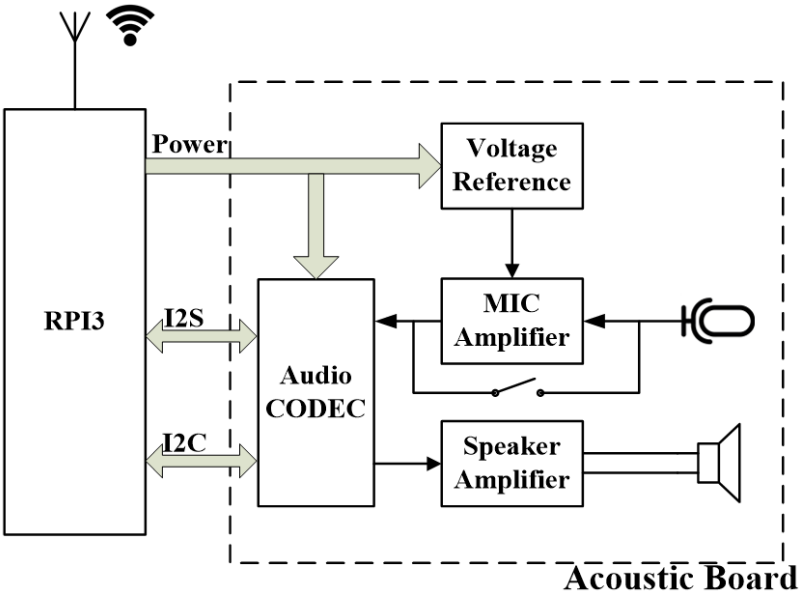


Figure 4.2: Anchor node block diagram.

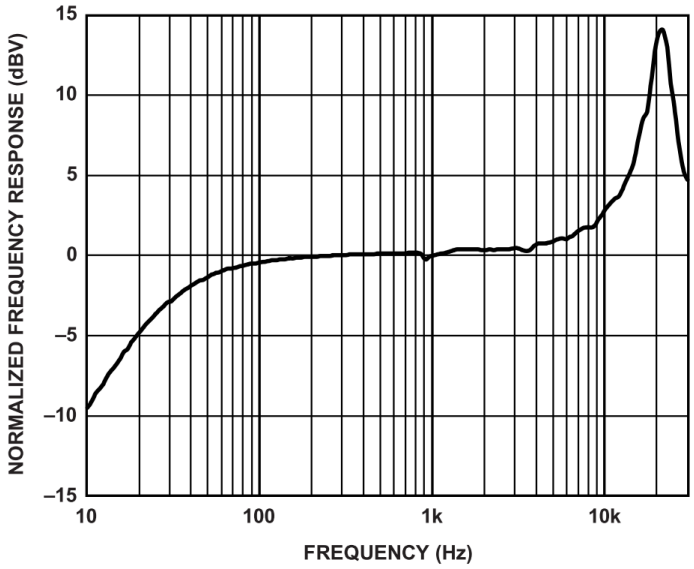


Figure 4.3: Typical frequency response of INMP411.

anchor nodes. The codec IC is controlled by RPI3 via I2C interface, and transfers to/from RPI3 the digital streams via I2S protocol.

To decode the beacon signal in the range of 17.5kHz – 21.5kHz, we chose INMP411 for the microphone [32]. INMP411 is a high performance, high sound pressure level (SPL), low noise, low power, and analog output bottom ported MEMS microphone. Its sensitivity specification makes it an excellent choice for near-field applications. The INMP411 has a linear response up to 131 dB SPL. Its normalized frequency response curve (Figure 4.3) shows the gain reaches nearly 15dBV around 20 kHz compared to 0 dBV in the range of 100 Hz to 3 kHz. Therefore, background audible sounds would not cause too much interference to beacon signals.

The selection of speaker is limited by the design requirement. ASE06008MR-LW150-R is chosen for its high output power and uniform gain in the operational frequency range [33] as shown in Figure 4.4. It is a 60mm enclosed speaker with 3W continuous, 4W maximum power handling and 8Ω impedance. In addition, its

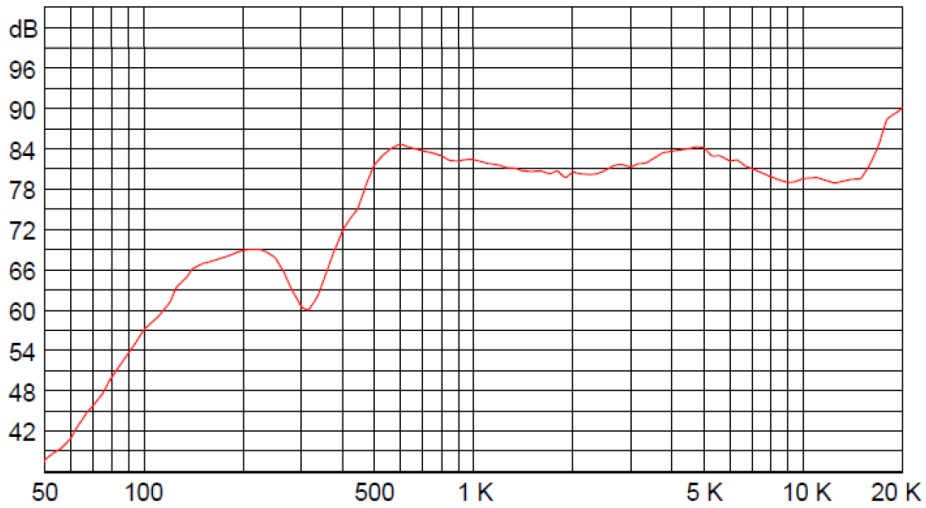


Figure 4.4: Frequency response of ASE06008MR-LW150-R measured at 0.5 meters away with 1 watt output power.

frequency response performance in the range of 17 kHz to 20 kHz is flat and can almost reach the peak of 90 dB.

## 4.2 Design of Acoustic Beacons

In this section, we present the detailed design of the waveform, modulation and demodulation, and message format of the beacon messages. The key considerations are, i) audibility of the signal, ii) message length, iii) preamble detection under low signal-to-noise ratio (SNR), and iv) timing extraction.

### 4.2.1 Beacon Message Format

To reduce beacon transmission time, we keep beacon messages as short as possible. As shown in Figure 4.5, the total length of beacon messages is 12-bits long with 4-bit identifier, 4-bit sequence number, and 4-bit CRC. A 4-bit identifier limits the

| Field | Preamble | ID | Seq | CRC |
|---|---|---|---|---|
| # of bits | - | 4 | 4 | 4 |
| Frequency | 21.5~17.5 KHz | 17.5Khz ~ 21.5KHz | | |
| Duration | 40ms | 30 ms / 2 bits | | |

Figure 4.5: Beacon message format.

maximum number of anchors in the system. If more anchors are required, more bits should be allocated in this field. With cycling sequence number, overflow is not a problem. A short sequence number requires urgent packet processing and we find 4 bits a good balance for the packet queues on the server. CRC-CCITT-4 is used in the 4-bit CRC. It will detect the following types of errors in a message [34]: (1) any one bit in error, (2) any two bits in error, (3) any odd number of bits in error, and (4) a burst of errors no longer than 4 bits. All other types of errors will be detected with the probability of $1 - 1/2^4 = 15/16$.

Each beacon message is preceded by a preamble. Preambles are used in data communication to extract clock information for coherent detection and to detect the start of data symbols. For TDoA estimation, preambles also serve the purpose of determining the arrival time of acoustic signals.

At the receiver side, if a message does not pass the CRC test, it will be discarded.

### 4.2.2 Waveform and Modulation

To determine the suitable frequency range of the acoustic signals used to transmit beacon messages, we have experimentally measured the channel response of smartphone microphones. Since Android OS continues to dominate the smartphone OS market

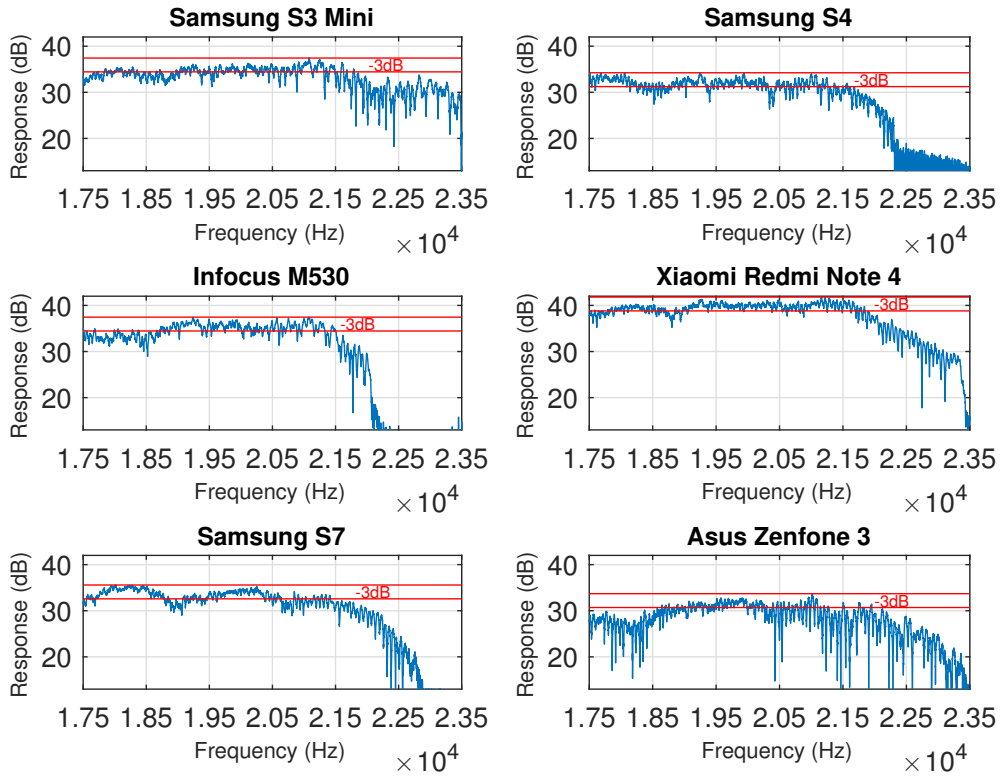


Figure 4.6: Microphone frequency response in 15.5kHz ∼ 23.5kHz of different mobile devices.

share (e.g., at 82.8% in 2015 Q2), it is reasonable to focus on Android devices. Tests with several human subjects show that acoustic signals above 17.5kHz are generally inaudible. Figure 4.3 and Figure 4.4 show the frequency response of the microphone and the speaker of the anchor node. Similarly, we test the frequency response of channels from the speaker on anchor nodes to the microphone on different phone models using a signal with flat frequency spectrum between 17.5kHz and 23.5kHz. From Figure 4.6, we do not observe a huge response decay for frequencies under 21.5kHz. As such, we choose the operating frequency range to be 17.5kHz – 21.5kHz.

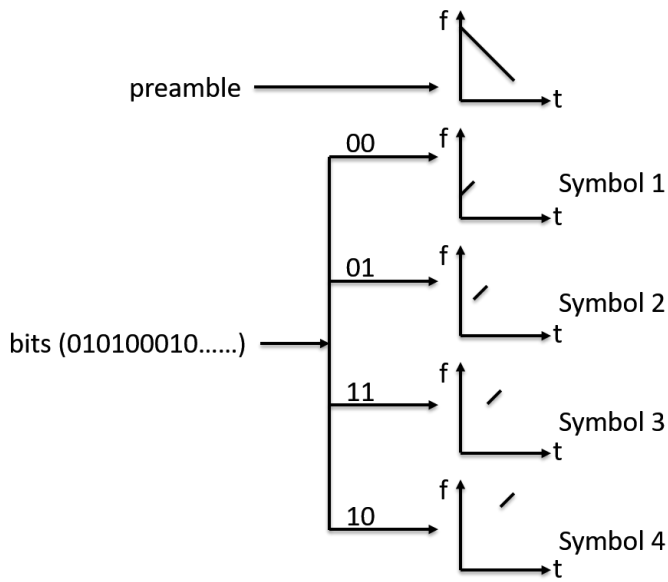


Figure 4.7: Linear chirp modulation of preamble and data symbols.

We adopt a linear chirp waveform in acoustic signal modulation for its nice property of pulse compression, which produces good ranging resolutions [10, 1]. Figure 4.7 illustrates the proposed modulation scheme.

A linear chirp waveform is mathematically represented by

$$x(t) = sin\left(\phi_0 + 2\pi(f_0 t + \frac{k}{2}t^2)\right),$$

and the instantaneous frequency $f(t)$ that varies linearly with time is given by

$$f(t) = f_0 + kt,$$

where $f_0$ is the starting frequency (at time $t = 0$), $k$ is the rate of frequency increase or chirp rate, and $t$ ends at $T$, which is the duration of the waveform. When $k$ is positive (negative), the chirp waveform is called an up (down) chirp. In our design, each data

symbol consists of 2 bits. We can alternatively represent the linear chirps using the starting frequency $f_0$ and ending frequency as shown in Table 4.1. An up chirp bears little similarity to a down chirp and hence has less interference in detecting a down chirp. Therefore we choose a down chirp for preamble and up chirps for data symbols to improve the preamble and data symbol detection rates.

Table 4.1: Chirp modulation parameters.

| **Symbol type** | **Starting freq.** | **Ending freq.** | **Duration (ms)** |
|---|---|---|---|
| Preamble | 21500 | 17500 | 40 |
| 00 | 17500 | 18500 | 30 |
| 01 | 18500 | 19500 | 30 |
| 11 | 19500 | 20500 | 30 |
| 10 | 20500 | 21500 | 30 |

### 4.2.3 Preamble Detection

Preamble detection is crucial for ranging measurements. It is important not only to detect the presence of preambles but also to extract the timing information accurately. The pulse compression property of chirp signals is achieved by correlating the received signal with the transmitted pulse.

To detect the preamble, we move a sliding window of duration $T$ along the received signal and calculate the normalized cross-correlation of the windowed signal with the stored samples of the preamble. If the resulting value exceeds a threshold, then a preamble is detected. Upon detection of the preamble, we continue sliding the window by $T$ and find the position where the maximum normalized cross-correlation is attained. To accelerate the peak finding process, the digital samples are stored in a buffer and searched block-wise.

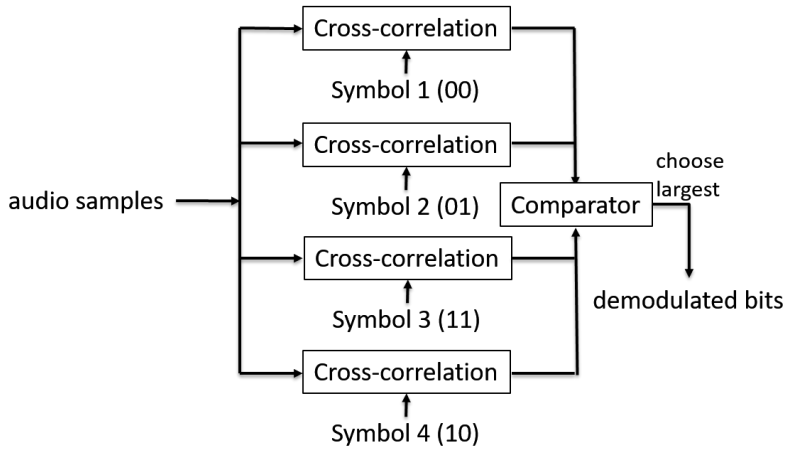


Figure 4.8: Block diagram of signal demodulation.

### 4.2.4 Demodulation

After the preamble is identified, the position of symbols in the beacon message is obtained. Demodulation of the beacon message is then performed using cross-correlation as shown in Figure 4.8. Normalized cross-correlation is not necessary in demodulating data symbols because all symbols have the same energy and the matched filter compares the outputs from cross-correlation between the same input signal and different symbols.

### 4.2.5 Mitigation of Audible Noise

Even though the instantaneous frequency of the transmitted signal is bandpass limited to the frequency range of 17.5kHz – 21.5kHz, the playback of the sound may still have audible components. Next, we discuss the contributing factors and the proposed solutions to mitigate the audible noise.

- **Volume:** High amplitude signals may saturate the speakers and introduce artifacts in the transmitted sounds. Furthermore, we observe some low volume

audible noise even when transmitting a single sinusoidal wave of high volume. Both can be mitigated by reducing the volume of transmitted acoustic signals.

- **Sharp amplitude changes:** At the start and the end of the transmissions, audible noise can be observed. This is caused by a phenomenon called speaker diaphragm inertia [35] and has been reported by several other works [10, 1, 13]. Essentially, there is a delay for speaker diaphragms to change from stationary to vibration states and vice versa. This effect can be mitigated by adding fade-in and fade-out signals before and after the transmitted signal to allow slow increment and decrement of acoustic signals. Longer fading duration mitigates noise more but compromises data transmission rate. In our experience, 10ms is a good choice and longer fading length does not further reduce noise for the speaker on our acoustic board. When reducing the fading length gradually to 5ms, the level of noise becomes more noticeable but is still acceptable. A further audio perception user study about the fading length can be found in [1].

- **Sharp phase changes:** Sharp phase transitions between symbols can also introduce unintended low frequency components. In our design, this issue is resolved by choosing the symbol duration according to the frequency range such that the start and the ending phase of each modulated symbol are 0 and $2\pi$, respectively.

## 4.3 Beacon Scheduler of the Localization System

Target nodes need to receive multiple beacons from different anchor nodes to localize themselves. Collisions among beacons result in not only the delay of localization

but also a less accurate result. Therefore, a beacon scheduler is necessary to avoid beacon collisions. The frequency of location updates is limited by the length of beacon interval. The shorter the interval is, the more frequent the location update is.

To provide timely location, the scheduler adopts time division multiple access (TDMA) for MAC with a high duty cycle. In TDMA, anchors take turns in transmitting beacons. To realize TDMA, clock synchronization is needed. Given that all anchors communicate with the location server, we set up a network time protocol (NTP) server on the same server and NTP clients on all anchor nodes. Within a wireless local area network (WLAN), NTP time synchronization accuracy is typically below or around 1-millisecond-level. This is sufficient for TDMA as we can include only a small guard interval for TDMA time slots. Note that for acoustic waves, 1 millisecond difference in the arrival time leads to 34.3cm difference in distance at 20 °C and hence we should not use NTP clocks as the synchronized clocks in Section 3.2.

To achieve high TDMA duty cycle, time slot length needs to be shortened to as close to beacon duration as possible. However, there are delays and jitters between tasks which worsen the performance. First, there are TDMA scheduling delay and jitter, which come from a timer trigger of the beginning of a time slot. Second, between the timer trigger and a beacon playing command, there are processing delay and jitter caused by other tasks. Third, between the beacon playing command and the actual time that the speaker starts playing, there are system delays and jitters from user space to kernel, and hardware delay of the audio interface. Next, when the beacon is broadcast, there are propagation delay and jitter depending on how wide the space is. Finally, if there are other unfinished tasks before the next time slot, they are also considered as processing delay and jitter. In a time slot, only jitters should

be considered. Deterministic delays do not have jitters and hence do not require extra time reserved for them in time slots. Among the above jitters, system jitter and propagation jitter are inevitable and cannot be shortened. However, other delays can be mitigated or even removed by optimizing the beacon scheduler.

For real-time operating systems (RTOS) such as Contiki, it is possible to schedule time-critical tasks with stringent timing requirements so the above jitters can be mitigated. However, the Linux distribution running on the Raspberry Pi is a non-real time OS. To reduce the scheduling jitter and the first part of processing jitter, we leverage multiple cores on the Raspberry Pi and use a dedicated thread with the highest scheduling priority to emulate a timer interrupt service routine (ISR) of a microcontroller (MCU). Hence, we can minimize the jitters between the start of the beacon time slot and the actual beacon transmission. Furthermore, other time-consuming tasks (microphone recording, acoustic signal decoding, beacon playing, etc.) are carefully scheduled so they can always be processed with the rest of the computation resources and do not interfere with the timer thread. In other words, the second part of processing jitter is removed.

The maximum propagation jitter depends on the largest distance between anchors and the room temperature. Take 15 meters and 20 °C for example, the propagation jitter is:

$$15(m) \div 343(m/s) \approx 43.7(ms) \tag{4.1}$$

Adding all jitters and the guard interval, we reserve 100ms and the beacon transmission duration in a time slot.

To this end, we summarize the key parameters of the beacon implementation in Table 4.2. The choice of preamble detection threshold will become clear in Section 5.1.

Table 4.2: Parameter settings.

| **Parameter** | **Value** |
|---:|---|
| Number of bits per symbol | 2 bits |
| Symbol duration | 30ms |
| Preamble duration | 40ms |
| Fade-in & fade-out duration | 10ms |
| Preamble detection threshold | 0.30 |
| Instant freq. range | 17.5 – 21.5kHz |
| Beacon message length | 1.5 bytes |
| Beacon transmission time | 0.4 sec |
| TDMA time slot length | 0.5 sec |

# Chapter 5

# Evaluation

In this chapter, we first provide numerical analysis of waveform design and then evaluate the proposed algorithms using two experimental testbeds. We are concerned with three aspects of the performance, namely, 1) accuracy and effectiveness of the algorithms, 2) whether there exists system bias in the localization results, and 3) impact of the number of anchor nodes used.

## 5.1 Comparison of Waveform Designs

In this section, we first evaluate the waveforms used for preambles and data symbols using numerical analysis and simulation study. It helps us to determine the threshold parameter in preamble detection (Section 4.2.3). For comparison purposes, we also implement the chirp modulation scheme in [1] with parameters listed in Table 5.1.

In comparison with Table 4.1, it is clear that the main difference is that the chirps for all the data symbols in [1] share the same instantaneous frequency ranges albeit with different chirp rates. In contrast, in the proposed chirp modulation schemes,

Table 5.1: Parameters of Chirp Modulation in [1].

| **Symbol type** | **Starting freq.** | **Mid freq.** | **Ending freq.** |
|---|---|---|---|
| Preamble | 21500 | - | 17500 |
| 00 | 17500 | 18300 | 21500 |
| 01 | 17500 | 19100 | 21500 |
| 10 | 17500 | 19900 | 21500 |
| 11 | 17500 | 20700 | 21500 |

we use different sub-channels to modulate different symbols. Table 5.2 shows the normalized cross-correlation among data symbols and the preamble with the scheme in [1] and the proposed one, respectively.

Table 5.2: Normalized cross-correlation among the chirp symbols in [1] and the proposed chirp symbols.

| normxcorr | scheme | preamble | 00 | 01 | 10 | 11 |
|---|---|---|---|---|---|---|
| preamble | symbols in [1] | 1 | 0.0824 | 0.0769 | 0.0763 | 0.0832 |
| | *proposed symbols* | *1* | *0.0874* | *0.0826* | *0.0853* | *0.0831* |
| 00 | symbols in [1] | 0.0824 | 1 | 0.2577 | 0.2180 | 0.3771 |
| | *proposed symbols* | *0.0874* | *1* | *0.0757* | *0.0038* | *0.0075* |
| 01 | symbols in [1] | 0.0769 | 0.2577 | 1 | 0.2939 | 0.2101 |
| | *proposed symbols* | *0.0826* | *0.0757* | *1* | *0.0078* | *0.0757* |
| 10 | symbols in [1] | 0.0763 | 0.2180 | 0.2939 | 1 | 0.2722 |
| | *proposed symbols* | *0.0853* | *0.0038* | *0.0078* | *1* | *0.0758* |
| 11 | symbols in [1] | 0.0832 | 0.3771 | 0.2101 | 0.2722 | 1 |
| | *proposed symbols* | *0.0831* | *0.0075* | *0.0757* | *0.0758* | *1* |

As evident from Table 5.2, the proposed modulation scheme has larger separations among data symbols. This would result in better data symbol decoding performance even under low SNRs.

The effect of the modulation schemes on data symbol decoding is further demonstrated in Figure 5.1. In this Matlab simulation, we randomly generate bit sequences

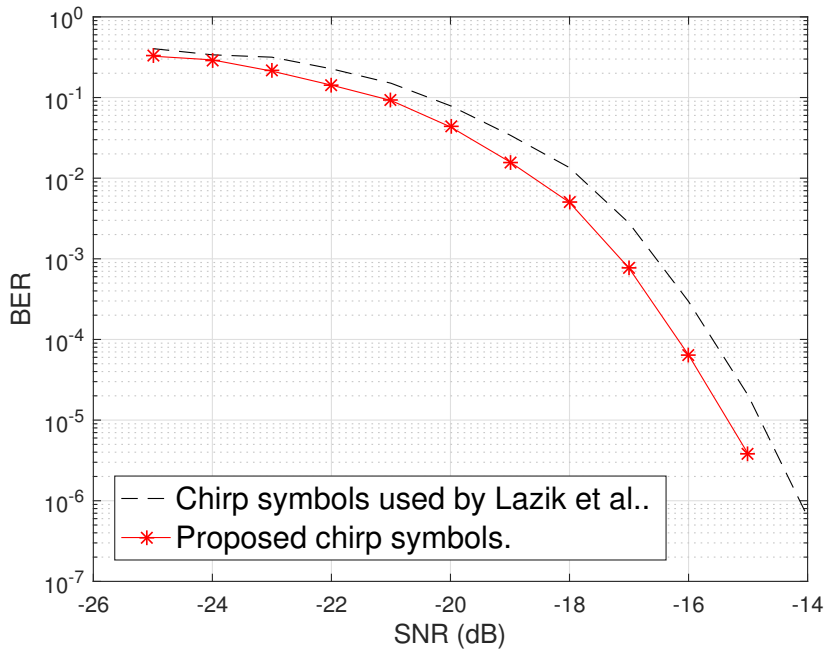


Figure 5.1: Simulated BER for different chirp modulation in AWGN channel with frequency range from 17.5kHz to 21.5kHz.

and map them to data symbols. Each data symbol is modulated using the proposed chirp modulation scheme and the one in [1]. The signal is passed through an additive-white-Gaussian-noise (AWGN) channel and demodulated using a matched filter decoder (Section 4.2.4). Figure 5.1 shows the bit-error-rate (BER) of the decoded data with respect to SNR using the two schemes. For the same BER, the proposed modulation scheme has around 1 dB gain compared to the approach in [1].


Figure 5.2: The first testbed: a general crowded office environment of size (10.67m × 7.76m)

## 5.2 Testbed Setup

The experimental testbeds consist of a location server, WiFi connectivity, and at least 4 anchor nodes. In the first testbed, as shown in Figure 5.2, eight anchors have been deployed near the ceiling of a 10.67m × 7.76m office space. The placement of the anchors can be found in Figure 5.7. For evaluation purposes, an NTP server runs on the location server to loosely synchronize the anchor nodes. The anchor nodes take turns according to a TDMA schedule of slot length 1 second in transmitting acoustic beacons periodically. The beacon transmission time is 0.4 seconds.

The second testbed was of size 15m × 15m and was deployed in a large convention center during the two-day Indoor Localization Competition in the IPSN'17 conference (Figure 5.4). As part of the requirements, only four anchors were allowed at the site[1]. To ensure sufficient coverage, external power amplifier and horn speakers were connected to the acoustic boards in Figure 5.3. Due to limited setup time, locations of

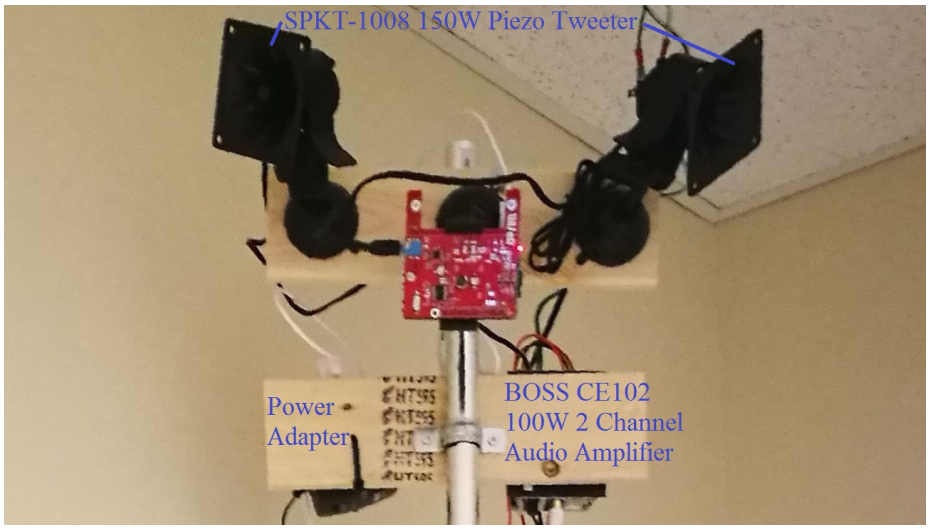


Figure 5.3: An anchor with external speakers used in the second testbed.

the anchor nodes were not optimized. A public WiFi network is used to communicate between the location server, the target mobile device, and the anchors. In both testbeds, a Samsung Galaxy S4 phone was used as the target device.

## 5.3 Indoor Localization Performance

For evaluation, we have implemented four variations of the proposed algorithm in Python on the location server. All variations apply the detection algorithm for outlier time offset to remove invalid beacons.

- Variation 1: All TDoA pairs, as shown in (3.7), without iterative outlier removal.

- Variation 2: Consecutive TDoA pairs without iterative outlier removal. In this scheme, we only consider TDoA pairs from beacons that arrive in ascending time

[1]A total of 8 anchors can be deployed but only 4 at each level of the two-floor space. Only evaluation results from the first floor are discussed as the second floor is of smaller size.

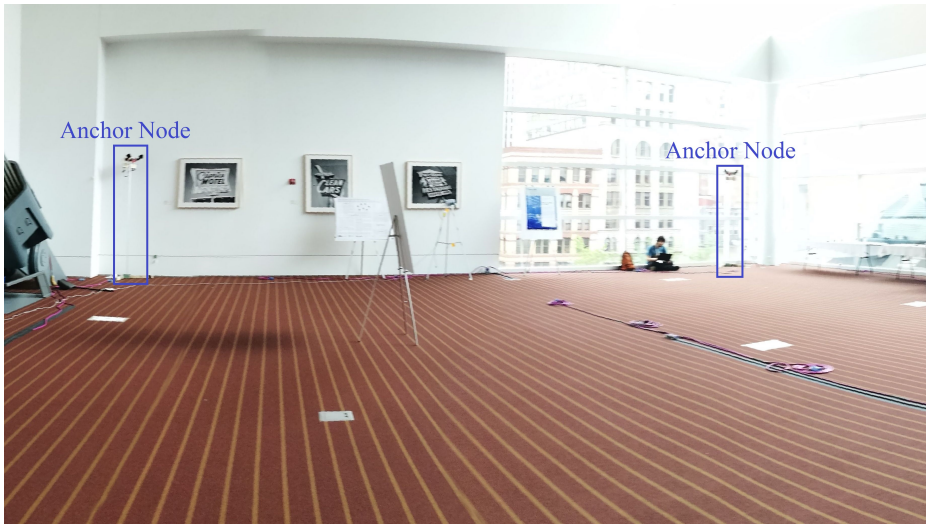


Figure 5.4: The second testbed: a larger and open-space environment of size about (15m × 15m) in a convention center

order at the target device in the time window. For instance, let $t_1 < t_2 < t_3 < t_4$ be the timestamps of beacon messages from anchors $A_1, A_2, A_3, A_4$. We only consider TDoA estimates of $TDoA_{12}$, $TDoA_{23}$, $TDoA_{34}$, and $TDoA_{41}$.

- Variation 3: All TDoA pairs with iterative outlier removal.
- Variation 4: Consecutive TDoA pairs with iterative outlier removal.

One location fix is computed using beacon messages received during a time window of 9 seconds (about twice the TDMA schedule length). In theory, 4 anchor nodes are sufficient for 3D localization. However, we found that since the anchors are deployed at similar heights, errors on the vertical $z$-axis are quite large. Therefore, only $x$, $y$ coordinates are reported. In solving the optimization problem in (3.7), we confine the feasibility region to the boundaries of the testbeds.

Figure 5.5 shows the Cumulative Distribution Function (CDF) of the location errors over 875 location fixes at six test locations in the first testbed. In this set

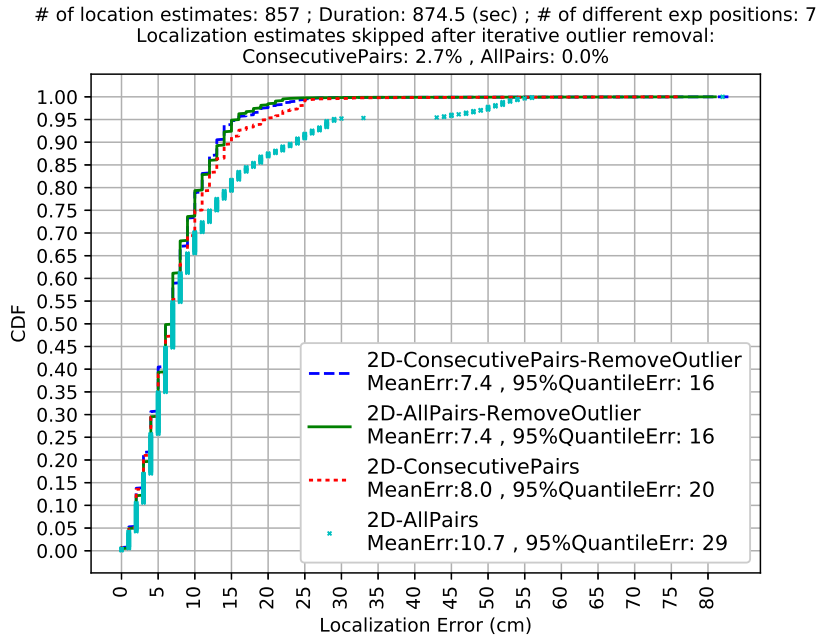


Figure 5.5: CDF of localization errors for different variations of the proposed algorithm in the first testbed.

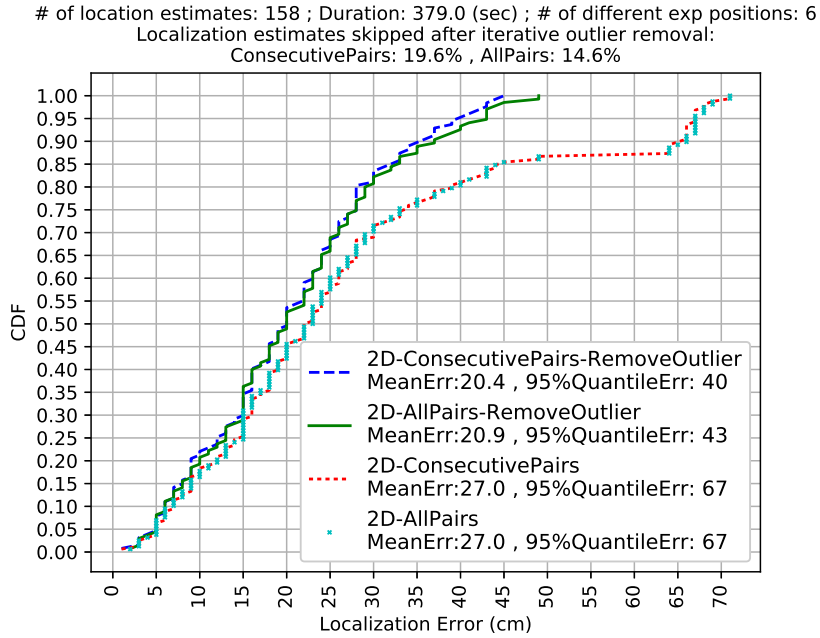


Figure 5.6: CDF of localization errors for different variations of the proposed algorithm in the second testbed.

of results, all 8 anchor nodes are used. The average localization errors for the four variations are 7.4cm, 7.4cm, 8.0cm and 10.7cm, respectively. The 95% quantile localization errors are 16cm, 16cm, 20cm and 29cm, respectively. A few observations can be made from Figure 5.5. First, iterative TDoA outlier removal can indeed improve localization accuracy. Its effect is more prominent when all pairs are used since with more pairs used, chances are that some provide inaccurate TDoA estimates. In the experiment, 2.7% of beacons are eliminated in the all-pair with outlier removal algorithm. Second, with iterative TDoA outlier removal, the performances of all-pair and consecutive-pair variations are comparable. This should come as little surprise. If all TDoA estimates are accurate, among eight beacons from eight anchor nodes, only seven TDoA estimates are linearly independent. Figure 5.6 gives the CDF of the location errors over 158 location fixes at five test locations in the second testbed. Similar observations can be made as in the first testbed. However, we see that the benefit of iterative TDoA outlier removal is more significant. As shown in Figure 5.6, in the absence of iterative TDoA outlier removal, location estimates in all-pair and consecutive-pair variations exhibit large errors – both have localization errors as large as 70cm! From both sets of results, we observe that Variation 3 has the best performance. Therefore, for subsequent evaluations, only Variation 3 is applied.

Figure 5.7 and Figure 5.8 are scatter plots of the localization results from the two testbeds, respectively. A bias is defined as the difference between the centroid of the location estimates of one test location and the ground truth. Biases can be attributed to many reasons such as NLOS paths, errors in ground truth location, and errors in anchor locations. The average biases are respectively, 7.34cm and 28.91cm in the two testbeds. We believe the larger bias in the second testbed is likely caused

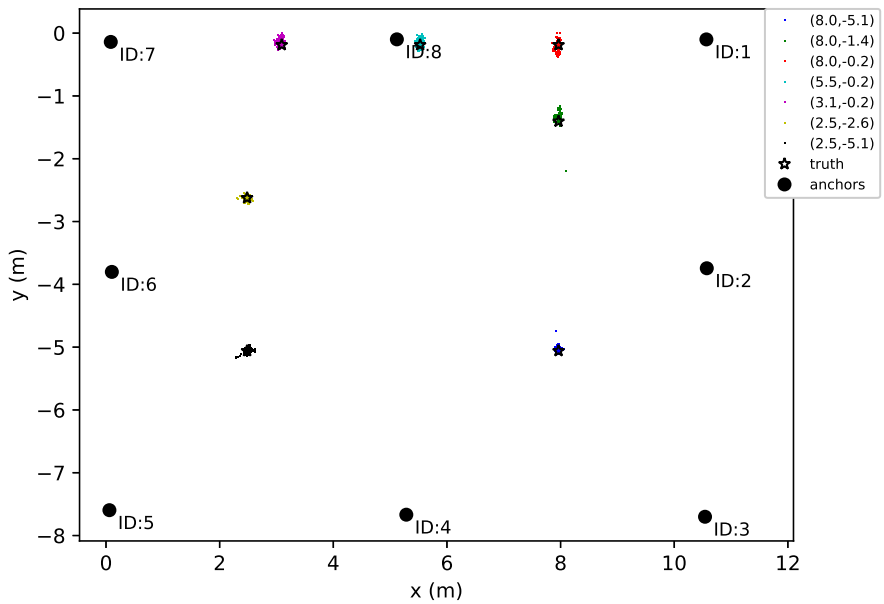


Figure 5.7: Scatters of localization result in the first testbed.

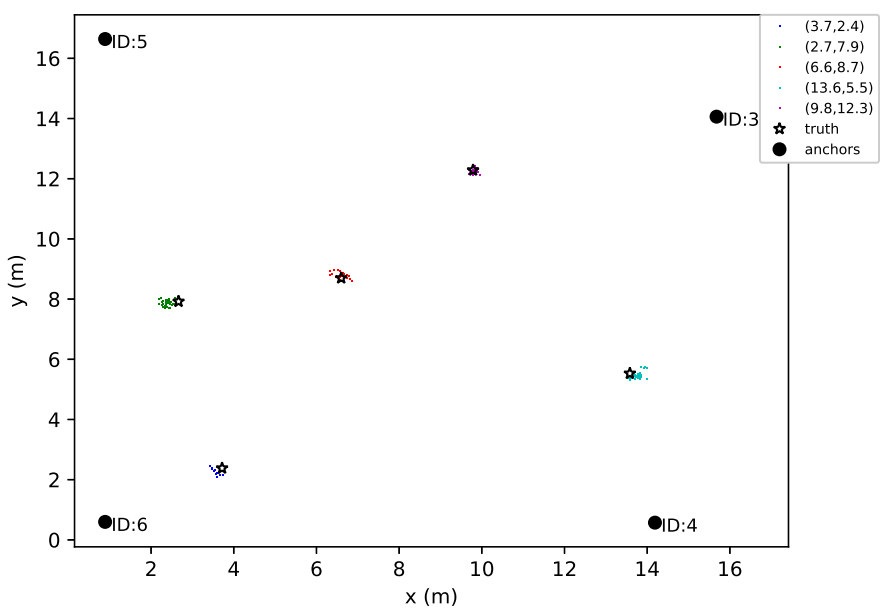


Figure 5.8: Scatters of localization result in the second testbed.

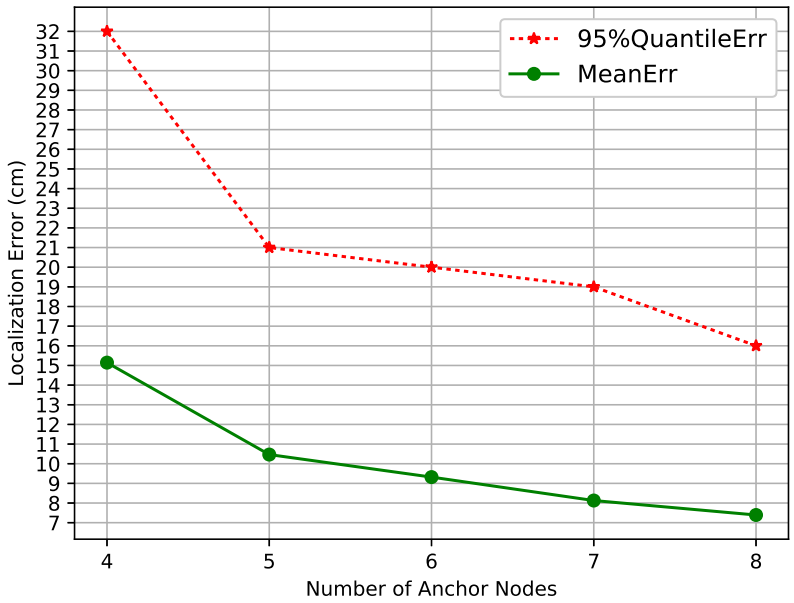


Figure 5.9: Localization errors for different number of anchors used in the first testbed.

by inaccurate measurement of the ground truth location, since we were holding a target device at the test locations as opposed to using a tripod as was done in the first testbed.

Table 5.3: Anchor ID information.

| # of anchors | anchor ID |
|---|---|
| 4 | 2,4,6,8 |
| 5 | 1,2,4,6,8 |
| 6 | 1,2,3,4,6,8 |
| 7 | 1,2,3,4,5,6,8 |
| 8 | 1,2,3,4,5,6,7,8 |

We also evaluated the impact of the number of anchor nodes used. The dataset collected from the first testbed was used by progressively removing the anchor nodes at the four corners and their respective beacon messages. The anchor ID information is

shown in Table 5.3. Figure 5.9 gives the average and 95% quantile localization errors. As expected, the more anchors used, the better the localization accuracy is. When there are only 4 anchors in the first testbed, the localization accuracy is comparable to the one in the second testbed. Furthermore, we observe while the performance degrades gradually from 8 to 5 anchors, there is a significant drop in localization accuracy when the number of anchors decreases to 4. This may be explained by the removal of one anchor that negatively affects the localization errors of the test sites at the upper half of the area.

Lastly, we show the comparison of 95% quantile localization errors with reported results in selected related works in Table 5.4. ARABIS outperforms all other works except the fact that Whistle [11] and ALPS [24] are evaluated in a slightly larger area. After connecting to two external speakers, which are still less powerful then the hardware used in ALPS [24], ARABIS in the second testbed still performs no worse than the previous two works but is evaluated in a much larger area.

Table 5.4: 95% quantile localization errors of selected approaches

| **Approaches** | **Scale ($m^2$)** | **Number of Anchors** | **Errors (cm)** |
|---|---|---|---|
| Wang et al. [14] | 1×1 | 4 | 7.5 |
| Guoguo [23] | 10×2 | 9 | 16.6 |
| Whistle [11] | 4×4 | 16 | 43.2 |
| Lazik et al. [13] | 4.5×5.5 | 3 | 10.9 |
| Lazik et al. [1] | 5×5<br>20×20 | 4<br>4 | 16.4<br>300.0 |
| ALPS [24] | 15×8 | 3 | 39.7[2] |
| our ARABIS [16] | 10.7×7.8<br>15×15 | 8<br>4 | 16.0[3]<br>40.0[4] |

[2]The error is reduced to 27.7cm with the combination of pedestrian dead reckoning (PDR).
[3]Each anchor uses one onboard speaker.
[4]Each anchor uses two external speakers.

# Chapter 6

# Conclusion

In this thesis, we proposed ARABIS, a robust and low-cost IPS using asynchronous acoustic beacons. It eliminates the need for tight clock synchronization among the anchor nodes. An extensible acoustic board is designed to support large operational ranges. With robust algorithms to remove outliers caused by low SNR and multi-path signals, experiments in two testbeds demonstrated the effectiveness of the proposed solution.

There are some limitations of ARABIS. First, both the detection of outlier time offset and iterative outlier TDoA removal rely on accurate locations of anchor nodes. Manually measuring the locations is labor-intensive and remeasuring is required when anchor deployment is changed. Though ARABIS can leverage the acoustic ranging capability of anchor nodes and apply rigidity theory [36] to automatically determine their locations, this is not implemented yet. Second, TDMA time slot length in the configuration is 0.5 seconds, which implies that targets need to stop for at least 2 seconds to get a location estimate. Once a target is moving, the iterative outlier TDoA removal approach is ineffective because recent TDoAs in the time window

become outliers and are not used until the target stops. Third, the single on-board speaker has narrow beam pattern which leads to low received SNR and difficulties when handling multiple paths. Although those problems are addressed in Chapter 3, we still find jitterish movement of stationary target nodes in noisy environment or at specific spots. As future work, we plan to devise a mechanism using a particle filter to handle those scenarios.

# Bibliography


[1] P. Lazik and A. Rowe, "Indoor pseudo-ranging of mobile devices using ultrasonic chirps," in *Proceedings of the 10th ACM Conference on Embedded Networked Sensor Systems (SenSys)*. ACM, 2012, pp. 99–112.

[2] eMarketer. (2016) Most smartphone owners use location-based services. [Online]. Available: https://www.emarketer.com/Article/Most-Smartphone-Owners-Use-Location-Based-Services/1013863

[3] D. Lymberopoulos, J. Liu, X. Yang, R. R. Choudhury, V. Handziski, and S. Sen, "A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned," in *Proceedings of the 14th International Conference on Information Processing in Sensor Networks (IPSN)*. ACM, 2015, pp. 178–189.

[4] Microsoft indoor localization competition - ipsn 2017. [Online]. Available: https://www.microsoft.com/en-us/research/event/microsoft-indoor-localization-competition-ipsn-2017/

[5] A. T. Mariakakis, S. Sen, J. Lee, and K.-H. Kim, "Sail: Single access point-based indoor localization," in *Proceedings of the 12th annual Mobile Systems, Applications, and Services (MobiSys)*. ACM, 2014, pp. 315–328.

[6] J. Xiong, K. Jamieson, and K. Sundaresan, "Synchronicity: Pushing the envelope of fine-grained localization with distributed mimo," in *Proceedings of the 1st ACM Workshop on Hot Topics in Wireless*. ACM, 2014, pp. 43–48.

[7] M. Leng, W. P. Tay, C. M. S. See, and S. G. Razul, "Gps-free localization using asynchronous beacons," in *Proceedings of the 8th International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*. IEEE, 2012, pp. 61–67.

[8] Microsoft indoor localization competition - ipsn 2015. [Online]. Available: https://www.microsoft.com/en-us/research/event/microsoft-indoor-localization-competition-ipsn-2015/

[9] C. V. Lopes, A. Haghighat, A. Mandal, T. Givargis, and P. Baldi, "Localization of off-the-shelf mobile devices using audible sound: architectures, protocols and performance assessment," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 10, no. 2, pp. 38–50, 2006.

[10] C. Peng, G. Shen, Y. Zhang, Y. Li, and K. Tan, "Beepbeep: a high accuracy acoustic ranging system using cots mobile devices," in *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems (SenSys)*. ACM, 2007, pp. 1–14.

[11] B. Xu, R. Yu, G. Sun, and Z. Yang, "Whistle: synchronization-free tdoa for localization," in *Proceedings of the 31st International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2011, pp. 760–769.

[12] V. Filonenko, C. Cullen, and J. D. Carswell, "Indoor positioning for smartphones using asynchronous ultrasound trilateration," *ISPRS International Journal of Geo-Information*, vol. 2, no. 3, pp. 598–620, 2013.

[13] P. Lazik, N. Rajagopal, B. Sinopoli, and A. Rowe, "Ultrasonic time synchronization and ranging on smartphones," in *Proceedings of the 21st IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS)*. IEEE, 2015, pp. 108–118.

[14] Y.-T. Wang, R. Zheng, and D. Zhao, "Towards zero-configuration indoor localization using asynchronous acoustic beacons," in *Proceedings of the 14th International Conference on Embedded and Ubiquitous Computing (EUC)*. IEEE, 2016, pp. 32–39.

[15] F. J. Álvarez, T. Aguilera, and R. López-Valcarce, "CDMA-based acoustic local positioning system for portable devices with multipath cancellation," *Digital Signal Processing*, vol. 62, pp. 38–51, 2017.

[16] Y.-T. Wang, J. Li, R. Zheng, and D. Zhao, "ARABIS: an asynchronous acoustic indoor positioning system for mobile devices," in *Proceedings of the 8th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2017.

[17] T. Sato, S. Nakamura, K. Terabayashi, M. Sugimoto, and H. Hashizume, "Design and implementation of a robust and real-time ultrasonic motion-capture system," in *Proceedings of the 2nd International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2011, pp. 1–6.

[18] I. Sonitor Technologies. (2016) Sonitor rtls technologies. [Online]. Available: http://www.sonitor.com/

[19] M. Minami, Y. Fukuju, K. Hirasawa, S. Yokoyama, M. Mizumachi, H. Morikawa, and T. Aoyama, "DOLPHIN: a practical approach for implementing a fully distributed indoor ultrasonic positioning system," in *Proceedings of the 6th International Conference on Ubiquitous Computing (UbiComp)*. Springer, 2004, pp. 347–365.

[20] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 6, pp. 1067–1080, 2007.

[21] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 13–32, 2009.

[22] I. Amundson and X. D. Koutsoukos, "A survey on localization for mobile wireless sensor networks," in *Mobile entity localization and tracking in GPS-less environnments*. Springer, 2009, pp. 235–254.

[23] K. Liu, X. Liu, and X. Li, "Guoguo: Enabling fine-grained indoor localization via smartphone," in *Proceeding of the 11th International Conference on Mobile Systems, Applications, and Services (MobiSys)*. ACM, 2013, pp. 235–248.

[24] P. Lazik, N. Rajagopal, O. Shih, B. Sinopoli, and A. Rowe, "ALPS: A bluetooth and ultrasound platform for mapping and localization," in *Proceedings of the*

*13th ACM Conference on Embedded Networked Sensor Systems (SenSys).* ACM, 2015, pp. 73–84.

[25] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The cricket location-support system," in *Proceedings of the 6th annual International Conference on Mobile Computing and Networking (Mobicom).* ACM, 2000, pp. 32–43.

[26] H. Schweinzer and M. Syafrudin, "LOSNUS: An ultrasonic system enabling high accuracy and secure tdoa locating of numerous devices," in *Proceedings of the 1st International Conference on Indoor Positioning and Indoor Navigation (IPIN).* IEEE, 2010, pp. 1–8.

[27] R. Bucher and D. Misra, "A synthesizable vhdl model of the exact solution for three-dimensional hyperbolic positioning system," *VLSI Design*, vol. 15, no. 2, pp. 507–520, 2002.

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

[29] K. Ho, "Bias reduction for an explicit solution of source localization using tdoa," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2101–2114, 2012.

[30] K. W. Lui, J. Zheng, and H. So, "Particle swarm optimization for time-difference-of-arrival based localization," in *Proceedings of the 15th European Conference on Signal Processing (EUSIPCO).* IEEE, 2007, pp. 414–417.

[31] A. Björck, *Numerical methods for least squares problems.* Siam, 1996.

[32] Inmp411 datasheet. [Online]. Available: https://www.invensense.com/wp-content/uploads/2015/02/INMP411.pdf

[33] Ase06008mr-lw150-r datasheet. [Online]. Available: http://www.puiaudio.com/pdf/ASE06008MR-LW150-R.pdf

[34] J. Stone, M. Greenwald, C. Partridge, and J. Hughes, "Performance of checksums and crcs over real data," *IEEE/ACM Transactions on Networking*, vol. 6, no. 5, pp. 529–543, 1998.

[35] "Loud speaker diaphragm," Sep. 22 1959, US Patent 2,905,260.

[36] T. Eren, D. K. Goldenberg, W. Whiteley, Y. R. Yang, A. S. Morse, B. D. Anderson, and P. Belhumeur, "Rigidity, computation, and randomization in network localization," in *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, vol. 4. IEEE, 2004, pp. 2673–2684.