

# Indoor Pseudo-ranging of Mobile Devices using Ultrasonic Chirps

Patrick Lazik      Anthony Rowe  
Electrical and Computer Engineering Department  
Carnegie Mellon University  
[{plazik,agr}@ece.cmu.edu](mailto:{plazik,agr}@ece.cmu.edu)

## Abstract

In this paper, we present an indoor ultrasonic location tracking system that can utilize off-the-shelf audio speakers (potentially already in place) to provide fine-grained indoor position data to modern mobile devices like smartphones and tablets. We design and evaluate a communication primitive based on rate-adaptive wide-band linear frequency modulated chirp pulses that utilizes the audio bandwidth just above the human hearing frequency range where mobile devices are still sensitive. Typically transmitting data, even outside of this range, introduces broadband human audible noises (clicks) due to the non-ideal impulse response of speakers. Unlike existing audio modulation schemes, our scheme is optimized based on psychoacoustic properties. For example, all tones exhibit slowly changing power-levels and gradual frequency changes so as to minimize human perceivable artifacts. Chirps also bring the benefit of Pulse Compression, which greatly improves ranging resolution and makes them resilient to both Doppler shifts as well as multi-path propagation that typically plague indoor environments. The scheme also supports the decoding of multiple unique identifier packets being transmitted simultaneously. By applying a Time-Difference-of-Arrival (TDOA) pseudo-ranging technique the mobile devices can localize themselves without tight out-of-band synchronization with the broadcasting infrastructure. This design is not only scalable with respect to the number of transmitters and tracked devices, but also improves user privacy since the mobile devices compute their positions locally. We show through user studies and experimentation on smartphones that we are able to provide sub-meter ( $95\% < 10\text{cm}$ ) accurate indoor positioning in a manner that is imperceptible to humans.

## Categories and Subject Descriptors

C.2.1 [Network Architecture and Design]: Wireless

Communication

## General Terms

Algorithms, Experimentation, Measurement

## Keywords

Localization, Sensor Networks, Acoustics

## 1 Introduction

Location tracking on mobile devices like smartphones has already begun to revolutionize personal navigation. Unfortunately, these services perform poorly indoors when GPS signals are no longer available. Highly accurate indoor location tracking would enhance a wide variety of applications including: building navigation (malls, factories, airports), augmented reality, location-aware pervasive computing, targeted advertising and social networking. In this paper, we propose an indoor location tracking system that can operate on existing smartphones using modulated Time-Difference-of-Arrival (TDOA) audio signals that are just outside the human range of hearing.

Studies have shown that humans can hear audio frequencies up to  $19 - 20\text{kHz}$  [1]. Microphones and Analog-to-Digital Converters (ADCs) on recent smartphones are capable of detecting frequencies as high as  $24\text{kHz}$ , leaving up to  $4\text{kHz}$  of bandwidth that can be used to transmit data. In practice, due to the non-ideal impulse response of speakers, humans can hear noises that fall into the lower frequency ranges when standard audio modulation approaches are shifted into these higher frequency ranges. These imperfections can be caused by frequency-delay phase inaccuracies, material resonance, energy stored in speaker drivers, and enclosure vibrations. In practice, if a speaker is instantly provided a high-amplitude tone, it will generate a loud clicking sound at both the start and end of the signal. One of the primary contributions of this paper is to provide a modulation mechanism that can send data utilizing these ultrasonic frequencies in such a manner that is imperceptible to humans. Furthermore, in order for the system to be scalable beyond just a few transmitters, we need to provide multiple-access support capable of both high-precision ranging at a high refresh rate, as well as the ability to send short sequences of data to identify transmitters. Our modulation technique is tailored to meet these specifications and has been implemented on real hardware. On the receiving side we have developed the means to demodulate, decode and perform ranging on our transmitted data. The increased processing capabilities found on current

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SenSys'12, November 6–9, 2012, Toronto, ON, Canada.  
Copyright © 2012 ACM 978-1-4503-1169-4 ...\$10.00

mobile devices make it possible to locally perform the required complex signal processing (similar to that of software defined radio) in real-time.

Our proposed approach uses a modulation scheme similar to Chirp Spread Spectrum (CSS). A chirp is a linearly frequency modulated pulse that increases or decreases over time between two frequency ranges. These waveforms are frequently used by the radar community due to an effect known as Pulse Compression. As discussed in Section 4.1, Pulse Compression is able to condense the width of a correlated signal (as compared to a sinusoidal signal) such that it provides tighter timing resolution and better Signal-to-Noise (SNR) ratios given the same amount of energy. In RADAR systems, this improves ranging resolution. The same approach can also be seen in nature. For example, many bat species will switch from generating constant frequency pulses to a form of chirp frequency modulation as they close in on their prey to enhance their ranging resolution [2]. As compared to modulation schemes like Direct Sequence Spread Spectrum (DSSS) and Frequency-Hopping Spread Spectrum (FHSS), chirps have gradual changes in frequency that produce fewer audible artifacts when played through speakers. In order to improve the amount of information contained in each chirp symbol, we employ a rate adaptation scheme on top of each chirp. This allows multiple speakers to simultaneously transmit over the same frequency band in a similar fashion as CDMA, which is essential for making a system that can scale across large indoor arenas.

In our system, unique signatures are broadcast from different speakers simultaneously. This is immediately applicable to theaters, PA systems, or concert venues that have surround sound systems already in place. A mobile device in the environment can record a short audio sample and then decode the TDOA of the various signals. Provided a mapping between each unique code and a physical location, the mobile device can localize itself in  $N$  dimensions assuming there are at least  $N + 1$  sources. Since the mobile device only knows the difference in arrival times with respect to each other and not an absolute point in time, this is called pseudo-ranging. A similar approach is used by GPS. This implies that the mobile devices do not need to synchronize (or explicitly communicate) with the infrastructure, which makes the approach highly scalable in the number of receivers and better for preserving privacy.

The main challenges associated with our system are:

- Efficiently utilizing a relatively small amount of audio bandwidth to accurately determine TDOA positions on mobile devices without additional hardware.
- Encoding a unique identifier into each signal such that data can be simultaneously transmitted in a scalable manner.
- Eliminating humanly perceivable audio artifacts (psychoacoustic properties).

## 1.1 Paper Organization

The paper is organized as follows. Section 2 discusses related work and how our approach differs. Section 3 describes

the use-cases and overall architecture of our system. Section 4 describes the details of our modulation primitive. Section 5 evaluates our system, Section 6 discusses its limitations, and Section 7 concludes and suggests future work.

## 2 Related Work

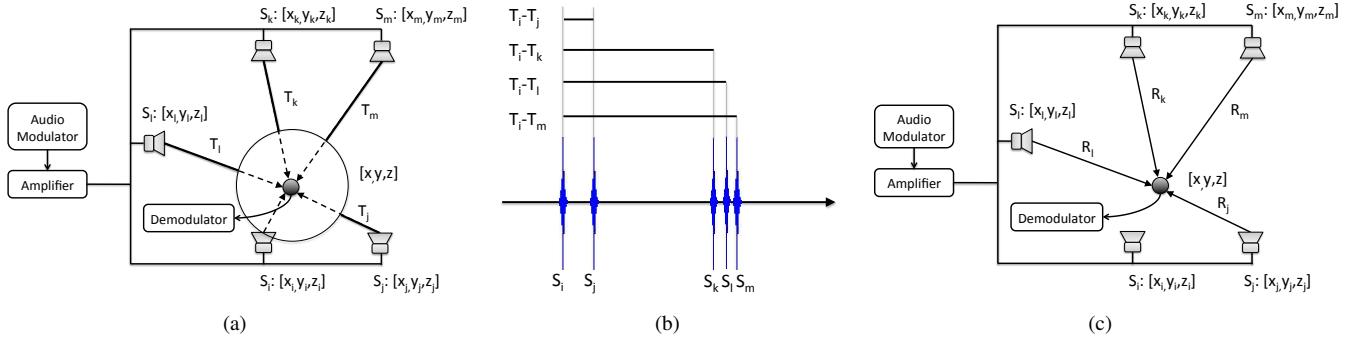
The large body of research on the topic of localization can be classified under the broad categories of range-based approaches [3, 4, 5] and range-free approaches [6, 7, 8, 9, 10]. Range-based approaches use measured distances or angular estimates between known anchor points to compute a position. Range-free approaches on the other hand typically attempt to match either synthetic or naturally occurring signatures to a particular location. It is also possible for a system to use a combination of both classes [11]. Location tracking systems all have trade-offs associated with their signaling technologies, line-of-sight requirements, levels of accuracy, cost and scalability. In this section, we will briefly touch upon a few related systems and focus more closely on those that are most similar to our approach. For a general overview, we refer to [12] for a more comprehensive survey.

### 2.1 Range-based positioning

Within the class of range-based localization approaches there are a few sub-categories based on how distances are computed. Time-of-Arrival (TOA) [13] systems compute distances based on how long it takes for a single signal type to propagate between a sender and receiver. For example [14, 15] both compute distances by measuring the Round-trip-Time-of-Flight (RTOF) by recording a signal's departure and the return time divided by the propagation speed. This assumes that the receiver will retransmit a return signal within a fixed amount of time. BeepBeep [15] uses this approach on cellular phones to compute inter-device ranges. At the time of publication, cellular phone microphones had significantly poorer frequency response as compared to what is available now, so the tones were well within the human audio range. The approach does not provide provisions for modulating data, computing absolute locations, supporting multiple-access and does not scale well beyond two devices.

Multiple systems have tried using communication mediums with two propagation velocities to measure a Time-Difference-of-Flight (TDOF). The MIT cricket [3] localization system uses the difference between RF and ultrasonic signals. The system is capable of extremely accurate range measurements, but requires line-of-sight communication, careful node positioning and high node density. It also requires tight synchronization between the infrastructure and the mobile devices. TOA and TDOF systems both require bi-directional coordination between the infrastructure and the device being tracked which generally limits scalability.

Time-Difference-of-Arrival (TDOA) systems can remove the requirement of knowing exactly when a signal was transmitted, using what is known as pseudo-ranging. As described in Section 3.1, pseudo-ranging computes distances by looking at the relative differences between the arrival of several signals, assuming they were all transmitted simultaneously. As compared to TOA and TDOF approaches, this requires one additional transmitter to allow the common distance from all broadcasting devices to be estimated. GPS



**Figure 1. System overview (a) with example signal detection (b) and final ranging distances (c).**

[4] is the most popular example of this ranging approach. Many modern cellular phones are equipped with GPS receivers, however they do not operate well indoors. The Dolphin [16, 17] system adopts a pseudo-ranging approach using ultrasonic modulation. Dolphin uses a  $50kHz$  carrier that is phase modulated by 511 bit Gold codes using DSSS. This is similar to our approach in that it uses pseudo-ranging, ultrasonic modulation and supports concurrent data transmissions. The major difference is that the system utilizes custom hardware that allows it to operate at much higher frequencies, making it naturally inaudible to humans. Based on our experimentation, using DSSS in the inaudible frequency band available on mobile devices would generate audible clicks between each frequency jump. Adding long fade-in and fade-out transitions between chips would severely impact throughput and makes such an approach impractical. Instead, we use rate-adaptive chirp symbols that are able to capture multiple bits of information with smooth frequency changes that do not create audio artifacts. Our system is also designed to work with standard speakers and ordinary mobile devices without additional hardware. [18] expands upon Dolphin (while still requiring custom hardware) by adding a self-training deployment approach based on filtered motion within the space. In the future we could apply a similar technique to our system to help simplify configuration and deployment.

In [19] the authors try to identify the location of a cellular phone in a car using ultrasonic pseudo-ranging from the car's audio speakers. The system identifies which seat within the vehicle the cellular phone user is occupying. The authors are only interested in ranging data within the car and use a tone-based TDMA system that does not have provisions to transmit additional data. This works well in a confined environment like a car, but limits the scalability of the approach to larger spaces. Simple fixed frequency tones generate audible noises that may not be noticeable in a car (especially if run for only a short period of time), but would be apparent in an environment like a museum or if left running continuously. In contrast, our chirps are tailored to be inaudible to humans and benefit from Pulse Compression, allowing for more precise timing resolution.

[20] is a study of cellular phones and their ultrasonic capabilities. The authors find that modern cellular phones are quite capable of transmitting and receiving ultrasonic tones

between  $20 - 22kHz$  with minimal distortion. We confirm their findings and also note that many phones have sampling rates as high as  $48kHz$ , providing a maximum frequency range of  $24kHz$ . We extend the concept of this work to show how one might modulate data and we show an initial evaluation of how these systems perform at TDOA ranging.

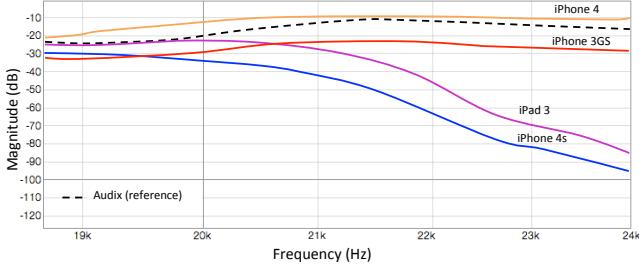
## 2.2 Range-free positioning

Signatures can be obtained from a variety of sources including (but not limited to) RF signals, background acoustics, magnetic fields and visual cues. These systems use pattern matching to help reduce issues like multi-path that cause problems in range-based systems. The RADAR [6] system was one of the first to use pre-recorded RF signatures from WiFi access points to build a database of signal strength values at particular locations. A mobile node receives signal strengths from the stationary beacons, which can then be compared with reference points to find the best match. The MoteTrack [8] system uses a similar approach with an emphasis on distributed operation in a sensor network. Environmental signatures such as background audio [9, 10] have been exploited to capitalize on the environmental regularity of many locations. For example, many rooms have characteristic noises that are created by machinery or ventilation patterns. In general, these approaches require a significant training phase and can lose accuracy over time due to subtle changes in the environment [11]. Though our work focuses primarily on the modulation primitive and TDOA ranging, we believe that ultrasonic fingerprints could also be an effective tool for augmenting signature-based approaches.

## 3 System Architecture

As shown in Figure 1, our system consists of a transmission infrastructure along with any number of mobile receiver devices. The transmitters are synchronized so that each speaker plays a unique audio signal (like surround sound speakers in a theater) simultaneously. Each audio signal represents a unique identification code. The ID code sound clip is simply replayed continuously in a loop. We envision that this could even be a simple hardware add-on that clips on near the speaker to mix in the audio sequence.

The required speakers could be part of an existing Public Announcement (PA) infrastructure or custom added to support localization applications. The receiver can be any mobile device that is able to record sound inside the  $19 - 24kHz$  frequency range. The frequency range is limited by the



**Figure 2. Frequency Response (1/12 Octave Smoothing)**

sampling rate of the audio hardware, along with filtering on the microphone’s analog front-end. Using a measurement microphone and a swept sine deconvolution technique commonly used by audio engineers to profile frequency responses, we measure the microphone response of various popular mobile devices above 19kHz. Figure 2 shows the frequency response for these devices (recording at 48kHz) as compared to a reference audio measurement microphone. The shape of the reference line is due to the frequency response of our test speakers. Since the microphone is calibrated to have a nearly flat frequency response, any line that follows its shape should share a similar characteristic.

### 3.1 Pseudo-ranging

In order to avoid requiring the mobile device to synchronize with the infrastructure, we use a TDOA pseudo-ranging technique. Each transmitter simultaneously sends an encoded identification message, each of which will arrive at the receiver at a slightly different time. As illustrated in Figure 1(b), each receiver is able to determine the relative timing (and ID) of each transmitter. We see the following relationships between distances  $R$  and arrival times  $T$  assuming a sound propagation speed of  $c$ :

$$\begin{aligned} R_i &= c * T_i \\ R_{ij} &= R_i - R_j \\ R_{ij} &= c * T_i - c * T_j \end{aligned} \quad (1)$$

Assuming the  $x, y, z$  coordinates of each speaker are known, we can setup the following system of equations:

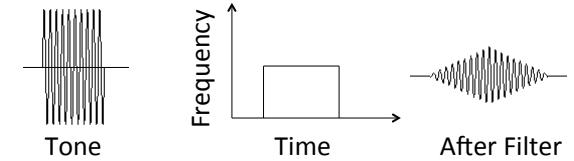
$$\begin{aligned} c * (T_i - T_j) &= \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} - \\ &\sqrt{(x_j - x)^2 + (y_j - y)^2 + (z_j - z)^2} \end{aligned} \quad (2)$$

$$\begin{aligned} c * (T_i - T_k) &= \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} - \\ &\sqrt{(x_k - x)^2 + (y_k - y)^2 + (z_k - z)^2} \end{aligned} \quad (3)$$

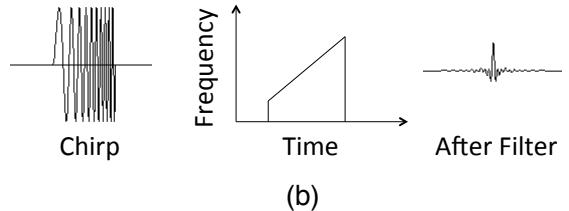
$$\begin{aligned} c * (T_i - T_l) &= \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} - \\ &\sqrt{(x_l - x)^2 + (y_l - y)^2 + (z_l - z)^2} \end{aligned} \quad (4)$$

With 4 receivers, we can construct the above system of 3 equations with 3 unknowns. We use the approach (and associated C code) from [21] to solve for the receiver’s  $x, y, z$  location.

Unlike in radio-based systems, the speed of sound depends on air temperature and is not always constant. As often mistakenly assumed, the speed of sound in air does not



(a)



(b)

**Figure 3. Pulse Compression Illustration**

depend on the amplitude, frequency, or wavelength of the audio signal, or the barometric pressure. Attenuation of sound, however does depend on the frequency range of the signal. We use the following formula to compute the speed of sound in  $m/s$ :

$$c = 331.3 + 0.606 * t \quad (5)$$

where  $t$  is the air temperature in degrees C. Many mobile devices are equipped with temperature sensors that can be used locally to calibrate the ranging measurements.

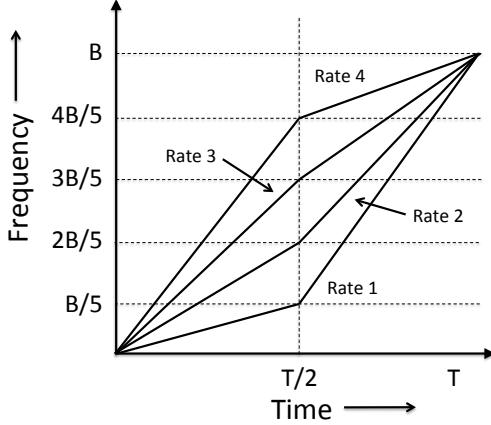
## 4 Ultrasonic Chirps

In order to send data and determine TOF ranging information, we require a modulation technique that provides precise timing resolution, resistance to multi-path fading, and can be easily distinguished from noise. Unlike communication systems that are typically focused on data throughput, the RADAR community has developed modulation schemes to increase sensitivity as well as timing resolution. A common approach is to use what is known as Pulse Compression on linear chirp signals.

We would also like a concurrent multiple access approach to simplify the tiling process of placing speakers across an indoor space with many overlapping domains (airports, malls, etc). A common technique in radio engineering for sending data is to use channel spreading techniques like DSSS. Unfortunately, most of these spreading techniques require rapid on-off transmission intervals as well as large frequency jumps that would generate audible artifacts when transmitted from speakers. In this section, we discuss the parameters and design of our ultrasonic chirp symbol that uses rate-adjusted Chirp Spread Spectrum (CSS) to concurrently transmit data and ranging information in an efficient manner.

### 4.1 Pulse Compression

Pulse Compression is a technique used in RADAR systems to increase range resolution as well as receiver sensitivity. When performing ranging using a standard sinusoidal pulse of constant frequency as a signal, the range resolution improves inversely proportional to the length of the pulse.



**Figure 4. Example of four chirp rates (2 bits per symbol)**

After correlating the reflected signal with the original tone waveform, a signal with a broad base similar to that in Figure 3(a) can be seen. The magnitude as well as the breadth of this signal increase proportionally to the length of the pulse, therefore increasing the received signal magnitude, but decreasing its range resolution. Pulse Compression on the other hand employs chirp waveforms that linearly increase (or decrease) in frequency as ranging signals. Now when the received signal is correlated with the original chirp, the width of the intercorrelated signals is smaller than what you would see from a standard sinusoidal pulse. Figure 3 shows an example of a tone and chirp before and after filtering. The peak value after filtering is identical, but the chirp appears to be compressed (hence the name Pulse Compression). This compression makes the signal simpler to detect as it effectively increases its SNR, which leads to lower amounts of timing jitter, hence improving the range resolution. The gain in SNR and the improvement in range resolution is given by the compression ratio of the chirp, which is equal to its time bandwidth product.

We now briefly summarize some of the key theoretical properties of Pulse Compression. A linear frequency modulation is described by the following equation:

$$s(t) = \sin(2\pi(f_c + \frac{k}{2}t)t) \quad (6)$$

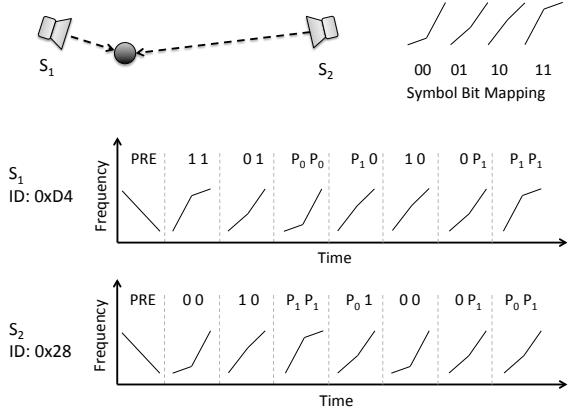
For  $0 \leq t \leq \tau$  where  $\tau$  is the pulse duration,  $k$  is the rate of frequency change,  $f_c$  is the starting frequency and  $t$  is time. The bandwidth  $\Delta f$  can be computed as:

$$\Delta f = k\tau^2 \quad (7)$$

The range resolution  $\rho$  can be computed as:

$$\rho = \frac{c}{2\Delta f} \quad (8)$$

where  $c$  is the propagation speed of the medium (in this case sound which is about  $340\text{ms}/\text{s}$ ). Given the  $4\text{kHz}$  of bandwidth available on mobile devices, the best range resolution we can expect in practice is  $4.25\text{cm}$ .



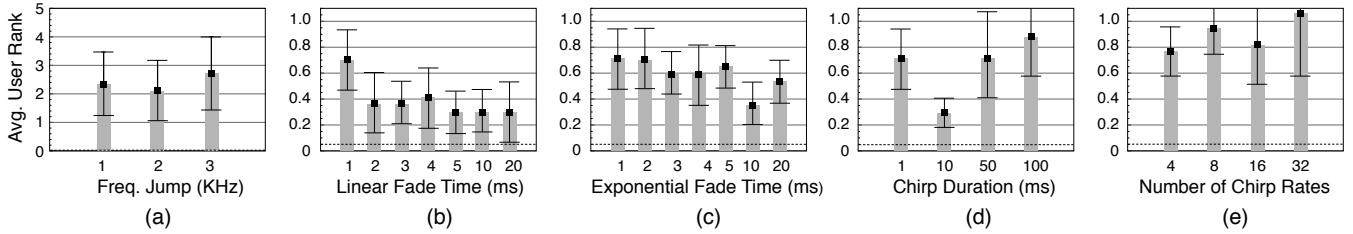
**Figure 5. ID transmissions**

## 4.2 Multiple Access

There are multiple ways to facilitate multiple-access transmissions using chirps including TDMA, FDMA and CSS. TDMA suffers from scalability and configuration issues since all chirps would need to be scheduled in a collision-free manner. Using frequency diversity isn't ideal since a chirp's timing resolution is directly related to the frequency bandwidth that it operates on. Ideally, you would like each chirp to cover the maximum bandwidth to achieve the highest ranging resolution. Chirp Spread Spectrum typically uses chipping codes composed of up-chirps and down-chirps to represent patterns of 1's and 0's. While promising, this approach can require long transmission times depending on the number of bits used in each code.

[22] introduces the use of chirp-rates as a mechanism to assign uniquely modulated chirp signals to users. This approach decomposes each chirp into two interconnected chirps with different frequency rates that change at the half-way point of the symbol. Figure 4 illustrates a scheme that supports four unique symbols across a shared bandwidth. Each rate represents a different signal waveform that is correlated with the received signal to extract the embedded sequences of data. We provide each transmitter with a unique ID, which is encoded as a series of up-chirps, each representing two bits. It is worth noting that [22] was based entirely on simulation. One contribution of this paper is a validation that such rate adaptation can work in practice.

Figure 5 shows a diagram where two transmitters are using our chirp modulation scheme along with chirp rate adaptation. Each transmitter ID is encoded as a sequence of two ( $7, 4$ ) Hamming codes, allowing us to transmit 256 unique IDs by using seven two-bit symbols. The error coding allows us to correct up to two single-bit errors and detect all single-bit, as well as two-bit errors. Furthermore, as a mobile device moves through a space, a map can be used to identify which transmitters are likely to be in range, allowing out-of-range IDs that were erroneously decoded to be discarded. Each data symbol is represented as a rate adapted up-chirp, and is prefixed by a preamble encoded as a constant-rate down-chirp. The preambles are used to mark the beginnings of data sequences and to measure high resolution TDOA information from. The modulation scheme can be easily adapted to



**Figure 6. Summary of audio perception user study**

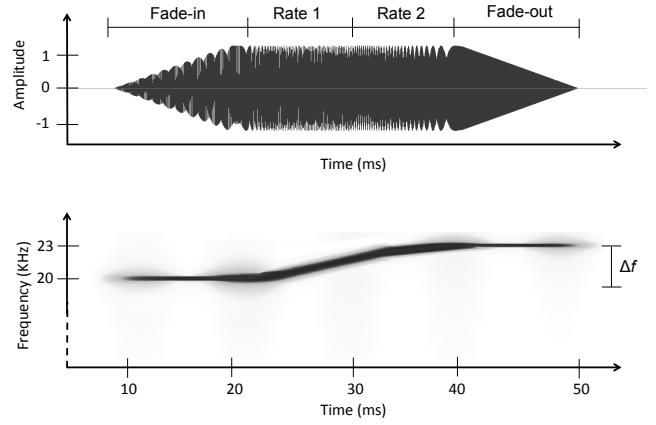
larger installations with more than 256 transmitters by employing a (15, 11) Hamming code and/or tiling transmitters. We argue that this approach is significantly more practical than centralized TDMA scheduling.

### 4.3 Receiving Data

The demodulation of the received signal is performed completely in software, in part by a process known as matched filtering. In matched filtering the incoming signal is convolved with a conjugated, time-reversed version of a signature signal that is expected to be contained within the received signal. This results in a distribution showing the similarity of both signals as they are slid across each other. Peaks of high magnitude denote a high cross correlation between both signals, therefore making it likely that an instance of the signature signal is located at the same location as the peak in the received signal. Therefore, by applying a matched filter for each rate adjusted chirp and the preamble, we are able to determine the starting locations of the signals as well as the time difference between them.

There are multiple ways of computing the matched filtering output. Running a matched filter in the time domain requires  $O(n \cdot m)$  operations, however when performed in the frequency domain, can be processed in  $O(n \cdot \log \cdot n)$  time. In the frequency domain, the FFT of the modulated signal is multiplied with a frequency domain representation of the signature signal and is then converted back into the time domain.

Since we are using sequences of multiple chirp rates as symbols to uniquely identify transmitters, the symbols need to be as orthogonal as possible (i.e. have low cross correlation properties between each other) in order to be differentiable after matched filtering. Rate adjusted chirps generally fulfill this requirement if the number of chirps within a set is kept to a reasonable number as can be seen in Section 5.2, but in practice we have found that the superposition of chirps of different rates that are staggered in time can become very difficult to detect. For this reason we prefix each data string with clearly identifiable preambles represented by down-chirps, which are highly orthogonal with respect to up-chirps. The preambles mark the beginning of the data sequences, which allows us to bound the region of signal that we perform matched-filtering on, and act as beacons for receivers to synchronize to a transmitter's broadcasts. Since each transmitter broadcasts identical sequences of data periodically, the ID of incoming data sequences can be predicted based on their arrival time with respect to a previous sequence. Therefore once a receiver is synchronized to a particular transmitter, TDOA ranging can be performed on



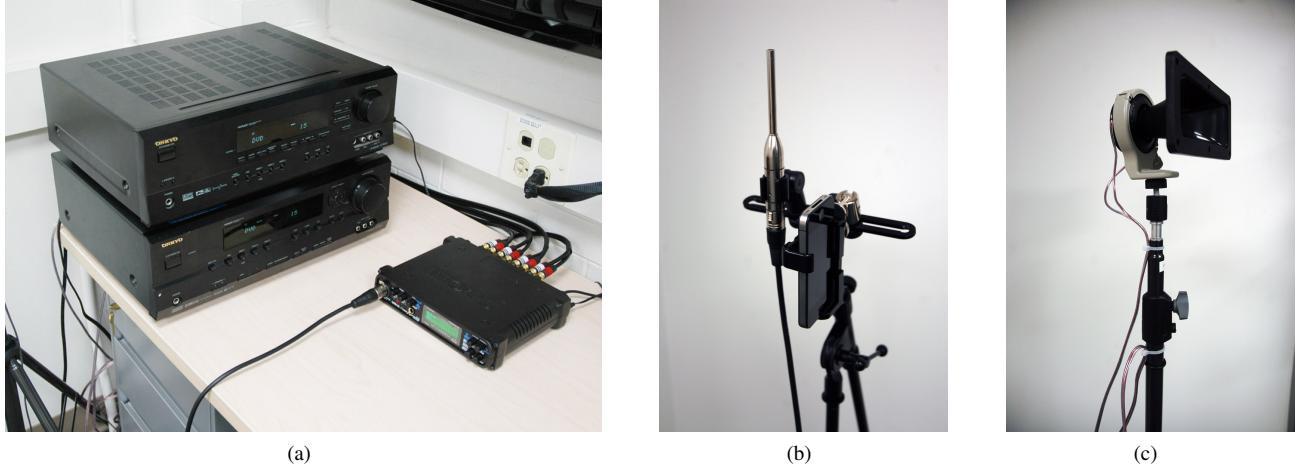
**Figure 7. Chirp Components**

each detected symbol, before the entire corresponding data sequence is decoded. This allows for significantly higher ranging update rates (but is not required). The time required to send an 8-bit ID is 240ms, but new ranging information can be obtained at up to every 30ms if successive broadcasts are transmitted in a gap-less fashion. The preambles also provide us with an estimate of the amplitude of the following data sequence. Since we know the location of the data symbols with respect to the preamble, we can now filter overlapping symbols according to their magnitude and position in time. In combination with the forward error correction, and discarding erroneously decoded transmitters that are likely to be out of range, we can achieve high packet reception rates as shown in Section 5.5.

Another technique that can be applied to help separate overlapping data sequences is Successive Interference Cancellation as described in [23]. Here the modulated signals of successfully decoded data sequences are reconstructed and then subtracted from the received signal in order of descending amplitude before any further decoding is performed. Furthermore the incorporation of a Rake Receiver as described in [24], or an Adaptive Matched Filter could improve robustness against multi-path interference.

### 4.4 Reducing Audible Artifacts

One of the main challenges associated with near sonic modulation over standard audio speakers is avoiding humanly perceivable artifacts. Since speakers are mechanical systems, they cannot instantly transition between gain settings without creating clicking noises. To alleviate these problems, our chirp signals require slow amplitude fade-in



**Figure 8. Test equipment (a) amplifier and DAC (b) measurement microphone and smartphone (c) piezo tweeter**

and fade-out changes, slow frequency changes and all adjustments are only made during zero-crossing points in the signal. Figure 7 shows the overall layout and various parameters associated with our chirp symbols. The spectrogram in the lower portion of the figure shows that the fades occur at a constant frequency followed by the two chirp rates, and then a fade-out at the highest fixed frequency. The chirp waveform that is used for correlation does not include the fade-in and fade-out periods since they interfere with the Pulse Compression.

In order to better understand the perceived effect of these attributes we conducted a user study where participants ranked the perceived loudness of different waveform configurations. The test was designed to evaluate the perceived audio artifacts associated with: (1) size of frequency jumps, (2) length of fade-in and fade-out, (3) linear versus exponential fades, (4) chirp duration, and (5) multiplexing of chirp rates. Participants were asked to watch a video that presented them with an intense click ( $70dB(A)$ ) as a loud reference and a soft click ( $10dB(A)$ ) as a low-level reference, followed by 35 different test sequences. Each test sequence was played three times per title slide in the video. All of the tests were randomized and some of the title slides contained no sound as a placebo value. Users would rank the intensity of the sound on a scale from 0 to 5 where 0 was considered silence and 5 was a loud sound. The experiments were conducted using high-end audio headphones connected to an external DAC at a fixed volume level. The histograms in Figure 8 summarize the results obtained from 35 users between the ages of 18 and 35. The dotted line at the bottom represents the average rating for the silent sequences. The error bars represent 1 standard deviation. Some of the tests included fixed tone with long fade-in sequences at different frequencies to act as a crude approximation of the user's frequency range. Our data showed a rapid fall-off around  $19kHz$ , which agrees with the standard hearing literature.

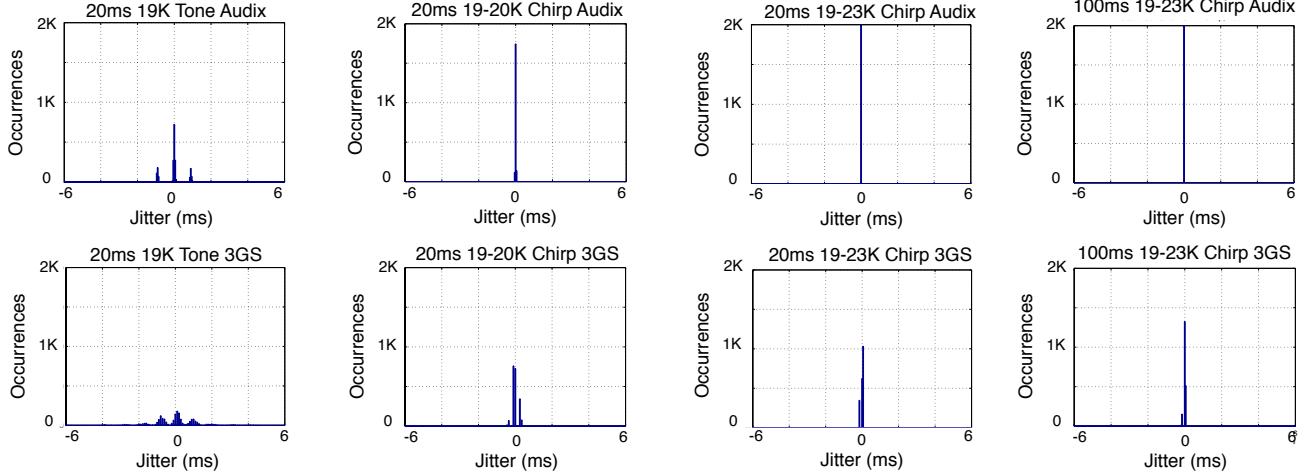
The first set of experiments was designed to determine if users could detect large frequency jumps. The test waveforms started off with a slow  $20ms$  fade-in to a fixed frequency tone between  $19$  and  $23kHz$ . The tone would run for

$20ms$  and then jump  $1,2$ , or  $3kHz$  to a higher fixed frequency followed by a slow fade-out. The slow fade lengths had earlier been selected since they were relatively unnoticeable. The idea was to simulate the types of frequency changes that would be apparent during Pseudo-random Noise (PN) DSSS modulation schemes. As can be seen in Figure 8 (a), even at  $1kHz$  the artifacts were quite noticeable with an average level above 2. This indicates that just using PN modulation would in fact be quite noticeable as compared to the chirps. Note, this histogram is on a scale of 1 to 5 while the remaining histograms are on a scale from 0-1.

The next sequence of tests compared linear and exponential fading lengths. Each fade-in and fade-out was added to front and back of a  $20ms$  chirp between  $19kHz$  and  $23kHz$ . The fade periods were on fixed tones and hence do not remove any amplitude from the main chirp. To our surprise it appears that exponential fade approaches tend to be significantly more noticeable than linear fading. Linear fading tends to decrease and then flatten-off at around  $5ms$ . To minimize transmission time, we chose a final fade value of  $5ms$ .

Using a long fading length of greater than  $20ms$ , we then test if the duration of the chirp has any impact on its perceptibility. All of the chirps swept between  $19kHz$  and  $23kHz$  with a rate configured by the desired test duration. As shown in Figure 8 (d), users could easily perceive very short chirps since they are quite similar to frequency jumps. Interestingly, as the length of the chirp increased, users began to notice a "swooshing" sound. For this reason, we sized our chirps to be at least  $20ms$  and no longer than  $200ms$ . In practice, the chirp should be sized to the excess delay of the channel which is usually around  $100ms$  at reasonable power output levels.

The final set of experiments evaluated the impact of applying rate adaptation to the chirps. In these tests, a worst-case chirp was generated for each chirp rate size where users were given the slowest rate followed by the fastest rate. As the second rate increases one would expect the user to perceive the rapid frequency changes. As can be seen in Figure 8 (e), there was a slight increase in perception due to increasing the number of possible rates (which leads to a higher rate



**Figure 9. Jitter Performance**

for one of the chirp sections).

#### 4.5 Coping with real transducers

Some mobile devices will have better frequency stability across the ultrasonic frequency band than others. Based on the frequency response of various mobile devices shown earlier in Figure 2, we see that some devices tend to drop-off above  $22\text{kHz}$ . Performing equalization based upon these measured curves can be used to improve performance. In the future, we could imagine a periodic calibration sequence where the mobile devices would record a known frequency sweep to not only calibrate their own microphones, but also accommodate for discrepancies in the speakers. In the experimental evaluation section, we include tests run using a measurement microphone with an extremely flat frequency response that is compared against the performance of various phones.

### 5 Experimental Evaluation

In this section we evaluate the performance impact of our modulation parameters on the bit error rate of data transmissions. We then show location accuracy results using our chirping scheme on a relatively simple TDOA system.

#### 5.1 Experimental Setup

Our experimental setup included a microphone stand, audio DAC/ADC and piezo electric tweeters. The microphone stand shown in Figure 8(b) consisted of an Audix TM1 omnidirectional measurement microphone and a smartphone holder. In Figure 8(b) we see the Audix mic on the left and an iPhone 4 in the holder on the right. The Audix microphone was chosen due to its extremely flat frequency response all the way up to  $25\text{kHz}$ . In order to generate high-quality test sounds we chose to use a Motu UltraLite-mk3 audio interface. The mk3 provides both a 24bit  $192\text{kHz}$  ADC and DAC with up to 10 channels of analog output. In Figure 8(a) we show the mk3 connected to two Onkyo HT-R540 amplifiers. Each Onkyo amplifier provides 7 channels of amplification, so we require two of them to utilize the entire 10 channels of output from the mk3. The HT-R540 has an extended frequency response mode that remains relatively flat

up to  $100\text{kHz}$ . Finally, we connect each output channel from the Onkyo to 10 Goldwood GT-1016 Dispersion Piezo Horn Tweeters. These are low-cost (\$2.45 each) tweeters that have a frequency response of up to  $27\text{kHz}$ . Figure 8(c) shows one of the tweeters mounted to a tripod stand.

For each test (unless specified otherwise), audio was transmitted from the speakers and recorded by both the Audix microphone using the mk3 ADC and by an iPhone 3GS. The iPhone used a wireless file-sharing program to push the recorded sound clips back to our main computer for processing. Test sequences could be remotely started and stopped using a VNC client on the iPhone. Streamlining this process enabled us to evaluate an extensive set of parameters. The following graphs were generated from over 25 hours of combined recording samples.

#### 5.2 Chirp Tuning

After performing a user study to understand the perceived impact of our chirp parameters, we then evaluate the performance impact on the system's ability to transmit data. First, we determine the impact of fading the signal in and out on the Bit Error Rate (BER) of data transmissions. In order to compute the BER, we transmit a modulated sequence of 1024 random bits using a  $20\text{ms}$  chirp (sweeping from  $19\text{kHz}$  to  $23\text{kHz}$ ) at different transmit powers while varying the length of the fade in and out durations. Each point in the plot represents 20 seconds of samples. The signal-to-noise ratio was computed based on the average intensity of the chirp signal as compared to the average intensity of the noise floor when there is no transmission. In both cases, the signal was high-pass filtered to remove audible noise. In the following tests, the microphone was mounted approximately  $2\text{m}$  from the speaker. As the transmit power is decreased, the SNR correspondingly decreases. At each bit interval, the receiver must decide if it correctly detects a 1 or 0 bit by correlating an up-chirp or down-chirp at the correct rate across the signal. A BER value of 0.5 corresponds to the expected value if the bits are decided by random chance (the signal is unreadable). In this experiment, all measurements were taken using the Audix microphone so as to determine the general trend.

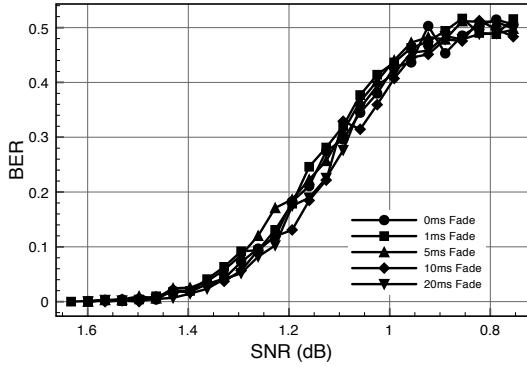


Figure 10. Impact of fade in/out on BER

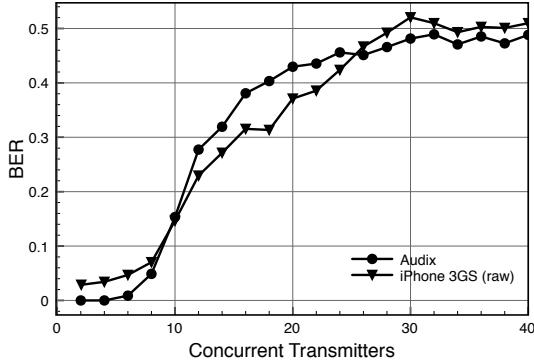


Figure 11. Impact of concurrent symbols on BER

Figure 10 shows that the fade lengths have almost no impact on the BER. This makes sense since the correlated input signal does not include the fade-in and fade-out regions.

Next, we evaluate how chirp length impacts BER. This test is similar to the one performed in the previous example, except now the fade period per symbol was set constant at 10ms (5ms fade in and out) while the chirp length was adjusted. We see that as the chirp length increases, the BER falls off at lower and lower SNR levels. This corresponds to the Pulse Compression equations that indicate that with longer chirps the signal should be distinguishable at lower SNR levels. Based on this performance graph as well as the user study, we select chirp lengths of 20ms for use in practice. In general, chirps should be greater than the excess delay of the channel in order to maximize performance under multi-path conditions. The excess delay can be determined by looking at the ultrasonic impulse response of a particular space. Sizing the chirps according to this parameter, however does lead to low data rates in most cases, requiring a compromise to be found.

We now evaluate how well the system scales with multiple concurrent transmitters by incrementing the number of rates used by the chirps. For each additional rate, we mix in a signal for all other rate values given a random offset around the signal that we are trying to decode. This corresponds to all other possible transmitters sending data simultaneously. For example, four concurrent transmitters would mean that the chirp can be modulated with four different rates and the

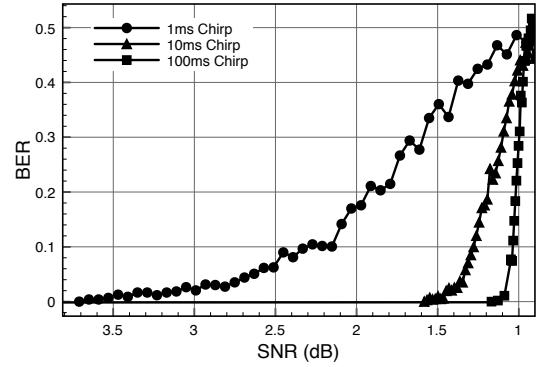


Figure 12. Impact of chirp length on BER

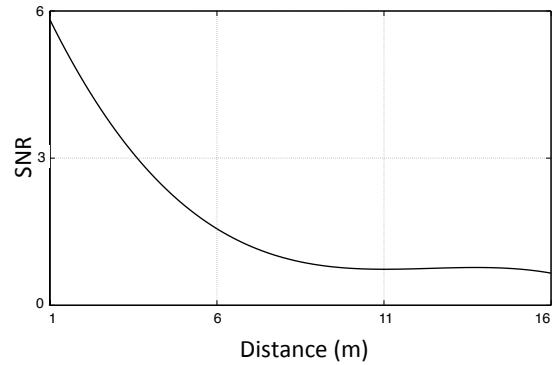
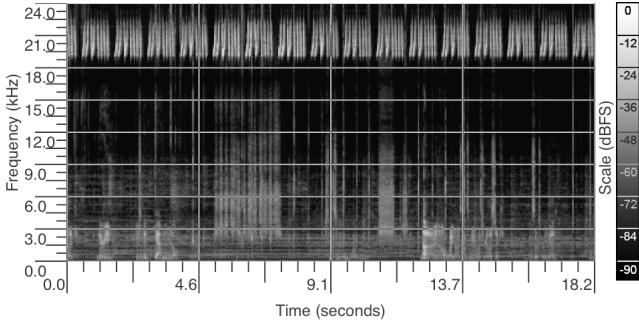


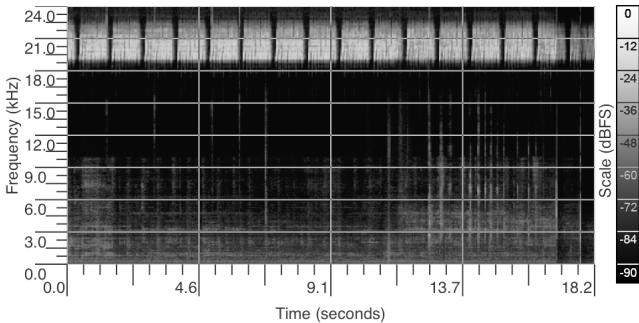
Figure 13. Distance vs SNR

three other possible rates are being mixed into the transmitted signal to act as simultaneous transmissions. Figure 11 shows the performance of the Audix reference microphone as well as an unequalized iPhone 3GS. First, we see that the Audix and iPhone perform comparably. We also see that the BER remains below 10% up to about 10 concurrent transmitters. In practice, most receivers will not overhear transmissions from all possible chirp rates within close proximity of each other, so this performance is quite pessimistic. Such a situation would only occur if many transmitters were placed in the same location. Above 16 rates, the system's BER begins to severely deteriorate. In order to support additional rates, we would require more bandwidth which is unfortunately limited by the microphone sampling rate. If in the future mobile devices could support higher sampling rates we would be able to support more concurrent transmitters and we would be able to achieve more precise ranging.

In order to estimate the transmission range of the system, we measure SNR versus distance as shown in Figure 13. For this test, the signal intensity at 1 meter away was measured at a modest 48dB(A), which corresponds to a volume level of about 5% of the maximum volume possible on the Onkyo amplifier and equivalent to a sound slightly louder than the humming of a refrigerator. At higher volume levels, we see transmissions as far as 50 – 100 meters. Depending on the deployment scenario, the transmit volume can be adjusted one way or another to aid in maximizing coverage and number of concurrent transmitters.



**Figure 14. Spectrogram with a single transmitter**



**Figure 15. Spectrogram with a four transmitters**

Figure 14 shows a spectrogram recorded on an iPhone of a 14 bit sequence from a single transmitter. Figure 15 shows the spectrogram when many transmitters are broadcasting simultaneously. The spaces between the sequences are still visible, but the other symbols are no longer distinguishable unless the signal is demodulated.

### 5.3 Timing Accuracy

We now evaluate the impact of various modulation parameters on timing jitter, which provides insight into ranging accuracy. Wide jitter distributions would result in poor distance estimates. During these experiments, we transmitted equally spaced chirps and then measured the distance between adjacent chirps. The jitter value is simply the difference between the detected chirp spacing and the transmitted chirp spacing. In Figure 9 we show plots for both the Audix and iPhone. The first histogram shows the performance of a fixed tone. We clearly see that without Pulse Compression the ranging resolution is quite poor (on the order of 2 – 4ms, which corresponds to 6 – 14 meters). Next, we use a 20ms tone and adjust the chirp bandwidth. We clearly see that additional bandwidth reduces the jitter. In the right-most histogram we then increase the length of the chirp from 20ms to 100ms. By increasing the chirp length we also see the jitter slightly reduce. These graphs verify the properties of Pulse Compression on chirps and show that our fade-in and rate adaptation is not introducing a significant degradation in quality.

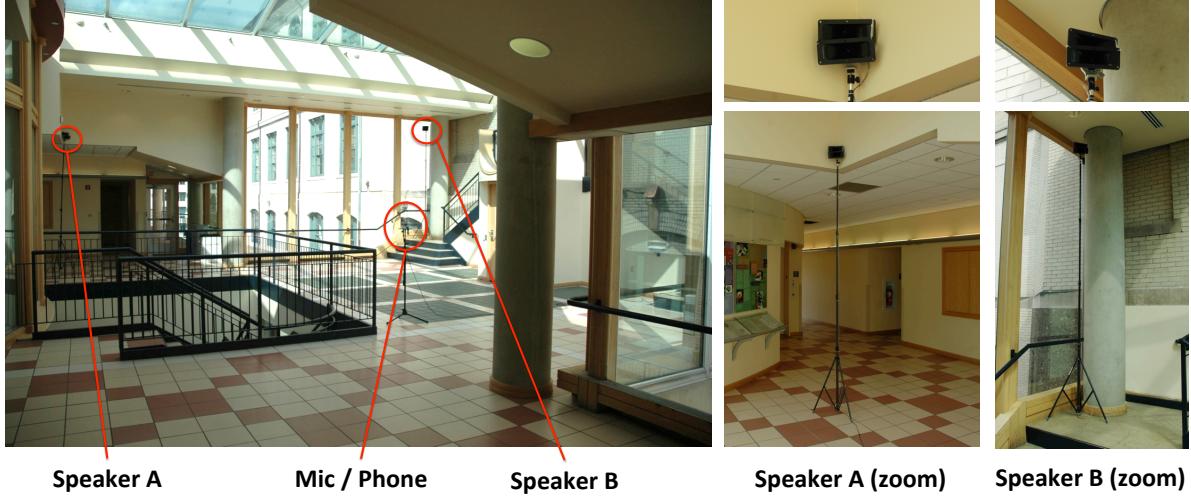
### 5.4 Location Accuracy

To demonstrate our ranging system’s ability to estimate locations, we perform a set of experiments that use TDOA from transmitters at known locations in the environment to compute the receiver’s position. We placed four speakers in

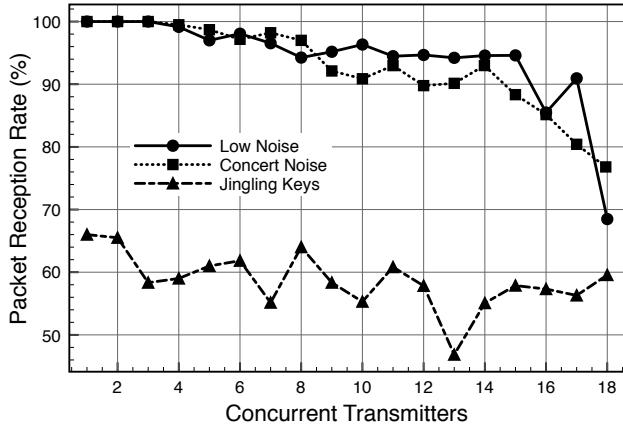
the corners of two spaces on campus. The first location is the 20m x 20m anonymous atrium shown in Figure 16. We chose the atrium location since it was similar to that of a museum environment with hard walls and tile floors. We then also chose a small 5m x 5m room with cement walls that exhibits a large amount of multi-path fading due to echoing. In order to capture the multi-path characteristics of each of these spaces, we record the impulse response at the center of the room. Figure 19 shows the excess delay in the atrium to be about 90ms, while Figure 21 shows an excess channel delay of about 60ms. One can see the multi-path is greater in the small room based on the number of echoes. One can also infer the size of the room by noting the spacing between echoes. We precisely measured the x,y, and z locations of the transmitter speakers. Using our microphone stand, we moved an iPhone 3GS and the Audix microphone to 25 different locations along a grid in the room. At each location we take five audio recordings to compute five position samples. We can then compute the ranging accuracy of the system by comparing our measured location as ground truth to each computed location. Figure 20 shows the distribution of errors across all of the samples. We see that 95% of the samples are within 100cm of the actual location with a worst-case overall error of 4m. As shown in Figure 22, the small room behaves similarly except with a sharper position accuracy fall-off due to the added channel fading. Figure 23 shows an example of the correlated signals at a particular location in the room. Each received signal is correlated with all of the different possible chirp rates. Each unique chirp rate that has a significant correlation is shown in a different color shade. The peak correlation points in the waveform as detected by our peak detector are denoted with a triangle. In some cases, the system was only able to detect three out of the four transmitter beacons if for example a barrier blocked one of the speakers. In these cases, our positioning algorithm will estimate the 2-dimensional location using the 3 detectable speakers. The step-like shape of the CDFs is attributed to having successfully detected all four transmitters in the vast majority of cases, but occasionally choosing the incorrect peaks for one or more of them when they are blocked or out of range. This results in the vast majority of errors being below 10cm, with only a handful of much larger ones. In practice, adding additional transmitters will help alleviate blind spots. This ability to scale the number of transmitters in an environment is one of the main motivations for our multiple-access symbols. More sophisticated positioning approaches like averaging samples over time and/or using Kalman filtering can be used to refine position values. However, even with simple TDOA ranging, we see the potential for significantly improved indoor positioning.

### 5.5 Data Reception

Up until now we have evaluated the performance characteristics of our system on the level of individual symbols, we will now do the same for the transmission of strings of data as described in Section 4.3. We have measured the packet reception rate (PRR) of a receiver using the Audix microphone, with respect to an increasing number of transmitters broadcasting their unique data sequences simultaneously. In this context the PRR is defined as the average number of trans-



**Figure 16. Photo of atrium environment (two speakers outside field of view)**



**Figure 17. PRR vs concurrent transmitters**

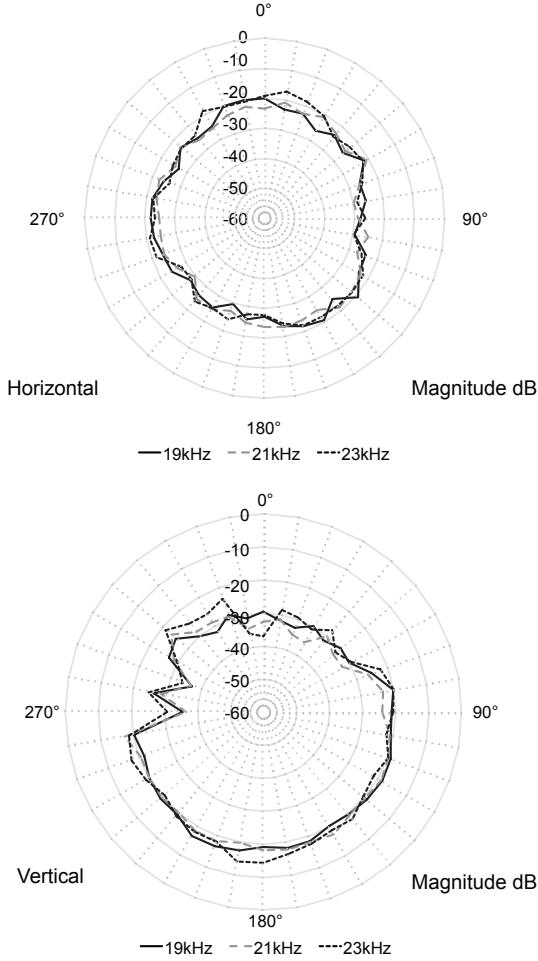
mitter IDs that were sent and successfully decoded, divided by the total number of IDs sent over 500 trials per increment in number of transmitters. To eliminate the need to set up an individual loudspeaker for each transmitter, we mixed several randomly time shifted data signals of random amplitudes, and broadcast them over a single speaker. The shifting of the signals in time was modeled after the shifts that can be expected to be received at various points in space inside a  $20m \times 20m \times 10m$  room with transmitters positioned along the top edges. It was assumed that the receiver was in range of 32 transmitters for each test, which could be used to filter some erroneously decoded IDs that were not caught by the forward error correction. Figure 17 shows that the PRR decreases by less than 15% for up to 17 concurrent transmissions. In comparison to the results shown in Figure 11, the increase of concurrent transmitters resulted in a relatively low increase of errors and dropped data. This highlights the improvements gained by employing the various error detection, correction and filtering techniques implemented in our receiver at the data layer. It also verifies that our data de-

modulation and decoding techniques are capable of detecting many more than the minimum four transmitters required to perform 3D localization by TDOA ranging, even if their signals overlap. This test was repeated with added noise in the form of music (to simulate a concert environment) and jingling keys to introduce wideband ultrasonic noise. The noise was played from an additional speaker pointed directly at the receiving microphone. The maximum amplitude of the music received at the microphone was approximately 32x that of the individual chirp sequences, while the key jingling sounds were approximately 4x louder. Since the music did not contain significant energy above  $19kHz$ , it had little impact on the PRR. The key jingling which contains a large component of ultrasound did significantly impact the PRR, but more than half of the packets were still able to be decoded.

Finally, it is worth noting the impact of synchronization error (between transmitters) on distance estimates. Existing sensor networking protocols have shown synchronization accuracy of well below  $10\mu s$  [25, 26, 27]. This corresponds to less than  $3cm$  of positioning error given typical acoustic signal propagation speeds.

## 6 Limitations

While promising, this approach has some limitations. First, all of the transmitters need to be precisely synchronized and individually addressable. In theater surround-sound systems this is usually the case, but in large public arenas like airports or shopping malls, PA systems have significantly fewer channels than speakers. This may change with innovations like LED lighting speakers [28]. Second, this approach requires that the speakers and mobile phones support frequencies up to  $24kHz$ . Most of the Android and iPhone models that we tested supported up to at least  $22kHz$  (the iPhone 4 remains relatively flat to  $24kHz$ ). We found that most speakers with tweeters tend to operate relatively close to  $24kHz$ , however some amplifiers that are not intended for use with movie sound systems can have  $20kHz$  low-pass filters. A related limitation is that systems operat-

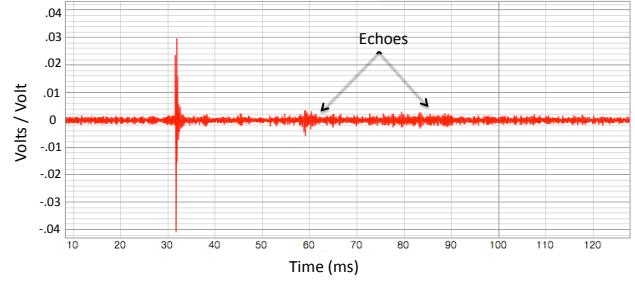


**Figure 18. iPhone 3GS microphone polar pattern**

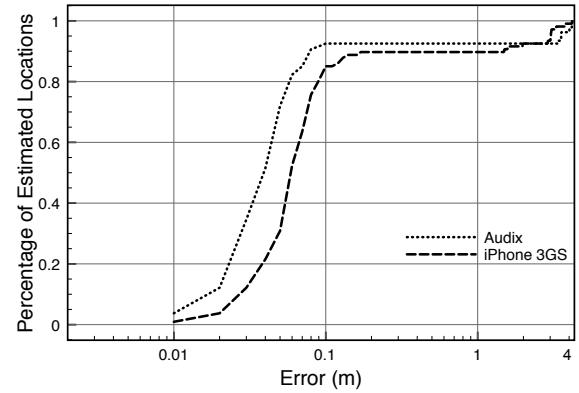
ing at just above the human hearing range could be problematic for pets that are able to hear these higher frequencies (see 6.3 for more details). Third, the system still suffers in extreme multi-path environments and interferences from clanging metal. Typically these sounds are intermittent as long as the system isn't operating in an industrial environment with loud machinery. Figure 24 shows a spectrogram captured at a grocery store during a busy weekend with many shopping carts moving around. We see that the high frequency ranges are relatively clear. Finally, in order to compute locations, the ranging system requires precise transmitter location information, which can be difficult to deploy in practice. We are currently investigating automated training procedures to help simplify that process.

## 6.1 Receiver Orientation

In order to ensure reliable signal reception, it is important that the receiver's microphone has an omnidirectional polar pattern in the 19 – 23kHz range. Figure 18 shows the polar pattern of the iPhone 3GS, where the front of the microphone is pointing towards the 180° marker in the vertical pattern, and the screen of the phone is pointing towards the 180° marker in the horizontal pattern. While the horizon-



**Figure 19. Impulse response in atrium environment**

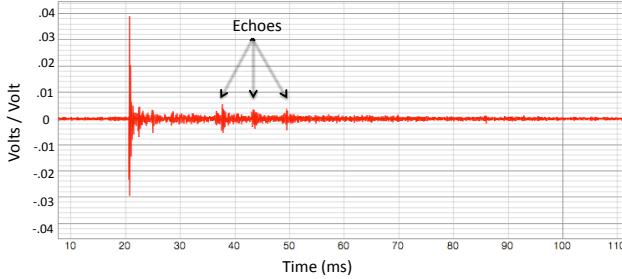


**Figure 20. Positioning error in atrium environment**

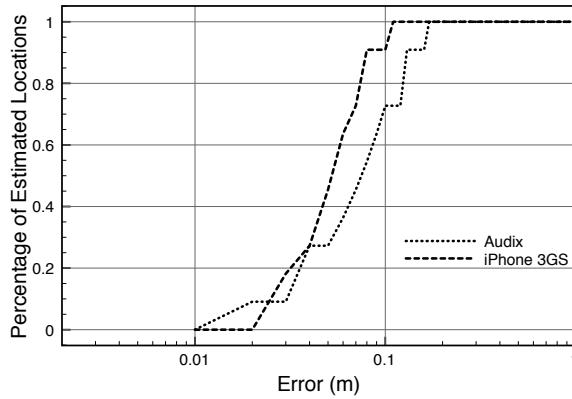
tal pattern is reasonably omnidirectional at all frequencies in the 19 – 23kHz range, the vertical pattern shows significant attenuation from 70° to 280° due to the back side of the microphone being blocked by the phone's chassis. We can expect several transmitters to be transmitting at this region, therefore additional processing like signal equalization may be beneficial to decoding the data in these signals. Newer phones such as the iPhone 4 and 4S contain an additional microphone on their top side, which could potentially be used in combination with the main microphone to achieve greater omnidirectionality.

## 6.2 Human Health Concerns

Extensive studies have been conducted to quantify safe volumes and exposure limits of ultrasound on humans. The Health Protection Branch of Health Canada published a report that summarizes multiple studies related to ultrasound [29]. This report suggests that for frequencies above 20kHz, the level should be kept below 110dB to prevent undesirable subjective effects of ultrasound. These effects include fullness in the ear, fatigue, headache and malaise. For reference, in the audio range, 110dB is approximately the loudness of a power saw from three feet away. Hearing damage can occur at above 95dB. The study also indicates that subjects are more susceptible to fixed frequency tonal sounds. For this reason, we believe that chirp pulses can be considered safe under prolonged exposure when kept at a reasonable volume. In all of our tests, the volume remained below 75dB. Even so, we believe that in certain environments, the transmission of data could be activated on demand by the mobile devices



**Figure 21. Impulse response in small room environment**

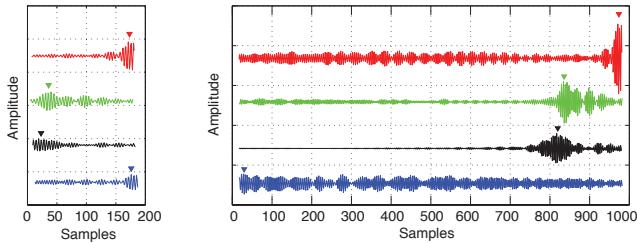


**Figure 22. Positioning error in small room environment**

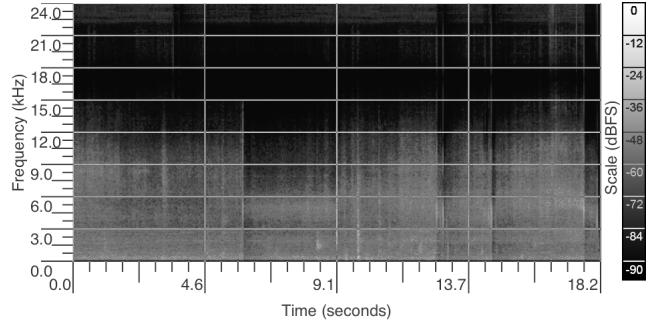
or aggressively duty-cycled to further mitigate potential annoyance to users (especially children which were not tested due to our IRB limitations).

### 6.3 Animal Exposure

Animals are known to have significantly greater hearing range than humans [30]. At the extreme, mice, bats, whales and porpoise can hear frequencies as high as  $90 - 150\text{kHz}$ . However, real-world deployments would be more concerned by the hearing range of household pets and service animals. Dogs and cats can hear frequencies as high as  $45\text{kHz}$  and  $64\text{kHz}$  respectively. It is difficult to ascertain the full extent of hearing attenuation at higher frequencies or if the sounds have a negative effect on the animals. While not extensively tested, we did play sample tones in a home with two cats. Initially, it was unclear if the cats could hear the tones since they exhibited no noticeable response. We then played a sample tone before feeding each cat for a few consecutive days. It then became apparent that the cats could in fact hear the



**Figure 23. Correlation of four different symbols**



**Figure 24. Spectrogram captured in busy grocery store**

tone based on their reaction once a food association was established. Significant further testing would be required to draw any real conclusions, but it appears that animals do hear the sound, but that limited exposure does not cause an (immediate) adverse reaction.

## 7 Conclusions and Future Work

High precision indoor ranging for mobile devices has the potential to enable a new set of pervasive computing applications. Current approaches either suffer from poor accuracy, reliability, or require additional hardware on the mobile device. In this paper, we presented an indoor ranging system for mobile devices that does not require any additional hardware on the receiver side and can operate on most smartphone devices currently on the market. The approach uses carefully modulated ultrasonic chirps just outside of the human hearing range to transmit small amounts of data and ranging information. This poses challenges associated with accurately determining the Time-Difference-of-Arrival of signals, making sure they are inaudible to humans, and facilitating an efficient multiple-access scheme so that the system can easily scale. Accurate timing data is achieved by use of Pulse Compression on chirps increasing linearly in frequency. A multi-rate modulation scheme within each chirp allows for the transmission of multiple unique signals. Data sequences are then modulated using up-chirp and down-chirp sequences at a transmitter defined rate. Since audio speakers have a non-ideal impulse response, we use a fading technique for suppressing clicking noises that would normally occur due to large signal changes. We show the impact of various modulation parameters on the audibility of the signal through a user study. We then evaluate the functional performance of each parameter through experimentation on a real hardware. Using a TDOA pseudo-ranging approach we are able to localize 95% of our test points to below 10cm accuracy in a large atrium space.

The focus of this paper is on the modulation technique and the system's ability to extract ranging data. For future work, we plan to improve upon the robustness and accuracy of the localization scheme by using more advanced techniques on top of our underlying ranging system. For example, Kalman filtering and integration with already existing localization services could be used to reduce error. This work will also include an in-depth evaluation of how both transmitter density and speaker geometry impact localization

accuracy. The next major challenge in these types of ranging systems is to streamline the deployment process. We believe that with a small set of known receiver coordinates, it should be possible to automatically determine the positions of the speakers. This becomes an even more challenging problem when considering large indoor areas with many speakers distributed across the space. We are also interested in applying this scheme to signature-based localization. Instead of determining a specific TDOA, the output of our correlated signals could be mapped to particular locations using learning techniques. This would be an easy approach for supporting location tracking in existing spaces with PA systems that do not provide individually addressable speaker channels. We eventually envision a hybrid approach that uses TDOA when speaker location data is available, otherwise it uses signature approaches.

## 8 Acknowledgments

This research was funded in part by the Bosch Research and Technology Center in Pittsburgh and the Intel Science and Technology Center on Embedded Computing. We would like to thank our reviewers and our shepherd Cecilia Mascolo for all of their great input.

## 9 References

- [1] Plack,C.J. *The sense of hearing*. Lawrence Erlbaum Associates, Inc., 2005.
- [2] G. Neuweiler. Evolutionary aspects of bat echolocation. *J Comp Physiol*, 2003.
- [3] Nissanka B. Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (Mobicom '00)*, pages 32–43, New York, NY, USA, 2000. ACM.
- [4] B.W. Parkinson and S.W. Gilbert. Navstar: Global positioning system - ten years later. *Proceedings of the IEEE*, 71(10):1177 – 1186, oct. 1983.
- [5] Gaetano Borriello, Alan Liu, Tony Offer, Christopher Palistrant, and Richard Sharp. Walrus: wireless acoustic location with room-level resolution using ultrasound. In *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services (MobiSys '05)*, pages 191–203, New York, NY, USA, 2005. ACM.
- [6] P. Bahl and V.N. Padmanabhan. Radar: an in-building rf-based user location and tracking system. In *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '00)*, volume 2, pages 775 –784 vol.2, 2000.
- [7] A. Ward, A. Jones, and A. Hopper. A new location technique for the active office. *IEEE Personal Communications*, 4(5):42 –47, oct 1997.
- [8] Konrad Lorincz and Matt Welsh. Motetrack: a robust, decentralized approach to rf-based location tracking. In *Proceedings of the 1st International Conference on Location- and Context-Awareness (LoCA'05)*, pages 63–82, Berlin, Heidelberg, 2005. Springer-Verlag.
- [9] U. Maurer, A. Rowe, A. Smailagic, and D.P. Siewiorek. ewatch: A wearable sensor and notification platform. In *International Workshop on Wearable and Implantable Body Sensor Networks (BSN '06)*, pages 4 pp. –145, april 2006.
- [10] Stephen P. Tarzia, Peter A. Dinda, Robert P. Dick, and Gokhan Memik. Indoor localization without infrastructure using the acoustic background spectrum. In *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services (MobiSys '11)*, pages 155–168, New York, NY, USA, 2011. ACM.
- [11] A. Rowe, Z. Starr, and R. Rajkumar. Using micro-climate sensing to enhance rf localization in assisted living environments. In *IEEE International Conference on Systems, Man and Cybernetics (ISIC '07)*, pages 3668 –3675, oct. 2007.
- [12] Isaac Amundson and Xenofon D. Koutsoukos. A survey on localization for mobile wireless sensor networks. In *Proceedings of the 2nd International Conference on Mobile Entity Localization and Tracking in GPS-less Environments (MELT '09)*, pages 235–254, Berlin, Heidelberg, 2009. Springer-Verlag.
- [13] Kaveh Pahlavan, Xinrong Li, Mika Ylianttila, Ranvir Chana, and Matti Latva-aho. An overview of wireless indoor geolocation techniques and systems. In *Proceedings of the IFIP-TC6/European Commission International Workshop on Mobile and Wireless Communication Networks (NETWORKING '00)*, pages 1–13, London, UK, UK, 2000. Springer-Verlag.
- [14] Zheng Sun, R. Farley, T. Kaleas, J. Ellis, and K. Chikkappa. Cortina: Collaborative context-aware indoor positioning employing rss and rtof techniques. In *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM '11 Workshops)*, pages 340 –343, march 2011.
- [15] Chunyi Peng, Guobin Shen, Zheng Han, Yongguang Zhang, Yanlin Li, and Kun Tan. A beepbeep ranging system on mobile phones. In *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems (SenSys '07)*, pages 397–398, New York, NY, USA, 2007. ACM.
- [16] Mike Hazas and Andy Ward. A novel broadband ultrasonic location system. In *Proceedings of the 4th International Conference on Ubiquitous Computing (UbiComp '02)*, pages 264–280, London, UK, UK, 2002. Springer-Verlag.
- [17] Mike Hazas and Andy Ward. A high performance privacy-oriented location system. In *Proceedings of the 1st IEEE International Conference on Pervasive Computing and Communications (PERCOM '03)*, pages 216–223, Washington, DC, USA, 2003. IEEE Computer Society.
- [18] Michael McCarthy, Paul Duff, Henk L. Muller, and Cliff Randell. Accessible ultrasonic positioning. *IEEE Pervasive Computing*, 5(4):86–93, October 2006.
- [19] Jie Yang, Simon Sidhom, Gayathri Chandrasekaran, Tam Vu, Hongbo Liu, Nicolae Cescan, Yingying Chen, Marco Gruteser, and Richard P. Martin. Detecting driver phone use leveraging car speakers. In *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking (MobiCom '11)*, pages 97–108, New York, NY, USA, 2011. ACM.
- [20] V. Filonenko, C. Cullen, and J. Carswell. Investigating ultrasonic positioning on mobile phones. In *International Conference on Indoor Positioning and Indoor Navigation (IPIN '10)*, pages 1 –8, sept. 2010.
- [21] R. Bucher and D. Misra. A synthesizable vhdl model of the exact solution for three-dimensional hyperbolic positioning system. *VLSI Design*, 15(2):507–520, 2002.
- [22] H. Shen, S. Machineni, C. Gupta, and A. Papandreou-Suppappola. Time-varying multichirp rate modulation for multiple access systems. *IEEE Signal Processing Letters*, 11(5):497 – 500, may 2004.
- [23] Sergio Verdu. *Multicuser Detection*. Cambridge University Press, New York, NY, USA, 1st edition, 1998.
- [24] R. Price and P.E. Green. A communication technique for multipath channels. *Proceedings of the IRE*, 46(3):555 –570, march 1958.
- [25] Miklós Maróti, Branislav Kusy, Gyula Simon, and Ákos Lédeczi. The flooding time synchronization protocol. In *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems, SenSys '04*, pages 39–49, New York, NY, USA, 2004. ACM.
- [26] Jeremy Elson, Lewis Girod, and Deborah Estrin. Fine-grained network time synchronization using reference broadcasts. In *Proceedings of the 5th Symposium on Operating Systems Design and Implementation (OSDI '02)*, pages 147–163, New York, NY, USA, 2002. ACM.
- [27] Saurabh Ganeriwal, Ram Kumar, and Mani B. Srivastava. Timing-sync protocol for sensor networks. In *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys '03)*, pages 138–149, New York, NY, USA, 2003. ACM.
- [28] Klipsch. Lightspeaker in-ceiling lighting and audio system. <http://www.klipsch.com/lightspeaker-in-ceiling-lighting-and-audio-system>, 2012.
- [29] Health Protection Branch Health Canada. Environmental health directorate. guidelines for the safe use of ultrasound: Part ii - industrial and commercial applications - safety code 24. *Published by authority of the Minister of National Health and Welfare*, 1991.
- [30] Fay RR. and Popper AN. Comparative hearing: Mammals. *Springer Handbook of Auditory Research Series*, 4, 1994.