

# Acoustical Ranging Techniques in Embedded Wireless Sensor Networked Devices

PRASANT MISRA, University of New South Wales and Swedish Institute of Computer Science  
NAVINDA KOTTEGE, Commonwealth Scientific and Industrial Research Organization  
BRANISLAV KUSY, Commonwealth Scientific and Industrial Research Organization  
DIETHELM OSTRY, Commonwealth Scientific and Industrial Research Organization  
SANJAY JHA, University of New South Wales

Location sensing provides endless opportunities for a wide range of applications in GPS-obstructed environments; where, typically, there is a need for higher degree of accuracy. In this article, we focus on robust *range estimation*, an important prerequisite for fine-grained localization. Motivated by the promise of acoustic in delivering high ranging accuracy, we present the design, implementation and evaluation of acoustic (both ultrasound and audible) ranging systems. We distill the limitations of acoustic ranging; and present efficient signal designs and detection algorithms to overcome the challenges of coverage, range, accuracy/resolution, tolerance to Doppler's effect, and audible intensity. We evaluate our proposed techniques experimentally on TWEET, a low-power platform purpose-built for acoustic ranging applications. Our experiments demonstrate an operational range of 20 m (outdoor) and an average accuracy  $\approx 2$  cm in the ultrasound domain. Finally, we present the design of an audible-range acoustic tracking service that encompasses the benefits of a near-inaudible acoustic broadband chirp and approximately two times increase in Doppler tolerance to achieve better performance.

Categories and Subject Descriptors: C.3 [**Special-Purpose and Application-Based Systems**]: Signal processing systems

General Terms: Design, Algorithms, Experimentation

Additional Key Words and Phrases: Ranging, Localization, Tracking, Ultrasound, Audible-range Acoustics, Linear Chirp, Envelope Detection, Near-inaudible Acoustic Signal Design, Doppler's effect

## ACM Reference Format:

Prasant Misra, Navinda Kottege, Branislav Kusy, Diethelm Ostry, and Sanjay Jha, 0000. Acoustical Ranging Techniques in Embedded Wireless Sensor Networked Devices *ACM Trans. Sensor Netw.* 0, 0, Article 00 (0000), 40 pages.

DOI: <http://dx.doi.org/10.1145/0000000.0000000>

## 1. INTRODUCTION

Determining the location of a device is a fundamental problem, and its importance has motivated a large body of research on localization for indoor and outdoor environments where the Global Positioning System (GPS) does not work well. While the applications of outdoor location information are widespread, our work is motivated towards indoor applications in the field of binaural science, acoustic source detection, location-aware sensor networking, target motion analysis, or mobile robot navigation. Location-

---

Author's addresses: P. Misra, Computer Systems Laboratory, Swedish Institute of Computer Science, Swedish ICT, Stockholm; N. Kottege and B. Kusy, Autonomous Systems Laboratory, CSIRO ICT Centre, Brisbane; D. Ostry, Wireless and Networking Technologies Laboratory, CSIRO ICT Centre, Sydney; S. Jha, School of Computer Science and Engineering, University of New South Wales, Sydney.

Permission to make digital or hard copies of part or all 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 show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or [permissions@acm.org](mailto:permissions@acm.org).

© 0000 ACM 1550-4859/0000/-ART00 \$15.00

DOI: <http://dx.doi.org/10.1145/0000000.0000000>

aware applications deployed in indoor and GPS-obstructed environments, such as a roof-covered canteen that is not strictly confined by physical boundaries from all sides, require a higher degree of accuracy than typical outdoor applications. However, accuracy of localization techniques in these environments remains a challenge.

The end-to-end process of location sensing consists of two sequential phases: (i) measurement, and (ii) positioning. For an active-cooperative location system [Savvides et al. 2004] wherein the target  $S$  probes the components of the system infrastructure  $R$  with a physical signal, the measurement phase consists of processing the received signal to estimate parameters of interest such as distance, angle or phase of arrival. The measurements are subsequently utilized in the positioning phase to compute the location coordinates. In this case, the measurement phase is invariably referred to as the *ranging* phase. Range estimation is a *crucial prerequisite* for reliable and high accuracy location information as a minor measurement bias will result in positioning errors that scale with increasing distance. The ranging performance depends on: (i) deployment and configuration of the location systems, and (ii) quality of the ranging waveform and measurement technique [Win et al. 2007]. Hence, an important research focus on fine-grained localization has been on robust distance estimation.

For high-accuracy range information, the most successful techniques are based on measuring the time-of-flight (TOF) of signals [Savvides et al. 2001]. Other common techniques, such as signal strength measurement or fingerprinting, tend to be highly susceptible to environmental interference [Zhou et al. 2004]; and so, are unreliable and are less preferred as standalone methods. In TOF approaches, there are two competing technologies: radio frequency (RF) and acoustics. Acoustic signals have been identified with a number of important features that provide significant advantages over RF for delivering high ranging accuracy.

**Ranging: Using Acoustics.** There are a number of factors that make acoustic attractive. Acoustic signals have low frequency components that are normally in the order of Hz/kHz rather than MHz/GHz typical for RF. Therefore, acoustic processing requires significantly lower sampling rates. Sampling rates between 40 kHz to 100 kHz are sufficient to adequately recover both audible and ultrasonic acoustic signals. Hence, the currently available commercial off-the-shelf (COTS) acoustic components are relatively inexpensive and simple to interface.

Compared to RF, acoustic waveforms have significantly slower propagation speed. With respect to ranging accuracy, this feature offers a key advantage that eases the synchronization requirements among the different components of the location system. This factor, not only provides better compensation of the timing errors, but also, creates a scope for the use of cheap and low-frequency clocks on acoustics devices that intrinsically consume less power than typical RF devices. The latter is an important metric for system design considerations, which lowers the overall cost of acoustic receivers and shifts the high power requirement to the acoustic transmitter device for long range detectability [Girod and Estrin 2001; Girod 2000].

Although, the aforementioned benefits make acoustics an attractive choice for location systems, it has several shortcomings. Its performance is limited by physical factors, such as reflections from the environment and ground, variation in air density caused by thermal effects leading to variation in sound speed, and propagation effects caused by non-uniformities in the atmosphere (due to wind/turbulence). Acoustic signals, below 20 kHz, feature psychoacoustics clues perceivable by humans. Therefore, acoustic systems that operate in the human audible range have not received wide acceptability as part of a general localization strategy or for covert operations [Kushwaha et al. 2005; Girod et al. 2006; Peng et al. 2007; Zhang et al. 2007; Kwon et al. 2005; Borriello et al. 2005].

In this regard, ultrasonic systems operating above the human perception range offer the convenience of inconspicuous location sensing, but introduce other complications. Ultrasound is more sensitive to atmospheric absorption and fading effects than frequencies in the audible spectrum, and consequently, may face the problem of reduced coverage range [Ash and Moses 2005]. Existing COTS ultrasound sensors exhibit a limited beam angle, which impacts their associated ranging performance that degrades with an increased angle offset between the transmitter and the receiver [Priyantha 2005]. In addition, higher sampling rates are required to process ultrasound, which inherently increases the cost, size/weight and power consumption of the sensing platform. Considering these limitations, previous ultrasonic systems were designed for applications restricted to only indoor environments that support dense deployment and require short coverage range [Harter et al. 1999; Priyantha 2005; Savvides et al. 2001].

**Ranging: Impact of Signal Design.** Apart from the physical waveform properties and hardware platform features, the design of the ranging signal also plays a key role in delivering the desired accuracy/resolution and coverage range. *Narrowband* signals have a relative bandwidth  $B_r$  (i.e., ratio of the bandwidth to the center frequency) of less than 0.2. Due to their limited frequency span, they are highly sensitive to environmental noise and also face difficulties in resolving multipath reflections [Weiss and Weinstein 1983].

To overcome these limitations, a broad range of frequencies (with  $B_r > 0.2$ ) can be used to essentially reduce the chances of the entire signal fading at any particular time [Klauder et al. 1960; Weinstein and Weiss 1984]. Klauder et al. [Klauder et al. 1960] described *pulse compression*, a signal processing technique that can both resolve multiple propagation paths as well as increase signal-to-noise ratio (SNR) of the direct path (that gives the range) without increasing the transmission power. As a result, a *broadband* signal can be processed to form a strong pulse at the line-of-sight (LOS) path without increasing the noise to the same extent. However, depending on the propagation channel conditions, this mechanism introduces an uncertainty in estimating the correct correlation peak due to the presence of numerous sidelobes (i.e., adjacent peaks surrounding the correlation peak) that may attain similar heights, and hence, contribute to ranging inaccuracy.

There are numerous broadband waveforms suitable for use with pulse compression such as chirps (linear/nonlinear), pseudo-noise (PN), or maximum length sequence (MLS) [Kottege and Zimmer 2011]. With regards to spectral simplicity and high processing gain, *linear chirps* offer the best tradeoff. Linear chirps exhibit reliable detection and ranging performance for stationary targets. However, their efficacy tends to degrade for moving targets maneuvering with high speed as they no longer achieve pulse compressibility with the introduction of Doppler's effect [Kelly and Wishner 1965; Yang and Sarkar 2006].

**Contributions.** We provide a comprehensive discussion of signal design and detection methodology that address a number of shortcomings of existing techniques, and make four research contributions.

*First*, we study the unidirectionality problem of narrowband ultrasound sensors, and its impact on coverage/range by undertaking a case-study of the Cricket indoor location system that pioneered the field of ultrasound localization. We draw insights from the various design aspects with respect to its ranging capabilities, and present an improved version of the existing Cricket system with detailed description of its design and implementation. The new hardware unit comprises of an omnidirectional ultrasonic receiver, an array of three ultrasound transducers, integrable with the existing Cricket motes, which becomes operational when configured as a listener. Empirical studies show a modest improvement, although the coverage range achieved

by both the original and modified versions of Cricket is quite limited.

In order to overcome the limited multipath resolving capability of narrowband ranging signals, we investigate the efficacy of broadband ultrasonic chirps, and study the tradeoff between range and accuracy/resolution that is dependent on the transmit pulse length and its bandwidth. We alleviate the peak ambiguity problem by proposing a signal detection algorithm for estimating the envelope of the correlated (compressed) pulse by a least-squares approximation technique. These various improvements in signal design and detection are, finally, combined into the design and implementation of TWEET, a mote-based ultrasonic broadband ranging system. The TWEET system consists of two separate hardware units: beacons and listeners, each consisting of a low-power sensor node, an audio codec, and a Blackfin DSP. Experimental results indicate that the system has an operational range of 20 m with an average accuracy of  $< 2$  cm with a 95% confidence interval of 2 cm. Using the TWEET platform, we demonstrate that broadband ultrasound is also a good candidature for long distance ranging (both indoors and outdoors), which is our *second* contribution.

Acoustic location systems operating in the human-audible range are often deemed unsuitable for general ranging applications, especially in indoor environments. This can be solved by shifting the signal frequencies to a range that is inaudible to humans. As our *third* contribution, we present the signal design features of a near-inaudible acoustic broadband chirp that combines the principles of human psychoacoustics with signal engineering techniques. Tests with human subjects suggest that the near-inaudible signal design is most effective when humans are not informed about the ambient chirping sound and they are involved in other simultaneous activities.

Tracking of mobile targets requires the acoustic location systems to withstand Doppler shifts in the signals, introduced by the relative velocity of the tracked and infrastructure nodes. Although, nonlinear chirps are more Doppler tolerant, we show that similar capability can be developed in linear chirps by using its sweep characteristic. Therefore, *finally*, we propose a detection algorithm that capitalizes on the linearly sweeping property of the linear chirp to measure the Doppler shifts caused by the moving target, and simultaneously estimates its relative speed and range. For tracking support, we present the design and implementation of TWEET-v2, an enhanced version of the TWEET system. Experimental results indicate approximately two times increase in Doppler tolerance levels of the (acoustic) linear chirp.

This article synthesizes and extends our prior work in this area, distilling our 3.5 year study down to a set of the most important findings and design challenges. Our earlier papers on this work examined the directionality problem of ultrasound sensors [Misra et al. 2011a] (Section 3.2.1) and introduced the TWEET ranging system [Misra et al. 2011c] (Section 4). This article collects all of these results and also adds new results and observations; the lessons and experiences of which will be helpful to other engineers working on similar projects.

The rest of the article is organized as follows: In the next section, we outline the related work followed by discussion and empirical study of the limitations of existing ranging techniques in Section 3. Section 4 presents signal design features, detection algorithm and implementation of TWEET as part of a broadband ultrasonic ranging system. Section 5 introduces the design of a near-inaudible acoustic signal and a Doppler tolerant detection algorithm as part of a tracking service with TWEET-v2. The final section suggests potential research directions and concludes with a summary of the areas covered in the article.

## 2. RELATED WORK

Determining the location of a target is based on two basic approaches: *landmarking* and *dead-reckoning*. The landmark based method requires selecting a set of three or more reference points (fixed or mobile) with known coordinates, obtaining their separation distance from the target, and finally, triangulating or multilaterating to obtain a position estimate within the selected coordinate system. On the other hand, dead-reckoning uses the motion dynamics of the target to determine its position with respect to some starting point. However, it suffers from drift since errors in measured dynamics (e.g., velocity, acceleration, odometry) accumulate when integrated over time. Therefore, most location systems are implemented using landmarks or a combination of landmarks and dead-reckoning.

Landmark-based systems determine the target's position based on its proximity to the reference points, which is derived by distance/angle measurements or signal strength signatures. The time-based distance measurement techniques such as time-of-arrival (TOA), time-difference-of-arrival (TDOA), round-trip time (RTT), and elapsed time between two time-of-arrivals (ETOA) are widely used approaches for location estimation. Since our primary focus is on acoustics, we only summarize the available systems within this scope, and refer our readers to the articles by Misra [Misra 2012] and Hui et al. [Hui et al. 2007] for a general review of location systems.

**Acoustic narrowband systems.** The Active Bat [Harter et al. 1999], Cricket [Priyantha 2005], AHLoS [Savvides et al. 2001], WALRUS [Borriello et al. 2005], Thunder [Zhang et al. 2007] and Kwon et. al. system in [Kwon et al. 2005] are existing narrowband systems. They share a common ranging technique, wherein the beacon transmits synchronous RF and sound (acoustic/ultrasound) pulses. The listener receives the fast propagating RF pulse (almost instantaneously) followed by the sound pulse; and computes the separation distance by measuring the time-lag between the arrival of these signals. However, these systems differ in their respective architecture, implementation technique, and hardware platform.

The Active Bat system uses a centralized controller to coordinate the ranging operation between the transmitter (called the Bat) and the receiver units that are placed at known locations on the ceiling of the instrumented rooms, and finally, computes their position through lateration. Its drawbacks of centralized control and high system maintenance cost were overcome by the Cricket system, wherein ranging distance from the beacon node (placed at predefined locations) were computed locally by the various listener nodes. Its decentralized administration with the protection of user privacy, and low system cost are its prime advantages. However, it has the drawback of limited coverage range characterized by its unidirectional ultrasonic transducers.

ALHoS removed the dependence on any fixed infrastructure (as required by its predecessors) by establishing a fully ad-hoc system with distributed localization algorithms running on every node. It overcame the unidirectional scope of the ultrasound transducers by creating an omni-directional unit by using six pairs of transducers arranged in a hexagonal pattern on the Medusa motes. This transducer array had a roughly constant response in azimuth, but a weaker response looking straight up and required an additional 9 V power source. Therefore, its design was altered in the next version Medusa-2 motes, wherein three pairs of transmit/receive transducers were inclined with respect to the base surface, while one of the transducer pair was placed at the center to provide a better response in the vertical direction. However, driving the transmit array (consisting of 4 transducers) requires four times the power to drive a single transducer, and hence increases the power consumption of the device.

The WALRUS system utilized easily available commercial off-the-shelf (COTS) components in an office environment (desktop PC with attached speakers, 802.11 wireless

infrastructure, and mobile devices) for localization. The system, though very simple, provides distance resolution with reference to the room-level, and therefore lacks in attaining fine-grained distance estimates.

In contrast to the previous systems, the Thunder and Kwon's system were developed for outdoor environments. Thunder requires a central entity to generate high-intensity (or loud) acoustic signals for its sensor nodes to receive at long distances, while the system proposed by Kwon utilized a COTS piezoelectric buzzer unit to generate acoustic signals of higher power, and it was augmented to the MTS310 sensor board interfaced with the Mica2 platform. However, a common drawback is the use of loud ranging signals that may be annoying to hear at constant intervals, and hence, is not suitable for quiet surveillance operations.

**Acoustic broadband systems.** The system proposed by Hazas et al. in [Hazas and Hopper 2006], Kushwaha et al. in [Kushwaha et al. 2005], AENSBox [Girod et al. 2006] and BeepBeep [Peng et al. 2007] are existing broadband systems. They share a common cross-correlation based signal detection technique; however, they differ in their signal design, synchronization schemes and methods to improve the received signal-to-noise ratio (SNR).

To mitigate the multipath resolution problem inherent with narrowband systems, Hazas et al. [Hazas and Hopper 2006] proposed a broadband ultrasonic localization system that was implemented on custom designed Dolphin devices. The 25 ms ranging signal was generated using a 50 kHz carrier wave modulated by Gold codes (of length 511 bits) using Binary Phase Shift Keying (BPSK). The sensitivity of the receiver was improved by using a transducer with a greater surface area (10 mm radius) rather than the general 5 mm transducer applied on the transmitter. The reported ranging results showed millimeter level accuracy that is comparable to the uncertainty in hand-measured distances, but it was only targeted for very short range ( $< 3$  m) indoor applications.

Kushwaha's system was based on the Mica2 platform with an attached custom 50 MHz DSP and an external speaker. The ranging signal was a Gaussian windowed linear chirp of 50 Hz-5 kHz. It employed a message time stamping technique. The SNR of the received signal was enhanced by adding a series of consecutive position-modulated chirps at the same phase and averaging these measurements.

The AENSBox system comprised of a custom designed acoustic sensor array that utilized beamforming to improve the received SNR, and time synchronization services to prevent clock skew and drifting. The ranging signal was a 2048-chip code modulated using BPSK on a 12 kHz carrier spread over 6-18 kHz. It differed from most of its predecessors in the use of separate synchronization service that maintained metrics to convert from one system clock to another on demand, rather than a synchronous radio and audio pulse. This approach is beneficial in scenarios where the audio range is greater than the radio range. The BeepBeep system used a 50 ms linear chirp of 2-6 kHz. It used a two-way sensing scheme (different from the round-trip time measurements) to avoid clock synchronization and was implemented on COTS mobile phones. We compare the results of our TWEET system and related characteristics to some of the related work in Section 4.4.1.

### 3. LIMITATIONS OF EXISTING RANGING TECHNIQUES

In this section, we study the limitations of existing acoustic (audible and ultrasound) ranging techniques to identify the potential areas for improvement. We distill the shortcomings of audible acoustic techniques, mostly from literature, and identify multipath and signal audibility as the two main problems of acoustic ranging. We provide an empirical case study with the Cricket ultrasonic platform. The main problems that

Table I: Mapping of Audible Intensity to Equivalent Common Sound

Acoustic System	Sound Pressure Level (dB)	Equivalent Common Sound
Thunder	73	Loud singing (at 0.9 m)
Kwon's	105	Power mower (at 0.9 m)
Kushwaha's	105 (at 0.1 m)	Power mower (at 0.9 m)
AENSBox	100	Diesel truck (at 9 m)

we observed in our experiments are the limited range and directionality of the ultrasound signal, in addition to being susceptible to multipath.

### 3.1. Limitations of Audible Acoustic Ranging Techniques

Audibility is one of the performance metrics for an acoustic ranging system. Table I shows the sound pressure level (SPL) of the acoustic pulse used in previous systems, and translates that to representative human hearing experience [Sonic Studio 1999]. The observations suggest that the audible intensity of these systems were of the magnitude that will annoy humans. Readers should note that most of the audible acoustic range-finders were designed for outdoor applications where the acoustic waveforms are not confined. For restricted and compact environments such as indoors, acoustic pulses emitted at SPL as low as 60 dB (which is equivalent to normal conversations) can be distinctly heard, and may be annoying if continued for a period of time. In addition, there is a large body of literature on ranging in indoor (room) acoustic environments [Chen et al. 2006a; Chen et al. 2006b], which attributes noise and multipath reflections as the main sources of ranging error.

### 3.2. Limitations of Ultrasound Transducers

We study the directionality, range, and multipath problems of narrowband ultrasound transducers by performing a case study of the Cricket indoor localization system. We show that, in a typical ranging setup, the first two problems are related. Specifically, by improving directionality of the Cricket mote, we can improve its coverage range. To explore the true limitations of ultrasound ranging, we design an omnidirectional extension of the Cricket receiver and study its performance empirically. We show that despite a modest improvement, the range achieved by both the original and modified versions of Cricket is quite limited.

*3.2.1. The Cricket Indoor Localization System: A Case Study.* The Cricket system estimates the range  $d$  by measuring the propagation time delay  $\delta t$  of the ultrasound signal from the *beacon* (transmitter) to the *listener* (receiver), i.e.,  $d \cong \delta t \times c$ . The accuracy of detection is dependent on  $\delta t$  that is measured by locating the leading edge of the ultrasound pulse after it crosses a preset threshold, and the speed of sound  $c$  that depends on ambient temperature and humidity. We analyze the effect of noise (due to multipath) on the coverage range and measurement accuracy of  $\delta t$ .

In case of a noise-free signal, the ultrasound pulse shape is rectangular with amplitude  $A$  and has a finite rise time  $t_r$  (as shown by the solid curve in Fig. 1(b)). However, with the addition of noise to the signal<sup>1</sup>, there is a shift  $\delta t_r$  in the time of threshold crossing that results in an error in estimating the time delay. When the SNR is large, the slope of the leading edge of the noise-induced pulse is nearly the same as the slope

<sup>1</sup>For simplicity, we assume a low frequency noise compared to the signal.

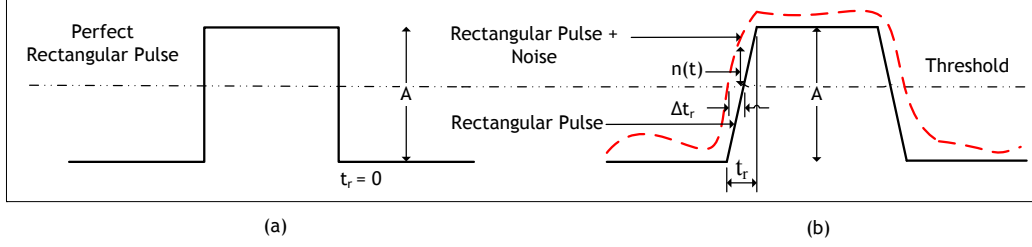


Fig. 1: Measurement of time-delay using the leading-edge threshold-based technique.

of the leading edge of the noise-free pulse<sup>2</sup>, and can be represented as:

$$\frac{n(t)}{\delta t_r} = \frac{A}{t_r} \quad (1)$$

where,  $n(t)$  is a continuous analytical function representing a random noise voltage that is characteristically “white”. Rearranging Eq. (1), we get:

$$\delta t_r = \frac{t_r}{A/n(t)}$$

The Root-Mean-Square (RMS) error of  $\delta t_r$  results in:

$$\delta T_r = [(\overline{\delta t_r^2})]^{1/2} = \sqrt{\frac{\overline{t_r^2}}{(A^2/n(t)^2)}}$$

If  $t_r$  and  $A$  are non-time varying functions, then  $[\overline{t_r^2}]^{1/2} = t_r$ ,  $[\overline{A^2}]^{1/2} = A$ , and:

$$\delta T_r = \frac{t_r}{\sqrt{(A^2/n(t)^2)}} \quad (2)$$

The denominator of Eq. (2) is the SNR of the received pulse, and equates to  $2S/N$  for a rectangular pulse [Minkoff 2002]. Therefore,

$$\delta T_r = \frac{t_r}{\sqrt{(2S/N)}} \quad (3)$$

If the receiver filter is of bandwidth  $B$ , then  $t_r = 1/B$ . If  $S = E/\tau$ , where  $E$  = signal energy,  $\tau$  = pulse width; and  $N = N_0B$ , where  $N_0$  = noise power per unit bandwidth, then:

$$\delta T_r = \sqrt{\frac{\tau}{2BE/N_0}} \quad (4)$$

Also, with the increase in range, the amplitude  $A$  of the pulse decreases due to the geometric spreading of the signal. Therefore, both range and accuracy of the Cricket system are dependent on large signal energy, where better accuracy also requires short pulses with large bandwidth.

The Cricket system provides a 12 dB signal amplification to attain a maximum transmit power. Higher amplitude signals can be generated, but at the expense of more

<sup>2</sup>The pulse shape of the noise-free pulse is rectangular as shown by the solid curve in Fig. 1(b). The pulse shape of the noise-induced pulse is near rectangular (with distortions) as shown by the dotted curve in Fig. 1(b). *This change of pulse shape from near-rectangular to rectangular depends on the SNR.* At high SNR (i.e., low noise), the pulse shape gets less distorted, and vice-versa. Therefore, in the case of minor distortions (at high SNR), the (low) noise-induced pulse is the most closest to its perfect replica of a rectangular pulse.



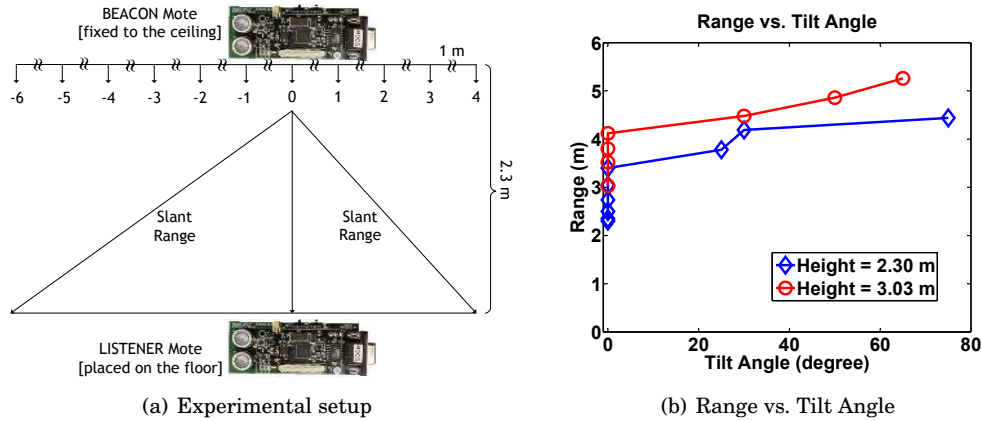


Fig. 2: The Cricket system: demonstration of the unidirectionality (or limited beam angle) of the narrowband ultrasound transducers. The tilting of the listener mote, with respect to the planar surface, does provide some improvement in range.

battery power and costlier ultrasound drive circuitry. A near perfect rectangular pulse with zero rise ( $t_r \approx 0$ ) and fall time (resulting in  $\delta T_r \approx 0$ ) can be achieved by increasing the bandwidth of the system. However, the limited bandwidth (2 kHz) of the utilized transducers does not support the generation of a large bandwidth signal, and therefore, Cricket uses a single frequency sinusoidal pulse of 40 kHz only.

We conducted the following experiment to check for the maximum signal detection range of the Cricket system within the scope of its current deployment architecture. We run two sets of experiments to demonstrate that the range and directionality problems are related. Fig. 2(a) shows the experimental setup. It was performed along the walkway inside the laboratory of dimensions  $[10 \times 1.5 \times 5]$  m. Two Cricket motes were used, wherein the first and second mote were configured as the beacon and listener nodes respectively. The beacon node was fixed near to the ceiling of the room (at a height of 2.30/3.02 m), while the listener node was placed at 11 different positions on the floor. Position 0 corresponds to the initial position of the listener, when it is placed directly below the beacon, to provide a direct LOS between them. The remaining positions, from  $[1 \rightarrow 4]$  and  $[-1 \rightarrow -6]$ , correspond to the listener position, when it is placed 1 m apart from its previous position, on either sides of the initial position 0. The range measured by the Cricket listener was the slant range; which is the path length from the beacon to the listener, rather than the horizontal range along the floor. Distance estimates were not logged by the system for position numbers  $[3, 4, -3, -4, -5, -6]$  and  $[4, -4, -5, -6]$  for the (beacon) ceiling height of 2.30 m and 3.03 m respectively. The best-case distance estimate (or the lowest estimation error) was recorded when the listener was placed directly below the beacon  $\theta = 0^\circ$ , as the sensitivity of the transducers is highest in the direction of the Z-axis, but again, the unidirectionality of the mounted piezo-electric transducers confines the coverage of the motes from  $[-40^\circ, +40^\circ]$  inclination with respect to the normal. Positions beyond this perimeter experience reduced or absolutely no coverage. The tilting of the listener mote, with respect to the planar surface, does provide some improvement in range (Fig. 2(b)), wherein previously unreachable positions can be estimated, but it is able to only partially recover from dead-spots.

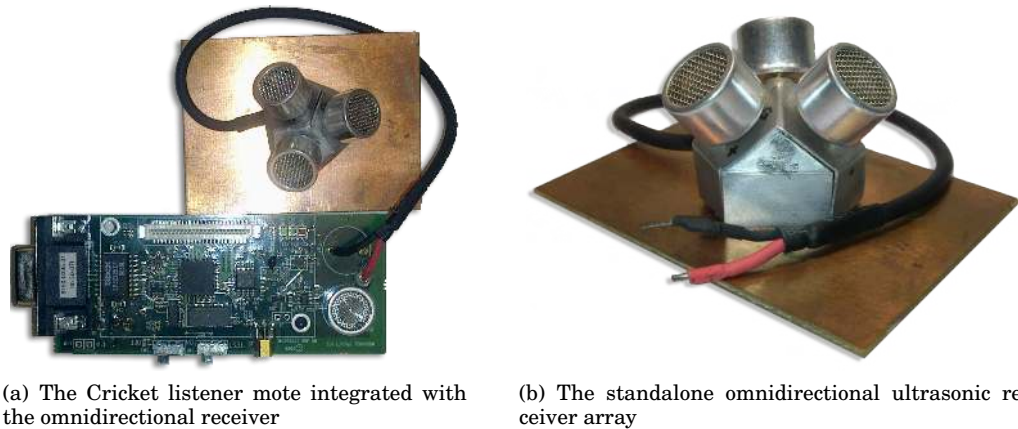


Fig. 3: The modified Cricket system.

**3.2.2. Omnidirectional Cricket.** With the understanding of the physical limitations of the Cricket system, we aim to improve the signal detection range, and hence, its coverage. A solution is to align a set of these transducers in different directions in order to create an omni-directional transducer unit, which can capture ultrasonic signals arriving from different directions, and hence, improve the reception strength of the direct LOS pulse irrespective of its azimuth angle. An improvement can also be achieved by reorganizing the transmitter unit to radiate the signal in all directions. However, providing a transmit array (consisting of  $x$  transducers) would take  $x$  times as much power to drive all the elements of the array, as to drive a single transducer. This would result in increasing the power consumption ( $x$  times) on transmit, and may require an additional power amplifier in the system. Therefore, a better alternative is to provide an omnidirectional coverage to the receiver.

While more details can be found in [Misra et al. 2011a], we present a basic overview of the improved version of the existing Cricket system. The new system is a combination of the existing Cricket mote with an omnidirectional ultrasonic receiver unit (Fig. 3). The data sheet of the ultrasonic transducers claims that the beam is typically about  $110^\circ$  wide ( $\pm 55^\circ$  wide at the half voltage points ( $-6$  dB)). The dodecahedron arrangement of the transducers was decided because the normals of the faces make an angle of about  $63^\circ$  to each other, close enough to  $55^\circ$ , so adjacent transducers would add their patterns to give a roughly constant response in the plane joining them.

**3.2.3. Evaluation.** Ranging experiments were performed with both original (Cricket) and modified (M-Cricket) platforms during the quiet period of the night in two different setups:

- **Case-A** - Indoor, High multipath: A narrow walkway ( $[10 \times 2 \times 4]$  m).
- **Case-B** - Indoor, Low multipath: A spacious corridor ( $[10 \times 10 \times 4]$  m).

The same experimental setup and procedure was followed as explained in Sec. 3.2.1. The beacon node was fixed at a height of 2.30 m. Our initial set of ranging results were promising: we could improve the Cricket range by up to 20%<sup>3</sup>. However, the results were highly dependent on the environment: only 5% improvement was observed in Case-A (note Cricket and M-Cricket (Before MPM) plots in Fig. 4). We suspected that

<sup>3</sup>The ranging error recorded at the maximum slant range was  $\approx 50$  cm.

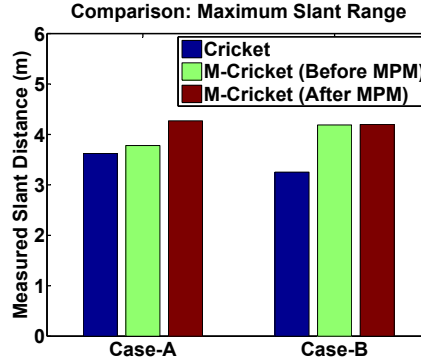


Fig. 4: Comparison of the maximum slant range: Cricket vs. M-Cricket (before multipath mitigation) vs. M-Cricket (after multipath mitigation)

multipath was the problem and implemented a Multipath Mitigation (MPM) version of M-Cricket.

An investigation into the disparity of this large range difference between Case-A and Case-B revealed interesting facts. In practice, the ultrasonic pulse emitted by the beacon node gets reflected off various surfaces inside the room, and therefore, the received signal at the listener node arrives through multiple paths. Since there is an unobstructed LOS between the devices, the receive array detects the LOS signal prior to other reflected signals, and it appears as the strongest impulse in the oscilloscope trace. The other echoes, dependent on their path lengths, arrive at different time-intervals and have varying amplitudes. We first estimate the time taken for the channel impulse response to decay completely. It was recorded to be between [800 - 1000]ms in Case-A and  $< 100$  ms in Case-B at various positions till the maximum reported range. Thus, Case-A was reported to generate more reflections with longer delay spread, which suggests that the received signal takes a longer decay time, thereby lengthening the channel impulse response. These excess delay impulses (i.e., echoes from the previous pulse that have not fallen below an acceptable level) strike against the next emitted pulse (after a random interval between [668 - 1332]ms). It creates a situation of fading or destructive signal addition where the current signal amplitude falls below the average level. This was the primary reason for the performance deterioration in Case-A.

If the time taken for the ultrasonic signal to travel the maximum range  $d$  at a speed  $c$  is at most  $d/c$ , and if the duration of ultrasonic transmission is  $t_{us}$ , then the signal must fade within time  $[d/c + t_{us}]$ . Thus, for  $t_{us} = 150\mu s$  (current pulse width utilized in Cricket) and  $d \approx 1$  m, the signal must completely die after 30ms. However, this assumption holds valid when there is only a single LOS path between the beacon and the listener, but as we have noticed, the signal reverberation time can be many times higher than the ideal situation depending on the environment.

Range estimation relies on finding the position of the first pulse. The basic requirement is that the pulse repetition frequency must be low enough so that the signal from one pulse has reduced to a small enough value by the time the next pulse is transmitted. The low ranging performance of the modified Cricket system was overcome by increasing the beacon interval time and maximum ultrasound time-of-flight to randomly choose within the range [1500 - 2000] ms and 65 ms respectively. With the introduction of these new parameters, the maximum range in Case-A increased to 4.27 m (an im-

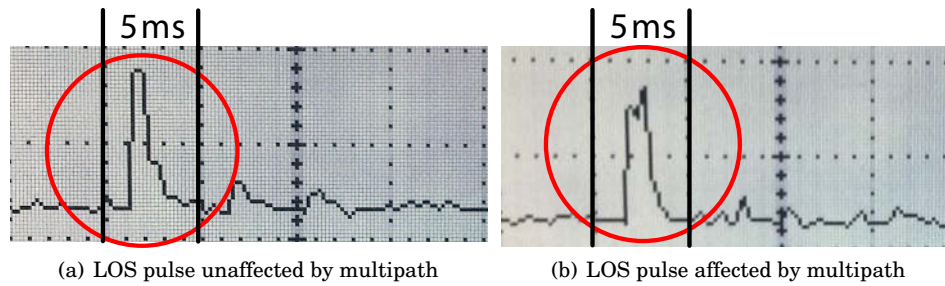


Fig. 5: Illustration of the widening of the narrowband pulse under the influence of multipath echoes.

provement of 18% over the Cricket system), while there was no change in the result for Case-B (shown as M-Cricket (After MPM) in Fig. 4).

**3.2.4. Discussion.** The modified Cricket system is similar in design to the ultrasonic board of the Medusa motes, but we only implement a omni-directional receiver unit rather than a omni-transmit/receive unit with lesser hardware components, and achieve better coverage range of more than 4 m with no additional power consumption. This system can be improved by providing multiple transducers, as we did, but instead of wiring them together, we could provide for the best one with the aims of increasing the signal strength, and reducing the influence of reflections. This would apply specifically to the transmitter, which would waste power if multiple transducers were connected together to operate at the same time. This was the main reason for building a receiver array only. Switching of the transducers was an approach that we considered and might still have good direction, although it would require both new electronics hardware and developing new protocols in the nodes.

The measurement of the time delay using the leading edge threshold-based technique, though simple and cost-effective, results in large estimation errors due to its high susceptibility to environmental noise that greatly impacts the pulse characteristics. Most of the energy in the multipath echoes is not resolvable and the received signal is reduced to a single wide pulse (Fig. 5). In addition, the use of a narrowband signal (single frequency sinusoidal) exhibits an inherent limitation characteristic to its design as they have unity product of bandwidth  $B$  and pulse time duration  $T$ , which creates a trade-off between range and resolution.

Range resolution depends on the bandwidth of the ranging signal. Narrow pulse width provides superior resolution and accuracy, but a small range. To achieve long distance, the transmitted signal should have larger values of  $E/P$  (i.e., signal energy / power spectral density). However, the amplitude of the signal reaching long range targets is considerably low and consequently has low  $E/P$ . Low-cost and low-power embedded systems (such as Cricket) have limitation on the maximum transmission power. Hence, sending high amplitude signals cannot be achieved beyond a certain threshold. An alternative is to send signals at a certain power levels ( $\approx$  amplitude) while increasing the duration of transmission  $T$ . However, increasing  $T$  would lead to a decrease in  $B$  which is a prime factor for resolution. A specific signal processing technique called *pulse compression* combines the benefits of higher energy of a longer pulse width with the high resolution of a short pulse width, and can be effectively applied to increase the ranging capability with lower transmission power. There are numerous waveforms suitable for use with pulse compression, and will be discussed in the following section.

#### 4. BROADBAND ULTRASONIC RANGING: SIGNAL DESIGN, DETECTION AND IMPLEMENTATION

We apply the lessons learned from the previous section to guide the design of the ultrasonic ranging signal. To achieve robustness against multipath and to improve signal detection latency, we selected a broadband signal with the bandwidth of 5 kHz. The frequency of the signal is selected just above 20 kHz, the maximum frequency audible to humans. This is significantly lower than the frequency of traditional ultrasonic techniques, which helps to improve directionality and coverage range of our technique. We further improve directionality limitations by using an omni-directional receiver on the receiver end.

##### 4.1. Signal Design and Analysis

Based on prior work in the field of acoustical localization in air, two classes of broadband signal designs were identified: *chirps* and *pseudorandom noise* (PN). Chirps are frequency modulated signals, wherein a sinusoidal wave of constant amplitude sweeps the desired bandwidth  $B$  within a certain time-period  $T$  in a linear or non-linear (for example following quadratic or logarithmic laws) manner. On the other hand, PN signals are (phase) modulated sinusoidal waveforms mixed with pseudorandom sequences. Broadband signals are detected using a *matched filter* implemented by correlation with a reference signal. A noteworthy point here is that the time-period  $T$  and bandwidth  $B$  of the signal control the output parameters of the filter, and have a key role in delivering the desired coverage range and resolution. In the following subsection, we explain this relationship for linear chirps, and then present comparison studies for the remaining waveforms.

*4.1.1. Analysis of Linear Chirp Waveform.* A linear chirp is represented by the bandpass signal:

$$s(t) = \begin{cases} \cos(2\pi(f_0 t \pm \mu \frac{t^2}{2})) & 0 < t < T \\ 0 & \text{elsewhere} \end{cases} \quad (5)$$

where,  $f_0$  is the center frequency in Hz, and  $\mu = B$  (Hz) /  $T$  (s) is the chirp rate that sweeps linearly from  $(f_0 - B/2)$  to  $(f_0 + B/2)$  between the time interval  $[0, T]$ . The  $\pm$  term defines its sweep direction.

When the signal in Eq. (5) is passed through its matched filter, the following output is generated [Cook and Bernfeld 1967]:

$$g(t) = \frac{T}{2} \cos(2\pi f_0 t) \left[ \frac{\sin(\pi \mu t (T - |t|))}{\pi \mu T t} \right] \quad \text{where } 0 < t < T \quad (6)$$

This is the approximate autocorrelation of the linear chirp  $s(t)$  and it provides two important results:

- The *peak value* (that signifies the energy of the signal) occurs at  $t = 0$  and is proportional to  $T$ .
- The *correlation envelope*, expressed as  $\left[ \frac{\sin(\pi \mu t (T - |t|))}{\pi \mu T t} \right]$  is approximately  $\left[ \frac{\sin(\pi \mu t T)}{\pi \mu T t} \right]$  for  $t \ll T$ , with its first zeros at  $t = \pm 1/(\mu T) = \pm(1/B)$ ; and is inversely proportional  $B$ .

This gives the *important* relationship that an increase in  $T$  increases the size of the post-correlation signal, and an increase in  $B$  gives better resolution by narrowing the envelope of the correlation peak.

*4.1.2. Comparison of Broadband Waveforms.* In this subsection, we compare the features of various types of broadband chirps based on  $B$  and  $T$ . For studying the *change of*

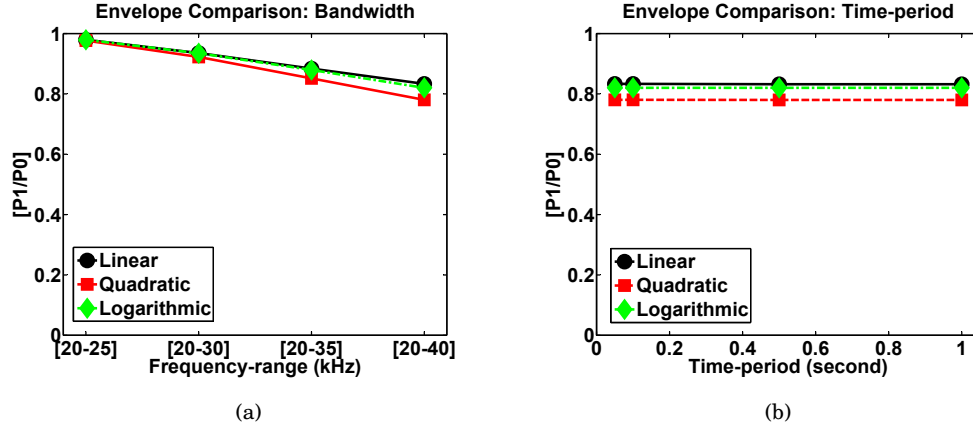


Fig. 6: Study of the change in (a) bandwidth and (b) time-period on the characteristics of the correlation envelope for linear, quadratic and logarithmic chirps.  $[P1/P0]$  denotes the ratio between the height of the first sidelobe  $[P1]$  to the correlation peak  $[P0]$ .

*bandwidth*, four types of chirp with constant time-period of 1 s and varying frequency range (and thus bandwidth) were designed: Chirp-1 [20-25 kHz], Chirp-2 [20-30 kHz], Chirp-3 [20-35 kHz] and Chirp-4 [20-40 kHz]. Fig. 6(a) shows the ratio between the height of the first sidelobe  $[P1]$  to the correlation peak  $[P0]$  denoted as  $[P1/P0]$  for linear, quadratic and logarithmic chirps for the different types (chirp-[1/2/3/4]). These peaks are related to the envelope of the correlation output, which is an important factor. A lower ratio of  $[P1/P0]$  signifies a narrower correlation envelope and is best supported by the highest bandwidth signal (i.e. Chirp-4). The linear and logarithmic chirps have similar correlation envelopes; however, the envelope cover of the quadratic chirp is even narrower. This suggests that although  $B$  does not define the acoustic pressure level of the chirp, a higher bandwidth signal is preferable due to its narrower correlation envelope that can improve the resolution of the range measurement.

Similarly, for studying the *change in time-period*, four types of chirp were designed with constant bandwidth of 20 kHz and varying time-periods: Chirp-1 [1s], Chirp-2 [0.5s], Chirp-3 [0.1s] and Chirp-4 [0.05s]. Fig. 6(b) shows that  $[P1/P0]$  is constant for all the different chirps, which suggests that the correlation envelope is independent of the time-period irrespective of the chirp type. As  $T$  controls the peak value of the correlated signal, one may consider choosing a longer signal duration that has higher energy to travel longer distances. However, a longer ranging signal increases the system reaction time, wherein the pulse repetition frequency has to be kept low; which means that the entire system is required to wait for one signal and all its echoes to de-

Table II: Chirps vs. PN Signal Characteristics

Signal Type	$[P1/P0]$	$[P2/P0]$
Chirp (Linear)	0.8332	0.4358
Chirp (Logarithmic)	0.8201	0.3961
Chirp (Quadratic)	0.7802	0.3030
Pseudonoise	0.8106	0.6211

can before transmitting the next pulse. Second, it increases the overall system cost, in terms of processing time, energy consumption and storage, thereby making its implementation difficult on resource-deficient sensor nodes.  $B$  determines the width of the correlation envelope, and therefore, determines the range resolution. Although, working in the ultrasonic domain provides the flexibility to use band of signal frequencies above 20 kHz, higher frequencies are more vulnerable to atmospheric absorption. This limits the use of an ultrawide bandwidth ultrasonic signal. The appropriate choice of  $B$  and  $T$  depends on the application requirements, but from the study presented in the previous subsection, it appears reasonable to choose a broadband signal of the highest bandwidth (for best detection accuracy) and shortest time-period (for long range incurring the least processing cost).

With regards to the choice between linear/nonlinear chirps and PN signal, we generated a 20-40 kHz/1 s pulse for each category, and compared them on the basis of their individual envelope cover (height of the first [P1] and second sidelobe [P2] to the correlation peak [P0]) and spectral complexity. For a PN signal of certain  $B$  and  $T$ , we observed that the correlation (peak and sidelobes) properties vary across different pseudorandom numbers, and therefore, we calculated the running average across 1000 randomly chosen PN codes.

Table II summarizes the overall statistics<sup>4</sup>. A lower value of [P1/P0] and [P2/P0] signifies a narrower signal envelope, and is best supported by chirp waveforms than PN signals. Of the different types of chirps, we choose a *linear chirp* (as our ranging signal) since it allows Doppler measurements (that is useful in tracking) despite its higher [P1/P0] and [P2/P0] value.

#### 4.2. Signal Detection and Post-processing

The system is presented with an indoor environment using a reverberation geometrical acoustic model<sup>5</sup> [Crocker 1998]. For the mathematical formulation, we adopt the following notation:  $s(t)$  and  $d(t)$  represent the signal emitted by the transmitter (Tx) and received at the receiver (Rx) respectively; the respective impulse responses of the transmitter, environment (channel) and receiver are represented by  $h_{tx}(t)$ ,  $h(t)$  and  $h_{rx}(t)$ ; and the white Gaussian noise in the channel is denoted by  $v(t)$ . We also assume that the system is linear and time-invariant.

$s(t)$  is a broadband signal in the form of a linear chirp and is transmitted at  $t = 0$  by Tx. The signal  $d(t)$  received at Rx is the convolution:

$$d(t) = s(t) * h_{tx}(t) * h(t) * h_{rx}(t) + v(t) * h_{rx}(t) \quad (7)$$

Assuming  $h_{tx}(t)$  and  $h_{rx}(t)$  are of unit magnitude (i.e., neither the transmitter nor the receiver change the signal characteristics):

$$d(t) = s(t) * h(t) + v(t) \quad (8)$$

<sup>4</sup>With respect to a cross-correlation based detection mechanism, narrowband signals result in a quasi-periodic output where it is nontrivial to distinguish unambiguously between adjacent peaks of the correlation function. Typically, signal bandwidth of 2-5 kHz (for ultrasonic narrowband systems) with respect to the center frequency of 40 kHz is only a small fraction of [0.05-0.125] which is  $\ll 1$ . As a result, adjacent peaks have very nearly equal heights, and hence, identification of the tallest correlation peak either requires a very large SNR or a long observation time-period - none of which are not desirable.

<sup>5</sup>Geometrical acoustic model is an approximation to the wave acoustic model, and is valid if: (a) the dimensions of the enclosure are large compared to acoustic wavelengths, and (b) the considered acoustic signal is broadband [Vorlander 2001].



$h(t)$  is modeled as the sum of  $M + 1$  impulses corresponding to the direct path with propagation delay  $\tau_0$  and  $M$  other possible paths between Tx and Rx as:

$$h(t) = \sum_{i=0}^M A_i \delta(t - \tau_i) \quad (9)$$

where  $A_i$  is the amplitude of the  $i$ -th ray and  $\delta(t - \tau_i)$  represents the delay in propagation from Tx to Rx. Ray  $i = 0$  is defined here as the direct sound ray from the source to the receiver, and rays  $i > 0$  are defined as reflected rays.  $\tau_i = d_i/c$ , where  $d_i$  is the distance traveled by ray  $i$  and  $c$  is the speed of sound under room conditions.

The received signal  $d(t)$  is processed using a matched filter implemented by correlating it with a reference signal  $s(t)$  (i.e., a locally stored copy of the original emitted signal), and result in:

$$\begin{aligned} y(t) &= [d(t) * s(t)] \\ y(t) &= [s(t) * s(-t)] * h(t) + v(t) * s(-t) \end{aligned} \quad (10)$$

$y(t)$  has its earliest component  $[s * s](t - \tau_0)$  (where:  $*$  implies correlation), whose peak can be used to determine  $\tau_0$  (direct path signal) with considerable precision, provided the other multipath components of  $d(t)$  are sufficiently weak and/or separated in time from  $t = \tau_0$ . The noise term  $v(t) * s(-t)$  may shift the peak at  $\tau_0$  from its actual timeline, which may result in an inaccurate estimate of the range information.

The signal detection scheme discussed in existing work provide resistance to multipath and low-noise signals [Kushwaha et al. 2005; Hazas and Hopper 2006; Girod et al. 2006; Peng et al. 2007] by using a peak detection approach, which under the condition of a direct line-of-sight (DL0S) between the transmitter and receiver, identifies the first tallest correlation peak that exceeds a preset threshold. However, from our study and observations, we noticed two potential problems. First, the correlation peak is surrounded by numerous sidelobes (i.e., adjoining peaks). Second, the correlation plot obtained from processing the band-limited signal is highly oscillating within its envelope cover. Both these conditions provide an inaccurate estimate of the correct detection peak in the vicinity of similar peaks of approximately equivalent heights under noisy conditions. Therefore, we propose a simple envelope detection mechanism that makes the role of sidelobes irrelevant and counters the effect of noise through the least-square curve fitting approach. In addition, it also provides the benefit of finer resolution that can be fractions of a sampling period.

The envelope detection approach estimates the maximum value of the envelope of the compressed (correlated) pulse that should give the best estimate of its position. A simple least-squares approximation technique has been used to find the envelope of the rectified signal, rather than the standard approach of calculating the magnitude of the analytical signal. The algorithm identifies the position of the local peak  $(t_2, y_2)$  that is greater than its left and right neighbor peaks at  $(t_1, y_1)$  and  $(t_3, y_3)$ , finds the parabola that passes through these points exactly, and finally, calculates the time coordinate of the maximum of this parabola ( $t = t_{peak}$ ). Therefore, fitting a parabola to these three points [Boucher and Hassab 1981; Moddemeijer 1991; Jacovitti and Scarano 1993; Jameson 2006] requires solving the following system of three linear equations for the three unknown  $[a, b, c]$ :

$$\begin{aligned} y_1 &= at_1^2 + bt_1 + c \\ y_2 &= at_2^2 + bt_2 + c \\ y_3 &= at_3^2 + bt_3 + c \end{aligned} \quad (11)$$



The corresponding representation in matrix form is:

$$\mathbf{A}\hat{\mathbf{x}} = \mathbf{B} \quad (12)$$

where  $\hat{\mathbf{x}} = [a \ b \ c]^T$  is the matrix of unknown parameters, and:

$$\mathbf{A} = \begin{bmatrix} t_1^2 & t_1 & 1 \\ t_2^2 & t_2 & 1 \\ t_3^2 & t_3 & 1 \end{bmatrix} \quad \mathbf{B} = [y_1 \ y_2 \ y_3]^T \quad (13)$$

Thus,  $\hat{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{B}$ , where  $\mathbf{A}^{-1}$  is the inverse matrix of  $\mathbf{A}$ . The maximum of the envelope occurs at  $t_{peak} = -b/(2a)$ .

In case of low-noise signals, there are more peaks surrounding the highest peak as shown in Fig. 7(a). The parabolic curve fitting using least-square approximation technique does the best to pass as near as possible to all the adjacent peaks, and thus, provides resistance to noise on the data points. To illustrate the least-square approximation process, suppose there are  $n$  data points that can be modeled by a system of  $n$  quadratic equations for the three unknown coefficients  $[a, b, c]$ . If  $n$  is greater than the number of unknowns (i.e., 3), then it is an overdetermined system, and is solved by the least-square parabolic fitting process that minimizes the summed square of the residuals.

Let the difference  $e_i$  between the  $i^{th}$  data point  $(t_i, y_i)$  and the fitted parabola be:

$$e_i = y_i - (at_i^2 + bt_i + c) \quad (14)$$

Then, the sum of squared errors is given by:

$$E = \sum_{i=1}^n e_i^2 \quad (15)$$

The goal is to minimize  $E$ , and is determined by differentiating  $E$  with respect to each parameter (or unknown), and setting the result to zero (i.e.,  $\partial E/\partial a = \partial E/\partial b = \partial E/\partial c = 0$ ).

Thus, we obtain the following three equations for the three unknowns  $[a, b, c]$ :

$$\sum_{i=1}^n y_i t_i^2 = a \sum_{i=1}^n t_i^4 + b \sum_{i=1}^n t_i^3 + c \sum_{i=1}^n t_i^2 \quad (16)$$

$$\sum_{i=1}^n y_i t_i = a \sum_{i=1}^n t_i^3 + b \sum_{i=1}^n t_i^2 + c \sum_{i=1}^n t_i \quad (17)$$

$$\sum_{i=1}^n y_i = a \sum_{i=1}^n t_i^2 + b \sum_{i=1}^n t_i + cn \quad (18)$$

This linear system can be solved (as explained before) for  $[a, b, c]$  to provide an estimate for the position of the peak:  $t_{peak} = -b/(2a)$ .  $t_{peak}$  is the estimate of the pulse position, and thus, provides a range estimate.

We simulated a custom environment and evaluated the performance of the proposed ranging algorithm. The simulator was designed to construct a virtual 2D rectangular room with (top,left) and (bottom,right) coordinates as:  $(-5, \zeta/2)$  and  $(\zeta + 5, -\zeta/2)$  respectively, where  $\zeta$  is the distance between Tx and Rx and is varied from 1-20 m for every set of measurements. Tx and Rx were placed at positions  $(0, 0)$  and  $(\zeta, 0)$ . It was

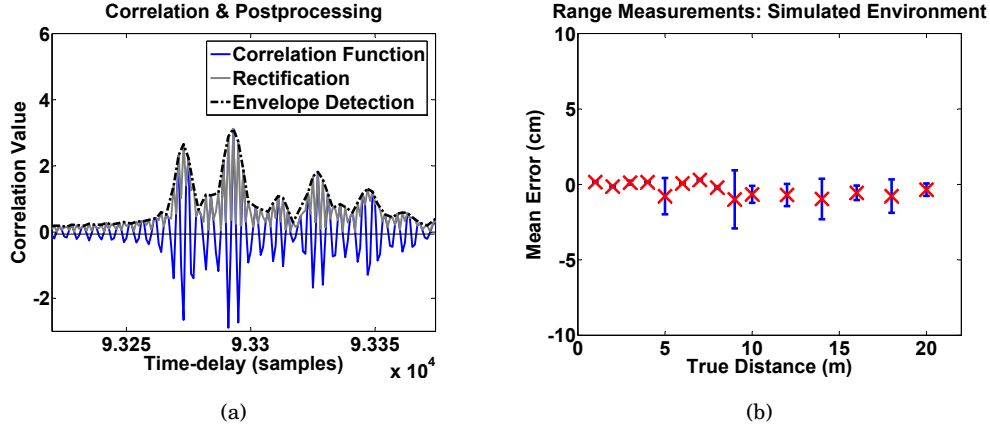


Fig. 7: (a): Signal detection and post-processing. Simulation results for range estimation. (b): Mean error with vertical bars representing 95% confidence intervals.

configured to generate fixed number of reflection points at random positions in the enclosed geometry and was programmed as per the described system model.

Measurements were taken at different positions inside the room for distances between 1-20 m. Every simulation was run for a length of 1000 iterations. The simulator was configured for 5 reflection points with attenuation factor of 0.9, and the transmitted signal of [20-40] kHz/50 ms (sampled at 96 kHz) was added with white Gaussian noise (SNR = 10 dB). The choice for these simulation parameters has been explained in [Misra 2012].

Fig. 7(a) shows the output from the correlation function, the result of rectification and envelope detection of the correlated pulse. Fig. 7(b) shows the distance estimation accuracy obtained from the simulation measurements. We observe that the magnitude of the mean errors is consistently less than 1cm for distances upto 20 m. Therefore, we conclude that our detection methodology is precise enough for fine-grained ranging.

#### 4.3. TWEET: System Implementation

We, initially, developed a proof-of-concept PC-based ranging system consisting of various COTS devices and custom designed units to experimentally verify the feasibility of our proposed scheme before incorporating them into an embedded design. Based on its understanding, we improved on its various stages to finally assemble all the different components onto a single mote-based ranging system named as the TWEET. In this section, we describe its hardware platform and software architecture.

*4.3.1. System Design: Hardware & Software.* TWEET has been implemented on CSIRO Audio nodes using its wireless sensor network platform: the Fleck-3z. Its main components include the Atmega1281 microcontroller, 1MB external flash memory and a low-bandwidth Atmel RF212 radio transceiver operating in the 900 MHz band. It supports a flexible range of digital I/Os and a daughter board interface, which allows the use of expansion boards to enhance its base functionality. The architecture relies heavily on the SPI bus, where the microcontroller acts as the SPI master and communicates with the remaining system components over the SPI interface. Fig. 8 and Fig. 9, respectively, show the architecture of the implemented TWEET system and its different components.

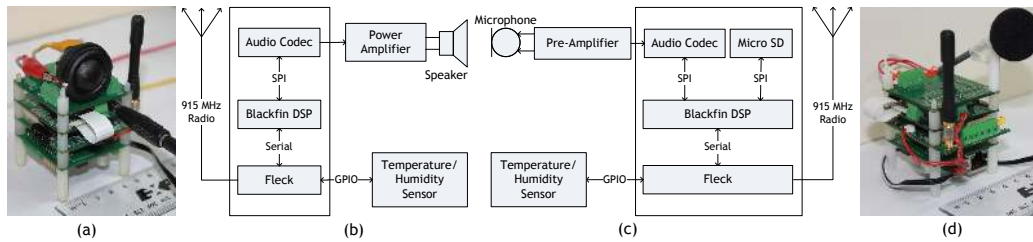


Fig. 8: The TWEET ranging system: architecture of the (a)-(b) beacon and the (c)-(d) listener.

The audio signal processing daughter board (designed by CSIRO) was used for ultrasonic ranging. It includes four TI TLV320AIC3254 audio codecs, each providing two audio I/O channels along with internal functionalities such as programmable gain amplifiers and software configurable filtering; micro SD flash memory card slot, and a connector to hold the CM-BF537E digital signal processor module manufactured by Bluetechnix. The CM-BF537E module combines a (Analog Devices) Blackfin DSP running at up to 600 MHz, a 32 MB RAM and an Ethernet interface. The DSP communicates with the low power Fleck mote through a serial interface. The power consumption of this daughter board is in excess of 1200 mA in the active state, and so (in its present implementation) Fleck-3z mote controls power to this board ensuring that the relatively high power consumption is only incurred during audio transmission and acquisition. There are two simple daughter boards that provide connector access to the audio I/O ports and an Ethernet socket.

The transmitting front-end of the beacon mote consisted of a power amplifier driving a tweeter transducer (VIFA 3/4" tweeter module MICRO), which is a speaker optimized for high frequencies. It was chosen due to its small size ( $[2 \times 2 \times 1]$  cm) and good high-frequency response compared to existing broadband transducers or piezoelectric ceramic/piezo film transducers reported in existing literature. The amplifier is a lightweight portable unit with a maximum output power of 0.5 W. It is powered by batteries and has a tunable gain controller.

The receiving front-end of the listener mote consisted of the Knowles microphone (SPM0404UD5 [Misra et al. 2011b]) fixed to the pre-amplifier PCB designed by CSIRO. The surface mount wideband ultrasonic sensor was chosen due to its omni

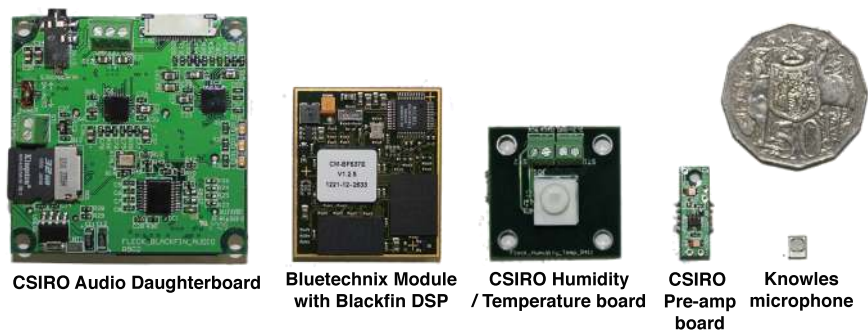


Fig. 9: TWEET components.

Table III: Power consumption of the audio node

Process	Power (mW)
Mote Idle	$60 \times 10^{-3}$
Mote Tx/Rx	60
Audio codec	4.2-21.9
DSP Startup/ Processing/ Idle	1200

directionality, high sensitivity and SNR within an extremely small form factor of  $[4.72 \times 3.76 \times 1.40]$  mm. The small PCB has been designed to operate in close proximity with the microphone in order to minimize the susceptibility of the low amplitude microphone output signals to corruption by electrical noise. The frequency response characteristics, for both the transmitting and receiving front-ends, have an approximately 20 dB acoustic pressure level above the noise floor for frequencies between [20-40] kHz [Misra 2012].

As temperature compensation is required for range measurements, a small form-factor PCB ( $[2.5 \times 2.5]$  cm) was designed to mount the Sensirion SHT15 temperature and humidity sensor (along with a filter cap and necessary discrete components such as capacitors and pull-up resistors), and was controlled by the Fleck-3z microcontroller via a GPIO digital interface. It consumes  $< 5$  mA of current, thus allowing it to be powered directly from the digital I/O ports of the mote.

Fleck-3z runs the TinyOS-2.x OS. The software performs the tasks of maintaining and executing a schedule of system operations, maintaining a persistent log of system actions and status, sampling from attached on-board/external sensors, controlling the operation and power switch of the audio signal processing daughter board. The software for the Blackfin DSP is responsible for configuring (sampling rate, gain, etc) and enabling the audio codec ICs, managing the incoming and outgoing digital audio stream, transferring data/information to the micro-SD card, command exchange from the Fleck-3z via serial interface, such as start/stop audio playback/recording, interrogate operational status, etc. The power consumption of the different components in the audio node are shown in Table III.

*4.3.2. Ranging Methodology.* For the TWEET system, a [20-25] kHz/50 ms ranging signal was chosen and the audio codes were configured to sample at 64 kHz. Although the audio codecs could support a maximum sampling rate of 192 kHz that can generate a signal upto 96 kHz, our system tests revealed that there was a significant drop in the audio output of approximately 30-40dB beyond 25 kHz frequency range.

TWEET uses the TDOA of RF and ultrasound signals to measure the beacon-to-listener distances. The beacon periodically transmits a RF message containing the measured ambient temperature and humidity. At the start of each RF message, the beacon transmits a broadband ultrasonic linear chirp pulse. The fast propagating RF signal leads its synchronous ultrasound pulse and reaches the listener almost instantaneously, which then measures the TDOA between them. The TOA of the ultrasound pulse is measured by cross-correlating the received chirp pulse with a copy of the reference signal stored in the receiver, and then, the range estimate is computed by the envelope detection technique (Section 4.2). Since the speed of sound has a relatively large sensitivity to temperature variations than relative humidity and atmospheric pressure, the final distance estimate is computed by the corresponding speed of sound obtained by averaging the ambient temperature measured at the beacon (sent in the RF message) and the listener (measured at the TOA of the ultrasound pulse). After carefully estimating the various system induced time-delays, a final calibration exercise was performed by conducting a series of ranging experiments for short distances between

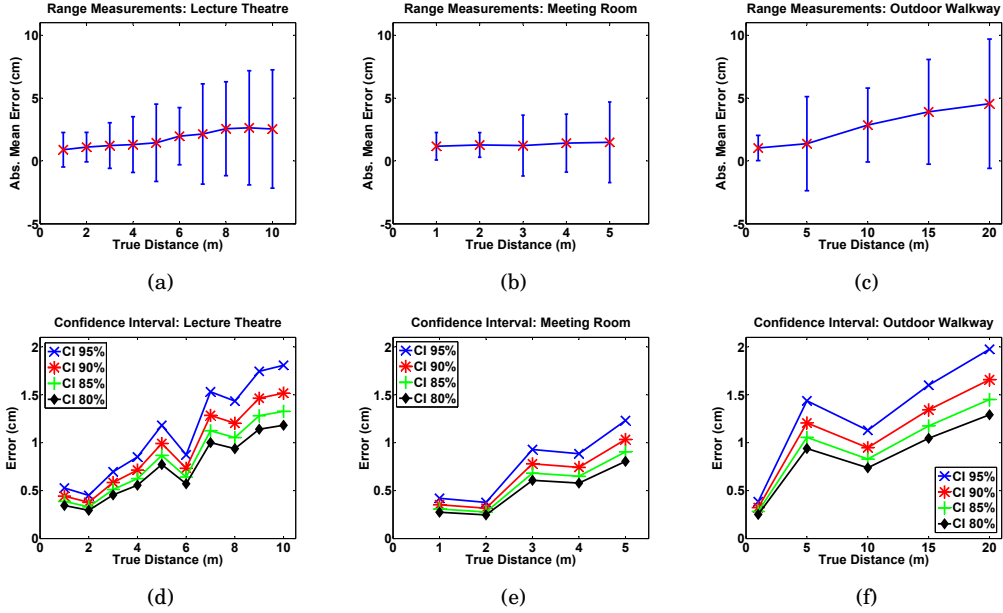


Fig. 10: The TWEET ranging system: accuracy in terms of mean error, deviation (shown in blue vertical lines), and confidence intervals for the different experimentation cases.

5-10 cm. The mean ranging error of 6.42 cm, obtained from its corresponding error distribution, was subtracted from the final result.

#### 4.4. Evaluation

To evaluate the performance of the TWEET system under different multipath conditions, we conducted ranging experiments in the following environments:

- **Case-C** - Indoor, Low multipath: A quiet lecture theatre ( $[25 \times 15 \times 10]$  m) with a spacious podium at one end of the large room.
- **Case-D** - Indoor, High multipath: A quiet meeting room ( $[7 \times 6 \times 6]$  m) with a big wooden table in the center and other office furnitures.
- **Case-E** - Outdoor, Very low multipath: A less frequently used urban walkway, and the weather being sunny with occasional mild breeze.

In all our experiments, the beacon mote was fixed while we performed a controlled and careful movement of the listener mote along the direct LOS using a measuring tape and markers for establishing the correct ground distance. The beacon was calibrated to chirp at 70 dB. The speed of sound used in distance calculation was according to the model:  $c_{air} = 331.3 + 0.6\theta$  ( $\theta$ : air temperature in  $^{\circ}C$ ). For every setting (i.e., different distances under different test cases), the experiments were repeated 30 times. The metrics used to evaluate the system were accuracy (difference between the ranging results and the true distance) and confidence interval for the measured errors. The accuracy and confidence measurements for the case-C setting are shown in Fig. 10(a) and Fig. 10(d), where we can observe that our system yields accurate and stable ranging results in a (less severe) multipath environment. The mean ranging error is within  $\pm[1-2]$  cm with a 95% confidence interval of  $< 2$  cm. High percentage of experiments recorded less than 2 cm accuracy, however, the performance deteriorates for distance

measurements at [9-10] m when the listener mote approaches close to the walls. Even then, the error deviation from the mean is  $< 5$  cm. Fig. 10(b) shows the accuracy for case-D, where the mean ranging error is 1.5 cm and the maximum deviation is 2.5 cm at its maximum measured distance of 5 m. Due to the multipath dominated environment, the reported error levels (for each measurement) is quite fluctuating, even for shorter distances. Nevertheless, the system was able to record a confidence interval of  $< 1.5$  cm (Fig. 10(e)).

The ranging statistics for case-E has been shown in Fig. 10(c) and Fig. 10(f). The system shows similar performance as reported for indoor case-A/B for distances  $< 10$  m with a maximum mean error and deviation of  $\approx 2.5$  cm and an estimated confidence interval of 2 cm. We also observe that the ranging error increases and shows a larger dynamic range for distance measurements at 15 m and 20 m, which is primarily due to the a lower SNR of the received signal caused by attenuation and non-uniformities in the atmosphere caused by wind. The measurements at 5 m shows a sudden drop in accuracy and an increased deviation, which is due to a strong breeze at that instant, but the system quickly recovers and attains a stable mean error.

For all the measurements, we observe that the absolute distance estimation error increases with the increase in separation between the transmitter and the receiver. In practice, distance information is not known a priori, therefore we plot the distribution of all ranging errors in the different test environments (reported in [Misra et al. 2011c]) so as to provide an overall system snapshot. The statistics suggest that the overall performance of TWEET is accurate with a mean error of 1 cm and deviation of 3.63 cm with the best performance obtained in indoor spacious facilities.

*4.4.1. Discussion.* Our work has similarities to the linear chirp based system designed by Kushwaha [Kushwaha et al. 2005]; but instead of using an additional Gaussian window to compensate for high correlation sidelobes, we use a simpler and effective envelope detection method. Our ranging precision (after temperature compensation) is generally 3 times better, where for ranges between [10-20] m, our standard deviation is 5 cm compared to [15-25] cm reported by Kushwaha et al. Further, Kushwaha's system achieved a maximum detection range of 30 m at a SPL of 105 dB at 10 cm (i.e., measured at the near-field of the speaker). Such near-field measurements may not provide the correct SPL representation, as they are dominated by the physical dimension of the speaker membrane and the volume of displaced air. Our experiences with two different speakers of different sizes used with our system (wherein we supplied the same power) showed different SPL values when measured in the near-field (10 cm), but were the same at 1 m. In contrast, TWEET attained at maximum range of 30 m, but its operational range was 20 m at SPL of 70 dB (SPL measured at 1 m).

Our system reports, approximately, the same level of accuracy as the BeepBeep [Peng et al. 2007] for ranges under 10 m. The authors do not provide any ranging analysis for distances over 10 m, which is where our system is more useful. There is no mention of the SPL for this system, therefore, comparing the ranges would be unfair. The techniques of EchoBeep and DeafBeep [Nandakumar et al. 2012] and Whistle [Xu et al. 2011] improve on the basic BeepBeep mechanism. While EchoBeep and DeafBeep, respectively, are provisioned for NLOS conditions and distance difference

Table IV: Resource Usage

System	Power (W)	Approx. Processing Cost
TWEET	2.5	2*FFT + 1*IFFT
AENSBx	3.6	5*FFT + 4*IFFT

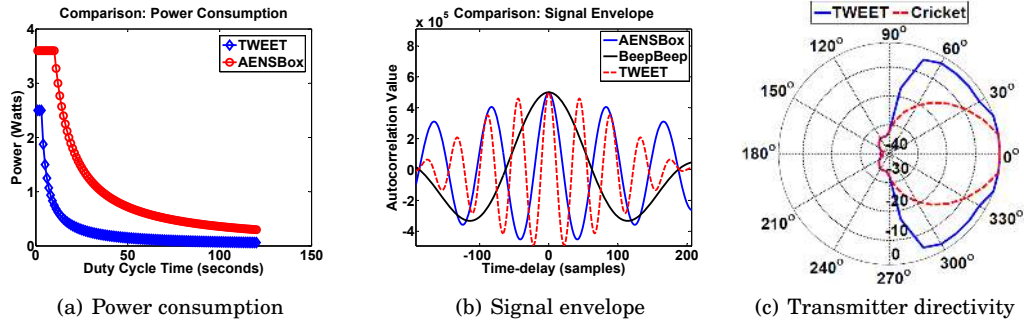


Fig. 11: System Comparison.

measurements (for devices without microphones), Whistle measures the TDOA. In the current implementation, TWEET has been designed to provide accurate TOA ranging (to within a few cm) in LOS environments, but will under perform in NLOS conditions. Nevertheless, TWEET's envelope detection mechanism can prove to be more efficient in catering to the above ranging conditions as the core idea of EchoBeep, DeafBeep and Whistle are based on peak detection, the disadvantages of which have been discussed in Section 4.2. Yang et al. [Yang et al. 2011] detect the arrival time of the ranging pulse by the sequential change-point mechanism, where the key idea is to identify the first arriving signal that deviates from the noise after filtering out background noise. This simple technique is useful when there is relatively little noise and interference from outside sources in the frequency range of the ranging pulse (such as inside a car), and therefore, cannot be used for general ranging applications. The results presented in [Yang et al. 2011] denote the distance difference from the receiver to two transmitters, and hence, a direct comparison with the TWEET system is not feasible.

The AENSBox system [Girod et al. 2006], calibrated to 105 dB SPL at 10 cm, attains a maximum range of 60 m with a mean error of 1.73 cm and deviation of 1.76 cm. It is 3 times better than the TWEET system in maximum operational range, but the accuracy is comparable to ours with a mean of 1 cm and deviation of 3.63 cm. This additional benefit in range comes at the cost of using a longer (1/3 s) ranging signal and an expensive matched-filter. However, under the condition of duty-cycling (Fig. 11(a)), our system scores over AENSBox in consuming 80% lesser power. Table IV compares the various resource usage between these two systems, where again, TWEET proves economical in saving power and processing cost. It is important to note that AENSBox was designed for outdoors, and therefore, the various design choices were made for delivering high power.

A comparison of the signal envelope (an important factor for precision) among TWEET, AENSBox and BeepBeep has been shown in Fig. 11(b), where we consider the signal length of the various systems to be 1 s (for easy analysis). We observe that TWEET has a narrower signal envelope than the other systems, which implies that if the same SPL is generated from all, then our system would be able to attain comparable detection range, but with better accuracy. In addition, the TWEET system reported better transmit directionality (Fig. 11(c)) than the Cricket system analyzed in Section 3.2.1.

The applicability of the parabolic curve fitting using the least square approximation technique used in TWEET can get limited by the SNR condition. If the SNR is very low, then the proposed interpolation method will introduce significant bias. In the outdoor experimentation environment, TWEET was able to cover a maximum range of 30 m, but its (useful) operational range was limited to 20 m only. The low SNR recorded

at distances beyond 20 m was a prime hurdle that affected the performance of cross-correlation, which in turn impacted the proposed envelope detection method.

The different system tests with TWEET showed that its performance was similar to AENSBBox [Girod et al. 2006] (operating below 20 kHz) except for improvements in power consumption, processing cost and (minor) accuracy. Although the performance of ultrasound was similar to audible signals in terms of accuracy, it provided major benefits in terms of inaudibility to human perception, which is a critical factor for urban deployments.

## 5. NEAR-INAUDIBLE TRACKING OF MOBILE TARGETS USING ACOUSTIC DOPPLER SHIFTS

There exists a wide range of audible acoustic ranging systems, but the fact that an audible signal is used makes the appeal of these techniques quite limited. In addition, although the efficacy of linear chirps has been successfully demonstrated for static point-to-point ranging, the capabilities of its simplistic spectrum has not been investigated thoroughly for improving Doppler tolerance in tracking devices maneuvering with high speed. In this section, we present the design of a tracking service that achieves better performance by combining the benefits of a near-inaudible acoustic broadband chirp and a dual-input receiver system. The key motivation to switch the sound signal from ultrasonic to audible domain was to make detection and tracking process faster as it will benefit by the need to process fewer samples.

We envision that the tracking system should locate humans and objects inside a building within a meter at any given time. We envision that sensor tags are attached to the targets, where each tag first broadcasts its identity and then starts the ranging process with a network of static anchor nodes that are installed inside the room of interest. The position of each tag is determined by the localization algorithm running on a centralized server. In the following sections, we discuss the challenges, solutions, and experiments conducted to evaluate our end-to-end tracking system.

### 5.1. Signal Design for Near-Inaudibility

Many signal sources sharing the same channel (or frequency spectrum) result in interference. The state of mutual coexistence of different signals is achieved by broadening the range of signal frequencies; thereby, alleviating the chances of the entire signal fading at a particular time, and decreasing the signal time-period to reduce the channel occupancy. However, audible-range acoustic signals add another constraint by operating in the frequency range that is easily perceived by the human auditory system. Therefore, achieving near-inaudibility is an important signal design feature for coexistence of acoustic signals in the audible domain.

The design of our proposed near-inaudible acoustic signal combines the principles of human psychoacoustics with signal engineering techniques. Psychoacoustics highlights the fact that the human ear shows a nonlinear frequency response between 20 Hz to 20 kHz, and records the highest sensitivity between [1-5] kHz (equal-loudness-contour)[Moore 2004]. The science of signal engineering suggests that the peak-to-average power ratio (PAPR) of a signal is a measure of its intensity. If the physical signal is acoustic, then a higher value of PAPR signifies loudness and vice-versa. Therefore, a simple solution to moderate (or reduce) the sound intensity is to broaden the frequency span of the signal that decreases the overall PAPR. Fig. 12 shows the relationship between PAPR and bandwidth  $B$  of a signal (which in our case is a linear chirp) with different time-periods  $T$  of 0.001 s and 0.01 s. We observe that PAPR decreases with the increase of bandwidth irrespective of the time-period of the signal. In addition, the signal with a smaller pulse duration (0.001 s) shows a lower PAPR value compared to the signal with a longer time-period (0.01 s). Therefore, designing a ranging signal with a very small pulse duration would result in a better near-inaudible



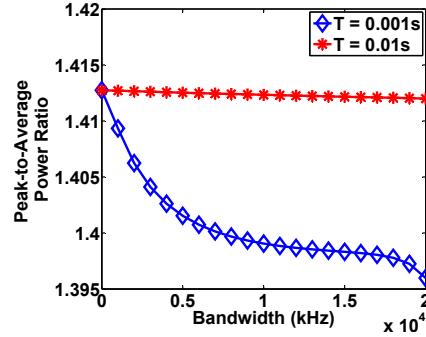


Fig. 12: Sound intensity (measured by peak-to-average power ratio) decreases with increasing bandwidth.

Table V: Evaluation of Inaudibility showing the percentage of humans that did not hear the signal.

Environment	Test-1	Test-2
	Uninformed (%)	Informed (%)
Env-1 (Quiet)	44.4	0
Env-2 (Noisy)	77.8	44.4

acoustic signal with significant reduction in sound intensity. However, a smaller  $T$  signifies lesser energy in the transmitted pulse, and hence, reduces the maximum range of the sensing system. This, in fact, introduces a trade-off between sound intensity (indicated by PAPR) and maximum range (dependent on  $T$ ), and their appropriate choice depends on the application requirements.

Based on these observations and our functional requirement of an operational range of 10 m, we use a linear chirp with the highest bandwidth in the audible-range [1 – 20]kHz and pulse duration of 0.01 s. In order to further limit the perceived sound intensity (for human hearing), the frequencies between [1-5] kHz were partially attenuated.

*5.1.1. Test of Near-Inaudibility.* To evaluate the near-inaudible signal design, we conducted sound perceptibility experiments<sup>6</sup> with 9 human test subjects in an office environment. Taking into account the different responses a human would have when subjected to informed and uninformed events, we formulate two specific test cases.

- **Test-1** - Uninformed: None of the subjects were advised about the periodic chirping signal inside the room.
- **Test-2** - Informed: All subjects were alerted a-priori.

These test cases were applied to the subjects in two different environments.

- **Env-1** - Quiet: Office setting wherein the subjects were not talking to each other for approximately 90% of the observation time.

<sup>6</sup>Human psychoacoustic analysis is necessary as humans can actually hear sounds in inaudible range due to the speaker response (for certain chirps).

- **Env-2** - Noisy: Office setting wherein the subjects were involved in discussions for approximately 90% of the observation time.

The total observation time for each set of experiment was 30 minutes. The transmitter was configured to chirp at 70 dB SPL.

Table V shows the results wherein the values represent the percentage of test subjects that did not perceive the presence of the chirping sound signal in the experimentation room. The results suggest that the near-inaudible signal design is most effective (more than 75% unperceived) when humans are not informed about the ambient chirping sound and they are involved in other simultaneous activities. However, the non-perceptibility rate decreases when humans are informed about their surrounding irrespective of whether their minds are engaged or relaxed. With respect to the tracking applications (human and object tracking inside buildings), the results are encouraging because the tracking service operators are not expected to reveal the tracking modality and the human subjects are expected to be performing their routine activities while they are being tracked ubiquitously.

## 5.2. Measuring Acoustic Doppler Shifts in Ranging

Doppler effect is a known phenomenon that is observed when objects exchange signals while moving relative to each other. The Doppler law states that if an object transmits a signal while being in motion relative to an observer, the frequency of the perceived signal will be Doppler shifted, and the magnitude of this shift depends on the frequency of the transmitted signal and the relative velocity between the object and the observer. Specifically, two effects will be observed in the received signal: first, the carrier frequency will be shifted proportional to the relative velocity; and second, pulse duration of consecutive received signal segments from the same range will be progressively time-shifted with respect to the time of transmission.

A criterion for validating the Doppler's effect on the linear frequency modulation of the linear chirp can be expressed in terms of its time-bandwidth product. If the signal duration is  $T$  and the target velocity is  $v$ , then the Doppler shifted received signal has a duration  $T(1 - v/c) = T - \Delta T$ . The Doppler's effect on the signal can be ignored only if [Kelly and Wishner 1965]:

$$\Delta T \ll \frac{1}{B} \quad \equiv \quad \frac{v}{c}TB \ll 1 \quad (19)$$

However, this change is noticeable if  $\Delta T$  is comparable to the inverse of the bandwidth  $B$ . At this threshold, the detection performance of the matched-filter degrades as the received signal does not stay matched with its reference copy. As a result, the output of the matched-filter does not show a distinct peak, but instead, a wider peak with its height significantly reduced and obscured by its sidelobes. This detection state introduces large errors in range estimation.

According to Eq. (19), the designed near-inaudible linear chirp signal of  $B=19$  kHz and  $T=0.01$  s will deteriorate in performance at an approximate source velocity of 3 m/s under Doppler's effect. This information is also conveyed in Fig. 13, where there is a significant increase in range estimation error and a corresponding decrease in the correlation peak when the velocity of the source exceeds 3 m/s. Fig. 13 also indicates that nonlinear chirps, following the quadratic or logarithmic laws, are more tolerant

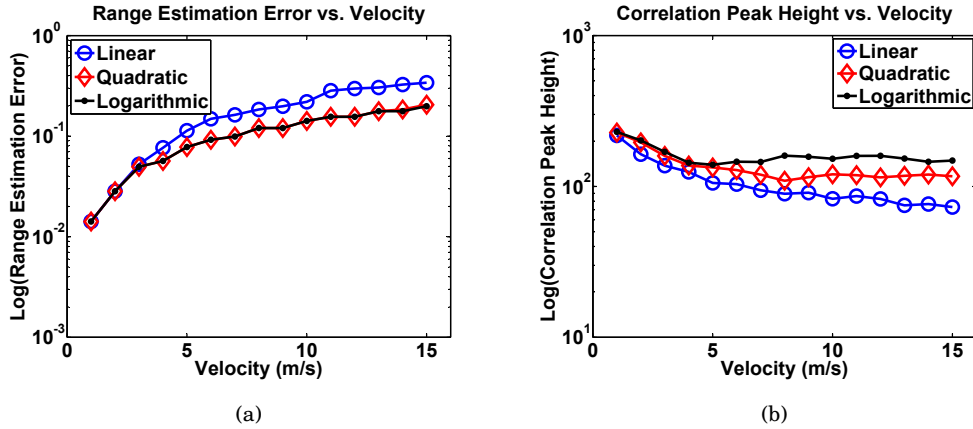


Fig. 13: Doppler's effect on linear and nonlinear chirps. For the designed near-inaudible linear chirp of  $B=19$  kHz and  $T=0.01$  s, there is a significant increase in range estimation error and a corresponding decrease in the correlation peak when the velocity of the source exceeds 3 m/s.

to ranging inaccuracies due to Doppler's effect<sup>7</sup>, and therefore, are better suited for tracking of very high velocity targets. However, with respect to typical indoor and semi-indoor applications where target velocity may not exceed 5 m/s, linear chirps are more befitting.

The Doppler shift of a signal changes the frequency and time-period of the transmitted pulse. If the signal is a linear chirp, then this change in frequency and time-period inevitably changes the sweep rate  $\mu$  of the pulse. We capitalize on the linearly sweeping characteristic feature of a linear chirp, and propose an algorithm to measure the Doppler shift.

**5.2.1. System Model.** The system model described in Section 4.2 is modified as follows to represent the operation of a tracking system. Fig. 14 shows a single transmitter node  $S$  moving with velocity  $\vec{v}$  in an area deployed with  $m$  receiver (i.e., anchor) nodes ( $R_i$  where  $i = 1$  to  $m$ ). We are interested in tracking  $S$ .

The magnitude of the Doppler shifts observed at node  $R_i$  depends on the relative speed of  $S$  and  $R_i$ , which can be found by projecting the velocity vector  $\vec{v}$  on the  $SR_i$  line. If the unit vector pointing from  $R_i$  to  $S$  is denoted by  $\vec{u}_i = \vec{R}_i S / \|\vec{R}_i S\|$ , then the relative speed of  $R_i$  and  $S$  can be defined as the following dot product.

$$v_i = \vec{u}_i \cdot \vec{v} \quad (20)$$

In Eq. (20),  $v_i$  is a scalar quantity with a positive value if  $\vec{v}$  points away from  $R_i$  (i.e., the transmitter is moving away from the receiver), and vice-versa. Moreover, the length of the projected vectors depends on both  $\vec{v}$  and the location of  $S$ , which means that vector  $\vec{v}$  can be same at different time instances, but the corresponding projected vectors will have different lengths.

<sup>7</sup>For a Doppler-invariant waveform, the instantaneous frequency of the signal at the time of transmission should remain unaltered at any instant of time. Nonlinear chirps preserve this property [Yang and Sarkar 2006], even under high velocity. Its explanation lies in the basic characteristic of the quadratic and logarithmic functions, which generate a nonlinear sweep (i.e., a curve) rather than a linear sweep (i.e., a straight line). This means that nonlinear chirps, inherently, have a long curve head of low frequencies (i.e., less Doppler shifted) followed by the short curve tail with high frequencies (i.e., more Doppler shifted). Hence, cumulatively, the impact of Doppler effect is less severe in nonlinear chirps.

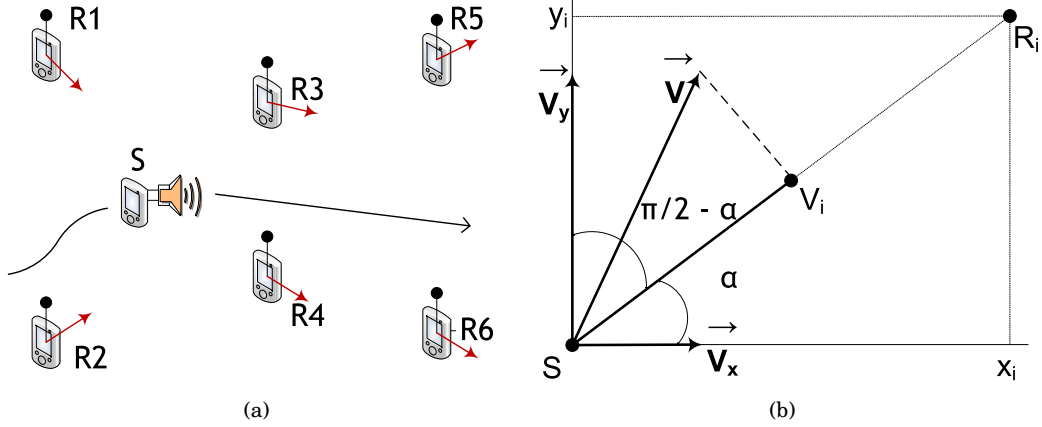


Fig. 14: The system model for tracking. (a): The receiver nodes  $\mathbf{R}_i$  are tracking the mobile target  $\mathbf{S}$ .  $\mathbf{R}_i$  calculates the relative speed of  $\mathbf{S}$  (red arrow) from the measured Doppler shifted signal.

(b):  $\mathbf{S}$  maneuvering with a velocity  $\vec{v}$  transmits a linear chirp signal.  $\mathbf{R}_i$  measures the Doppler shifted signal that depends on  $v_i$  (the relative speed of  $\mathbf{S}$  and  $\mathbf{R}_i$ ).

Let  $S$  transmit a linear frequency modulated chirp (Eq. (5)) with start frequency  $f_1$ , end frequency  $f_2$  (i.e., center frequency  $f_0 = (f_1 + f_2)/2$ ), and time-period  $T$ . Since  $S$  moves relative to a number of static nodes  $R_i$ , the frequencies in the received waveform  $r_i^d$  are Doppler shifted by a factor  $b_i$  at each node  $R_i$ .  $b_i$  is referred to as the expansion/compression factor, and depends on the relative speed  $v_i$  of  $S$  and  $R_i$ . If  $c$  is the speed of sound in the medium, then  $b_i$  is given as:

$$b_i = \left( \frac{c}{c + v_i} \right) \quad (21)$$

At each node  $R_i$ , the Doppler shifted start frequency  $f_{1i}^d$ , end frequency  $f_{2i}^d$  and time-period  $T_i^d$  of the received pulse (linear chirp) is given as:

$$f_{1i}^d = b_i f_1 \quad (22)$$

$$f_{2i}^d = b_i f_2 \quad (23)$$

$$T_i^d = \frac{T}{b_i} \quad (24)$$

Consequently, their Doppler effected sweep rates  $\mu_i^d$  are given as:

$$\mu_i^d = \frac{f_{2i}^d - f_{1i}^d}{T_i^d} \quad (25)$$

Substituting Eq. (22), Eq. (23) and Eq. (24) in Eq. (25), we obtain:

$$\mu_i^d = b_i^2 \mu \quad (26)$$

where,  $\mu$  is the sweep rate of the transmitted linear chirp.

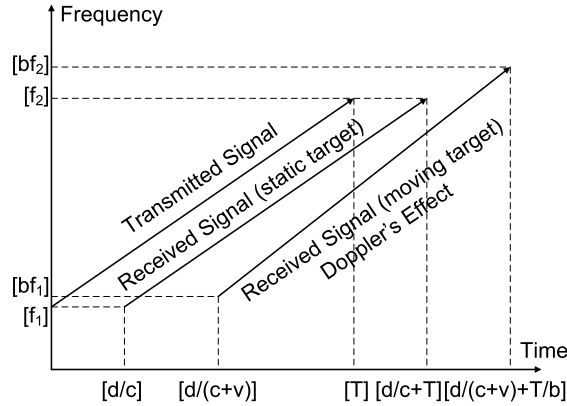


Fig. 15: Illustration of Doppler's effect on the properties (i.e., bandwidth and time-period) of a linear chirp. The instantaneous frequency and time-period of three signals are shown: the transmitted signal  $(f_2 - f_1)/T$ , the received signal  $(f_2 - f_1)/T$  from a static transmitter, and the Doppler shifted received signal  $(bf_2 - bf_1)/(T/b)$  from a mobile transmitter.

Using Eq. (21) and Eq. (26), we derive the following expression for calculating  $v_i$  (where the '+' sign for the square-root term is ignored as it will result in imaginary values<sup>8</sup>):

$$v_i = -c \left[ 1 - \sqrt{\frac{\mu}{\mu_i^d}} \right] \quad (27)$$

The only unknown parameter that needs to be computed for estimating  $v_i$  using Eq. (27) is  $\mu_i^d$ . In the following subsection, we propose an algorithm using the characteristic linear sweep feature of the linear chirp to measure  $\mu_i^d$ , and hence, the Doppler shift. A major advantage of our approach is that the speed and range information is derived implicitly and simultaneously from the same observation.

**5.2.2. Measurement Algorithm for  $\mu^d$ .** The measurement algorithm needs to be executed on every stationary receiver node associated with the tracking service, or at a higher level data fusion node that has the information of the received pulses at each individual receiver. For description simplicity, we explain the functionality at a single receiver (say  $R_1$ ). As a result of the Doppler shift, the sweep rates (or their respective slopes in the time-varying spectral representation as shown in Fig. 15) of the transmitted linear chirp and its received copy would be different. Therefore, the objective of this algorithm is to estimate the sweep rate of the received signal  $r_1^d$  (i.e.,  $\mu_1^d$ ) from the sweep rate of the reference (i.e., locally stored copy of the original transmitted) waveform  $\mu$  that can be computed from the known parameters  $f_1$ ,  $f_2$  and  $T$ .

The key idea is to progressively change the sweep rate of the reference signal by an increment factor  $\delta$  until it best fits with the sweep rate of the received signal  $r_1^d$ , thus providing an estimate of  $\mu_1^d$ . In a time-varying spectral representation, this process is equivalent to rotating the transmitted signal line in steps of  $\delta$  until its slope matches completely with the received signal line. However, the question here is: of the three parameters  $f_1$ ,  $f_2$  and  $T$  that control the sweep rate, which one should be chosen for controlling the sweep rate?

<sup>8</sup>The maximum value of  $\sqrt{\mu/\mu_i^d}$  in Eq. (27) can be 1. In this case, using the '+' sign for this square-root term will result in:  $v_i = -2c$ , which is not possible.

According to Doppler's theory, the magnitude of the Doppler shift observed at a given velocity is proportional to the frequency of the respective transmitted signal, i.e., the higher the frequency, the larger the observed frequency shift. Alternately, decreasing the frequency of the signal results in smaller Doppler shifts, but requires improved measurement accuracy. With respect to our designed linear chirp signal where  $f_1 = 1$  kHz,  $f_2 = 20$  kHz and  $T = 0.01$  s, the magnitude of the Doppler shift at  $f_1$  will be very small compared to  $f_2$  and will be difficult to precisely quantify. Therefore, we fix  $f_1$  and progressively change  $f_2$  in steps of  $\delta$  to estimate  $\mu_1^d$ .

The measurement algorithm is as follows, and requires iterating the steps 1-4 until the generated sweep rate matches with the Doppler shifted sweep rate.

- **Step 1:** Generate a new sweep rate (using the existing information in the transmitted pulse) by changing only  $f_2$  to  $f_2 + \delta$  (i.e., by an increment factor of  $\delta$ ).
- **Step 2:** Generate a new copy of the transmitted pulse using Step-1.
- **Step 3:** Using a matched-filter, cross-correlate the new copy of the transmitted pulse obtained in Step-2 with the received signal  $r_1^d$  and obtain the correlation peak.
- **Step 4:** If the magnitude of the correlation peak obtained in Step-3 (*which is an indication of the success/failure of the matching process*) is greater than the one observed in the previous iteration, then goto Step-1; else terminate the process by notifying that the best match for the Doppler shifted sweep rate has been derived. *An increase in the correlation peak height with every iteration signifies an additional scope for better matching, and so, the process is iterated. However, a decrease of the same is a trigger for termination as it indicates worsening of the matching process.*

For a respective transmitter-receiver pair  $(S, R_1)$ , the results of the measurement algorithm are interpreted as follows:

- **Result 1:** The magnitude of the tallest cross-correlation peak and its corresponding index among all iterations provides the *time-of-arrival* estimate (Fig. 16(a)), and thus, the *range*.
- **Result 2:** The sweep rate (or  $\mu_1^d$ ) at which the tallest cross-correlation peak was obtained (for Result-1) provides the *relative speed* estimate using Eq. 27.

The choice of the increment factor of  $\delta$  defines a trade-off between estimation time and measurement accuracy. A smaller  $\delta$  provides better range estimation accuracy but prolongs the processing time, and vice-versa for a larger  $\delta$  (Fig. 16). However, a smaller  $\delta$  increases the probability of false detection, wherein the estimator gets stuck in the local maxima and terminates the process without reaching its global maxima (Fig. 16(b)).

Fig. 17 shows the impact of  $\delta$  on the estimation accuracy of distance and speed. For accurate distance estimates, a smaller  $\delta$  is desirable (Fig. 17(a)) as it provides a fine-grained representation to approximate  $\mu$  by  $\mu_1^d$ . However, with respect to velocity estimation, a smaller  $\delta$  destabilizes the ratio  $(\mu/\mu_1^d)$  in Eq. (27), and increases the probability of frequent (positive and negative) oscillations about  $\mu$ . This effect can be observed by the large deviations at smaller  $\delta$  (Fig. 17(b)), which translates to an overall increase in error for the velocity estimate. Based on these observations, we choose the pivotal value of  $\delta = 50$  as the best trade-off between (distance,speed) accuracy and estimation time.

### 5.3. Tracking with dual-input receivers

In the design of the tracking system, we assume that a fixed number of anchor nodes are deployed with sufficient accuracy at known locations in the area of interest (i.e., their locations are known a-priori). A moving target node transmits a *linear chirp* signal with known sweep rate parameters. The received signal is measured by the

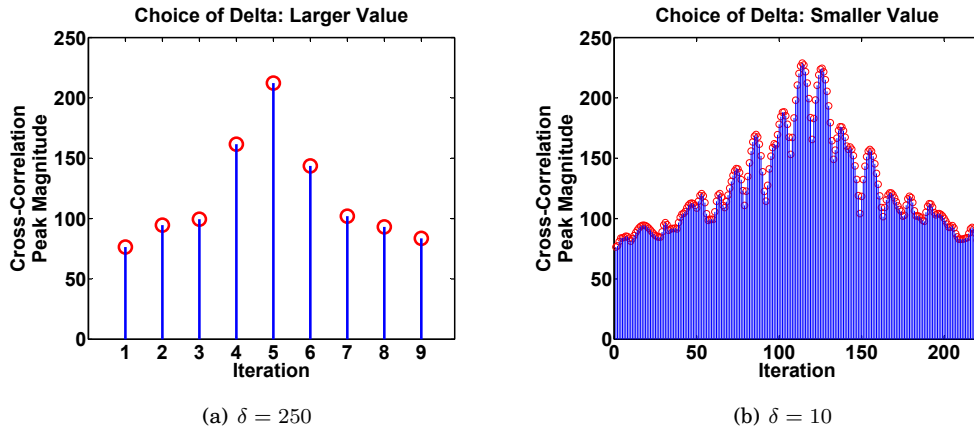


Fig. 16: (a): Illustrates the effect of progressive increase and subsequent decrease of the match percentage (indicated by the cross-correlation peak magnitude) of the rotated version of the transmitted signal with the received signal to compensate for Doppler's shift. (b): Illustrates the problem of increased false detection with smaller  $\delta$ . Observe that the estimator may wrongly indicate a local peak as the correct solution without executing further to reach the global peak.

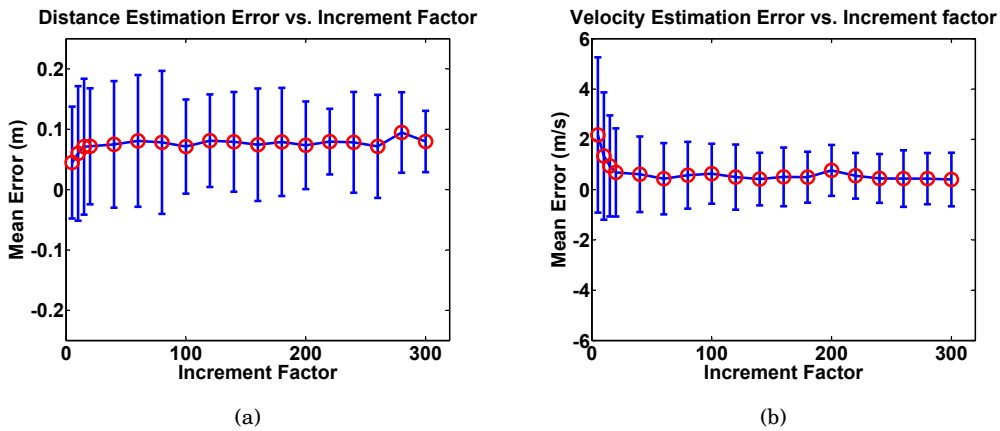


Fig. 17: Impact of increment factor  $\delta$  on (a): distance and (b): speed estimation error.

anchor nodes, and they assist the moving node to find its location and speed. The tracking service consists of three phases.

- **Coordination phase:** The tracked node sends a synchronous RF/acoustic pulse to synchronize its local clock with those of the anchor nodes. This is required for compensating the various timing errors in the ranging process as well to allow for data fusion of the ranging data.
- **Measurement Phase:** The location, bearing and/or speed of the tracked node are computed from range measurements, which are collected locally at each anchor node, and then, transferred to a data fusion node/workstation for trajectory estimation.

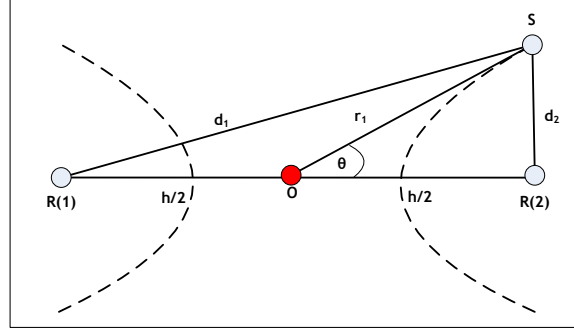


Fig. 18: Angle-of-Arrival derivation.

The tracked node initiates every measurement by synchronously broadcasting a radio packet followed by a (near-inaudible) linear chirp [1 – 20] kHz/0.01 s, while the multiple anchor nodes listen for incoming acoustic waves. The respective ranges are estimated by measuring the TDOA between the radio packet and the acoustic pulse, which is detected using a matched-filter and is identified by the first tallest peak.

- **Tracking Phase:** A Particle Filter (PF) is used to estimate the location and velocity of the tracked target. In each subsequent time step, the filter reduces the measurement (i.e., range, bearing and speed) error by combining these data measured at multiple anchor nodes to confidently predict the movement of the target.

We assume that a subset of infrastructure nodes will be equipped with more capable acoustic ranging hardware to estimate the angle-of-arrival (AOA) of the acoustic signal. Figure 18 demonstrates this scenario with a transmitter at  $S$  and two receivers at  $R(1)$  and  $R(2)$ . The signal from  $S$  takes two different paths to reach  $R(1)$  and  $R(2)$ , traversing path lengths of  $d_1$  and  $d_2$  respectively. The corresponding path length difference can be denoted as:

$$\Delta d = d_1 - d_2 \quad (28)$$

Considering the spacing between the two receivers as  $h$  and following the standard definition of a hyperbola, the locus of  $S$  can be defined as lying on the branch of a hyperbola with its foci at  $R(1)$  and  $R(2)$ , and eccentricity  $h/|\Delta d|$ . If  $r_1$  denotes the distance from  $S$  to the midpoint  $O$  of the two receivers, and considering the polar coordinates of  $S$  as  $(r_1, \theta)$  where the polar axis lies along  $R(1)R(2)$  with the pole at  $O$ , the polar coordinate equation of the hyperbola provides the AOA of a signal originating from  $S$  as:

$$\theta = \pm \arctan \left( \frac{\sqrt{(4r_1^2 - \Delta d^2)(h^2 - \Delta d^2)}}{\Delta d \sqrt{(4r_1^2 + h^2 - \Delta d^2)}} \right) \quad (29)$$

Usually,  $r_1$  is not available during measurements. However, if the transmitter is positioned sufficiently far away compared to the spacing between the receivers (in practice  $r_1 > 2h$ ), then AOA can be expressed using the polar coordinate equation of the asymptotes of the hyperbola as:

$$\theta = \pm \arctan \left( \frac{\sqrt{h^2 - \Delta d^2}}{\Delta d} \right) \quad (30)$$

where  $\tan \theta$  is the gradient of the asymptotes.



In the following section, we discuss a Particle Filter (PF) based tracking method to fuse the range and relative speed information to estimate the trajectory (i.e., location and velocity) of the moving target.

*5.3.1. Tracking Model.* Among the different available variants of the PF, we use a recursive Bayesian tracking algorithm that incorporates measurements from all sensors (in our case static anchor node referred to as receivers) into the estimation of the target (or transmitter) state. The PF is a suboptimal nonlinear filter that estimates the probability density function of the target state using sequential Monte Carlo methods. For the purpose of this work, the design of the filter has been limited to estimating the state of a *single target*.

The system state (or parameters) that need to be estimated are the location  $(x, y)$  and the velocity vector  $\vec{v} = (v_x, v_y)$  of the tracked object as it maneuvers with a constant velocity in the environment. At each discrete time step  $k$ , the tracked object estimates its distance from the static nodes. The current state  $\mathbf{x}_k$  of the system is calculated recursively from the previous state  $\mathbf{x}_{k-1}$  and a new observation vector  $\mathbf{c}_k$ . The parameter vector  $\mathbf{x}_k$  of the current state at time step  $k$  to be estimated is defined as:

$$\mathbf{x}_k = (x, y, v_x, v_y)^T$$

The system state  $\mathbf{x}_k$  evolves according to:

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{w} \quad (31)$$

where  $\mathbf{x}_{k-1}$  is the previous state,  $\mathbf{F}$  models the system dynamics, and  $\mathbf{w}$  is the process noise with covariance  $\mathbf{W}$ .  $\mathbf{w}$  represents the noise in our state model wherein the tracked node is expected to move with a constant velocity. The state transition matrix  $\mathbf{F}$  is given as:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & k & 0 \\ 0 & 1 & 0 & k \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (32)$$

Assuming that  $m$  static nodes measure the Doppler shifted acoustic signal, the observation (or measurement) vector  $\mathbf{c}$  is defined as:

$$\mathbf{c} = (\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_m)^T$$

where  $\mathbf{o}_i = (d_1, d_2, v_i, \theta)$  for  $i = 1, \dots, m$ .

$\mathbf{o}_i$  contains the distances  $d_1$  and  $d_2$  between the tracked object  $S$  and stationary node  $R_i$  to each of its two (left and right channel) receivers, relative speed  $v_i$ , and AOA  $\theta$ . Considering  $h$  to be the spacing between the two receivers, the calculated distance  $r$  is given as [Kottege 2009]:

$$r = \sqrt{\frac{d_1^2 + d_2^2}{2} - \left(\frac{h}{2}\right)^2} \quad (33)$$

Therefore, the measurement vector  $\mathbf{c}$  for each infrastructure node reduces to:

$$\mathbf{c} = (\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_m)^T$$

where  $\mathbf{o}_i = (r, v_i, \theta)$  for  $i = 1, \dots, m$ . The  $\pm$  sign of  $v_i$  is resolved in  $\mathbf{o}_i$  by taking  $\theta$  into account, where its increasing/decreasing value provides an indication regarding the movement of the target towards or away from the respective infrastructure node.

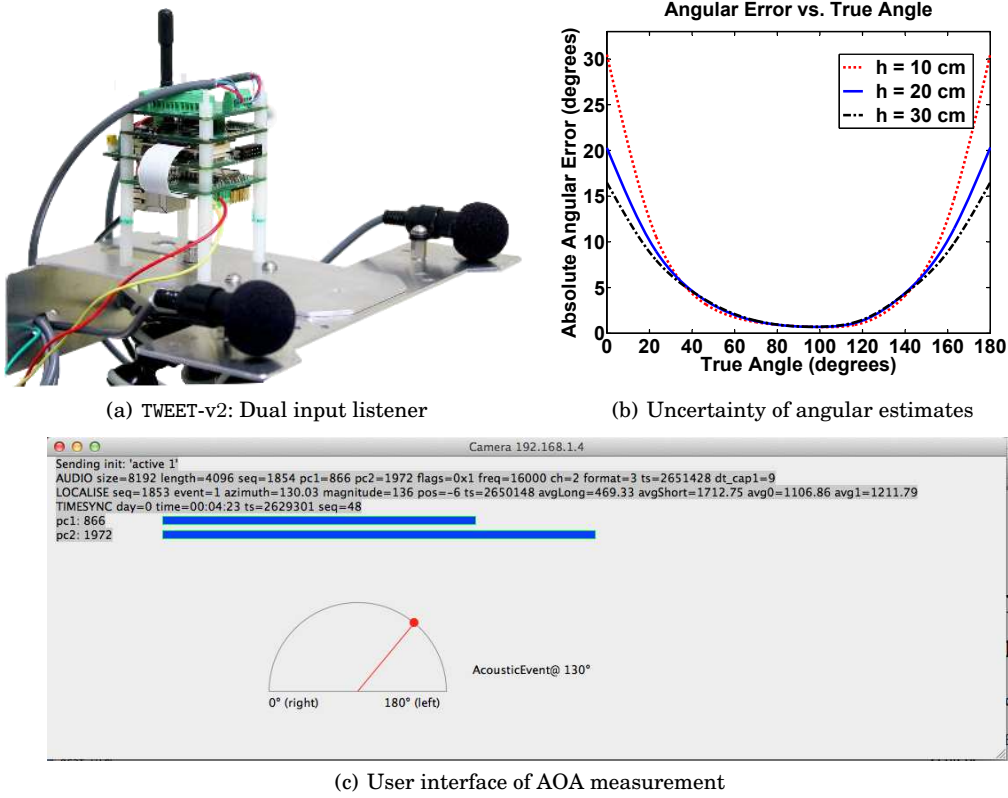


Fig. 19: The TWEET-v2 ranging system.

The observations  $\mathbf{c}_k$  of the current state is modeled as:

$$\mathbf{c}_k = H(\mathbf{x}_k) + \mathbf{u}_k \quad (34)$$

where,  $H$  is the nonlinear measurement function, and  $\mathbf{u}_k$  is the normally distributed measurement noise with covariance  $\mathbf{U}$ . The function  $H$  is a vector function consisting of  $m$  functions  $H_i : R^4 \rightarrow R^m$ , and is given as:

$$\begin{aligned} H_i(\mathbf{x}_k) &= \frac{v_x(x_i - x) + v_y(y_i - y)}{r_i} & i = 1, \dots, m \\ v_x &= v_i \cos \theta \\ v_y &= v_i \sin \theta \end{aligned} \quad (35)$$

The standard technique of Monte Carlo sampling is used to estimate the current state  $\mathbf{x}_k$ , and therefore, we refer our readers to [Ahmed et al. 2010] for information on the functional mechanism.

#### 5.4. Implementation with TWEET-v2

The tracking service was implemented using TWEET-v2: an improved version of the TWEET system. This newer version consisted of the same beacon mote, but featured a more capable listener mote with dual omnidirectional acoustic receivers for estimating the AOA (Fig. 19(a)).

The angular resolution depends on the separation distance  $h$  between the two receivers on the listener mote. It can be improved by increasing  $h$ , but will enlarge the receiver array size. Apparently, a more compact design can be obtained by reducing  $h$ , but at the cost of lower AOA measurement accuracy. This, in fact, introduces a design trade-off between angular resolution and form-factor. Fig. 19(b) characterizes the uncertainty of angular estimates for different values of  $h$  on the TWEET-v2 listener mote. The result indicates that, irrespective of  $h$ , highly inaccurate angle estimates appear in the range  $[0^\circ-20^\circ]$  and  $[160^\circ-180^\circ]$ ; but tend to improve in accuracy as the true AOA approaches  $90^\circ$ . Moreover, the best AOA estimates were observed for  $h = 30$  cm, while the worst were recorded for  $h = 10$  cm. Therefore,  $h$  was fixed at 20 cm so as to strike a good trade-off between AOA measurement accuracy and the dimensional compactness of the listener platform. Fig. 19(c) shows the respective UI screen shot captured for AOA measurements.

We conducted experiments in two different environments to calculate the accuracy of the location and velocity estimates of the tracked target.

- **Case-F** - Semi-indoor, Very low multipath: Roof covered canteen with the experimental area covering approximately  $[7 \times 7]$  m.
- **Case-G** - Indoor, Low multipath: A quiet lecture theatre ( $[12 \times 15 \times 10]$  m) with the experimental area covering approximately  $[10 \times 7]$  m.

We have utilized three infrastructure (anchor) nodes in each of the two experimental settings to track the mobile target node. A major challenge in evaluation was to measure the ground truth (both spatially and temporally) for the moving node with sufficient accuracy. We simplified this task by restraining the trajectory of the moving node to a series of straight line segments. A person carrying the transmitter node was maneuvering at approximately constant speed at each of the segments with a variable change in speed noticed while moving onto a different segment. The ground truth of the entire track was consolidated as the individual ground truths for each segment. The location and speed in a given line segment was estimated by recording the times when the moving node passed the endpoints of a given line segment. Its speed was calculated as the length of the respective line segment over its total time taken to span the segment, and its location was computed by interpolating the line segment with its corresponding measurement time. Therefore, for any given ranging measurement, we were able to reconstruct the ground truth location of the target as well as its velocity vector. The update interval was 250 ms, which was the time interval between subsequent chirps.

Fig. 20 shows the location and speed errors of the tracked node with respect to the ground truth. In both the experimentation environments, the mean location error varied between  $[0.22 - 0.50]$  m and the mean speed error varied between  $[0.45 - 1.20]$  m/s for different maneuvering speeds between  $[1.1 - 5.0]$  m/s. As observed from the results, the tracking accuracy degrades with increasing velocity, but is within a meter at any given time. We also ran two other tracking methods on the same data set: (i) simple multilateration using only the measured distances (TOA), and (ii) hybrid: combination of distance (TOA) and AOA estimates. The results show a significant improvement in tracking performance using PF over these two methods. Without compensating for Doppler effect, a 5 times increase in tracking error was recorded for all the three methods for target speed exceeding 3 m/s. Although, we have covered a relatively small area for tracking, our goal was to show that linear chirps can be tailored for tracking targets above their breaking point (which in our case was  $3 \text{ ms}^{-1}$ ) using the detection algorithm presented in Section 5.2.2. We show how the design of a linear chirp and its characteristic sweep feature can be used as an additional tool for curtailing measurement errors arising from Doppler's effect.

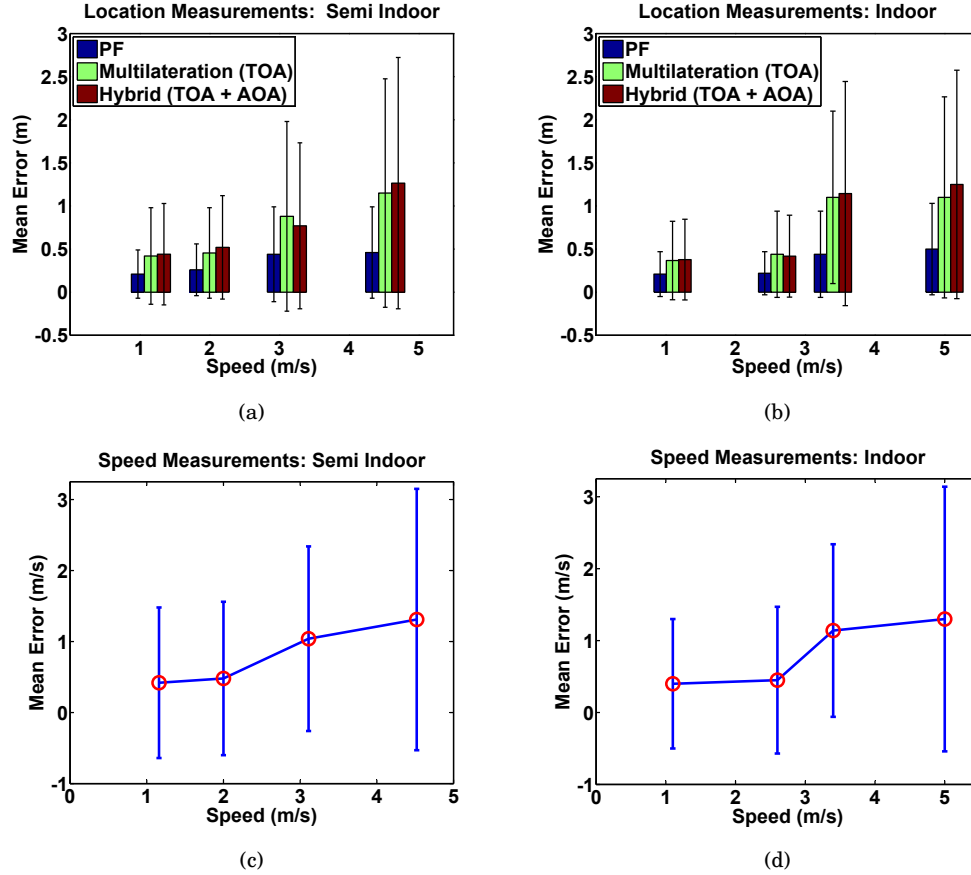


Fig. 20: Tracking performance for target speeds between [1 – 5] m/s using a linear chirp.

Relying on distances between nodes only, a simple setup with three anchor nodes requires three distance measurements to determine the location of the target in a two-dimensional plane. The additional information of the angle of the target movement is also able to solve the localization problem with two anchor nodes. However, measurements to additional anchor nodes with multiple ranging parameters improve the positioning accuracy. The near-inaudible capability is a useful addition to the tracking mechanism as it facilitates mutual coexistence with humans. For overcoming the inaccuracies from multipath effects in the Doppler shift measurement algorithm, we resort to the idea of rotating the transmitted signal (typically, a multipath free copy) rather a noisy trace of the received signal. This mechanism imparts robustness to the detection process and makes it less susceptible to multipath.

## 6. FUTURE WORK AND CONCLUSION

Although our work closely studies the two extreme ends of the acoustic spectrum (i.e., audible-range and ultrasonic) and two different signal design techniques (narrowband vs. broadband), there exist several directions for future study. While the performance of these ranging systems are impressive, they require hardware components that are costly and power exhaustive. This is a major drawback for the new fields of pervasive

Table VI: Comparison: M-Cricket vs. TWEET vs. TWEET-v2

Features	M-Cricket	TWEET	TWEET-v2
Ranging Signal	Narrowband (40 kHz/150 $\mu$ s)	Broadband ([20-25] kHz/50 ms)	Broadband ([1-20] kHz/10 ms)
Maximum Range	10 m	30 m	–
Operational Range	10 m	20 m	–
Ranging Accuracy	< 5 cm	< 4.5 cm	< 4.5 cm
Ranging Resolution	10 cm	5 mm	7 mm
Coverage	Omni	Omni	Omni
Obtrusiveness	3 Mic. Array Inaudible	1 Mic. Inaudible	2 Mic. Array Near-inaudible
Processing Hardware	MCU	Audio codec & DSP	Audio codec & DSP
Cost	$\approx$ \$400	$\approx$ \$1000	$\approx$ \$1000
Operational Environment	Indoor	Indoor/Outdoor	Indoor/Outdoor

computing and WSN that aim to achieve similar functional capability on low-cost and low-power hardware. Generally, resource constraints of WSNs limit the implementation of complex algorithms. Therefore, an interesting future direction is to design light-weight signal detection and post-processing algorithms that can be a reliable substitute to high-performance signal processing operations of high-dimensional data such as cross-correlation. Another direction is to revisit the conventional cross-correlation method (whose implementation has a running time of  $O(n^2)$  in the time domain and  $O(n \log n)$  in the frequency domain), and investigate into the design of a light-weight cross-correlation algorithm that has the same runtime complexity but a much smaller space complexity (competent with the mote constraints).

In this article, we studied the limited coverage problem of existing narrowband location sensing systems (such as the Cricket) due to the unidirectionality of available COTS (narrowband) ultrasound sensors. To overcome it, we designed and developed an omnidirectional ultrasonic receiver unit that was interfaced with the standard Cricket mote and showed improved coverage range results. With the understanding of the limited performance of narrowband systems, we transcended to broadband signal designs and studied the dependence of their bandwidth and time-period on resolution and range. Using a matched-filter (a benchmark signal processing method for TOA estimation), we also analyzed the uncertainty in estimating the correct correlation peak. We alleviated this problem by a simple proposition of estimating the envelope of the correlated pulse that make the role of correlation sidelobes irrelevant (which are a prime noise introduction factor in simple correlation). We combined these benefits onto a mote-based platform called the TWEET ranging system, and demonstrated the efficacy of ultrasound for long distance ranges (contrary to anecdotal records) in both indoor and outdoor environments. As the performance of the TWEET system was almost similar to standard audible-range systems (such as AENSBox) in ranging accuracy, and no additional benefits were obtained except for its (human) inaudible property, we presented the design an acoustic signal (in the audible range) with near-inaudible perception feature. Finally, we showed that linear chirps succumb to Doppler's effect under mobility. Hence, we proposed an additional algorithm for its mitigation when the velocity of the moving target exceeded the breaking point of linear chirps, and presented its working as part of a standard tracking service using the TWEET-v2 platform. A comparison among the three systems (i.e., M-Cricket, TWEET and TWEET-v2) proposed in this work has been tabulated in Table VI.

In the past decade, the sensor networks community has been successful in replicating the functionality of traditional location systems into mote scale devices. Now that

the small form factor systems exist, our next task is to make the existing detection algorithms work efficiently on these restricted platforms.

## 7. ACKNOWLEDGMENT

P. Misra and S. Jha acknowledge the partial support in this work from the Australian Research Council, under grant DP110104344. A part of this work was also carried out by P. Misra during the tenure of an ERCIM “Alain Bensoussan” Fellowship Programme funded by the *European Union Seventh Framework Programme (FP7/2007-2013)* under grant agreement n<sup>o</sup> 246016.

## REFERENCES

- N. Ahmed, M. Rutten, T. Bessell, S. S. Kanhere, N. Gordon, and S. Jha. 2010. Detection and Tracking Using Particle-Filter-Based Wireless Sensor Networks. *IEEE Transactions on Mobile Computing* 9, 9 (2010), 1332–1345.
- Joshua N. Ash and Randolph L. Moses. 2005. Acoustic time delay estimation and sensor network self-localization: Experimental results. *Acoustical Society of America Journal* 118, 2 (2005), 841–850.
- G. Borriello, A. Liu, T. Offer, C. Palistrant, and R. Sharp. 2005. 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)*. ACM, New York, NY, USA, 191–203.
- R. Boucher and J. Hassab. 1981. Analysis of discrete implementation of generalized cross correlator. *IEEE Transactions on Acoustics, Speech and Signal Processing* 29, 3 (1981), 609–611.
- J. Chen, J. Benesty, and Y. Huang. 2006a. Time delay estimation in room acoustic environments: an overview. *EURASIP Journal Applied Signal Processing* 2006 (Jan. 2006), 170–170.
- J. Chen, Z. Yu, and W. Xiao. 2006b. A Framework for Context-Aware Acoustic Source Localization. In *Proceedings of the 6th International Conference on ITS Telecommunications (ITSTS'06)*. 959–964.
- E. C. Cook and M. Bernfeld. 1967. *Radar Signals: An introduction to theory and application*. Academic Press.
- M.J. Crocker. 1998. *Handbook of Acoustics*. Wiley.
- L. Girod. 2000. Development and Characterization of an Acoustic Rangefinder. (2000).
- L. Girod and D. Estrin. 2001. Robust range estimation using acoustic and multimodal sensing. In *RSJ (IEEE/RSJ '01)*. IEEE, 1312–1320.
- L. Girod, M. Lukac, V. Trifa, and D. Estrin. 2006. The design and implementation of a self-calibrating distributed acoustic sensing platform. In *Proceedings of the 4th international conference on Embedded networked sensor systems (SenSys '06)*. ACM, New York, NY, USA, 71–84.
- A. Harter, A. Hopper, P. Steggle, A. Ward, and P. Webster. 1999. The anatomy of a context-aware application. In *Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking (MobiCom '99)*. ACM, New York, NY, USA, 59–68.
- M. Hazas and A. Hopper. 2006. Broadband ultrasonic location systems for improved indoor positioning. *IEEE Transactions on Mobile Computing* 5, 5 (2006), 536–547.
- Liu Hui, H. Darabi, P. Banerjee, and Liu Jing. 2007. Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews* 37, 6 (2007), 1067–1080.
- G. Jacovitti and G. Scarano. 1993. Discrete time techniques for time delay estimation. *IEEE Transactions on Signal Processing* 41, 2 (1993), 525–533.
- I. Jameson. 2006. *Time Delay Estimation*. Technical Report DSTO-TR-1705. Defence Science and Technology Organisation, Department of Defence, Australian Government.
- E. J. Kelly and R. P. Wishner. 1965. Matched-Filter Theory for High-Velocity, Accelerating Targets. *IEEE Transactions on Military Electronics* 9, 1 (1965), 56–69.
- J. R. Klauder, A. C Price, S. Darlington, and W. J. Albersheim. 1960. The Theory and Design of Chirp Radars. *Bell System Technical Journal* 39, 4 (1960).
- N. Kottege. 2009. *Underwater Acoustic Localisation in the context of Autonomous Submersibles*. Ph.D. Dissertation. Australian National University.
- N. Kottege and U. R. Zimmer. 2011. Underwater acoustic localization for small submersibles. *Journal of Field Robotics* 28, 1 (2011), 40–69.

- M. Kushwaha, K. Molnár, J. Sallai, P. Völgyesi, M. Maróti, and Á. Lédeczi. 2005. Sensor node localization using mobile acoustic beacons. In *Proceedings of the 2nd International Conference on Mobile Adhoc and Sensor Systems (MASS'05)*. IEEE, 483–491.
- Y. Kwon, K. Mechitov, S. Sundresh, W. Kim, and G. Agha. 2005. Resilient Localization for Sensor Networks in Outdoor Environments. In *Proceedings of the 25th IEEE International Conference on Distributed Computing Systems (ICDCS '05)*. IEEE Computer Society, Washington, DC, USA, 643–652.
- J. Minkoff. 2002. *Signal Processing Fundamentals and Applications for Communications and Sensing Systems*. Artech House.
- P. Misra, D. Ostry, and S. Jha. 2011a. Improving the Coverage Range of Ultrasound-based Localization Systems. In *Wireless Communications and Networking Conference (WCNC '11)*. IEEE, 605–610.
- P. Misra, D. Ostry, S. Jha, and N. Kottege. 2011b. *An Envelope-detection based Broadband Ultrasonic Ranging System for Improved Indoor/Outdoor Positioning*. Technical Report UNSW-CSE-TR-1106. The University of New South Wales.
- P. Misra, D. Ostry, N. Kottege, and S. Jha. 2011c. TWEET: an envelope detection based broadband ultrasonic ranging system. In *Proceedings of the 14th ACM international conference on Modeling, analysis and simulation of wireless and mobile systems (MSWiM '11)*. ACM, New York, NY, USA, 409–416.
- P. K. Misra. 2012. *Acoustical localization techniques in embedded wireless sensor networked devices*. Ph.D. Dissertation. University of New South Wales.
- R. Moddemeijer. 1991. On the determination of the position of extrema of sampled correlators. *IEEE Transactions on Signal Processing* 39, 1 (1991), 216–219.
- Brian C. J. Moore. 2004. *An Introduction to the Psychology of Hearing*.
- R. Nandakumar, K. K. Chintalapudi, and V. N. Padmanabhan. 2012. Centaur: locating devices in an office environment. In *Proceedings of the 18th annual international conference on Mobile computing and networking (Mobicom '12)*. ACM, New York, NY, USA, 281–292.
- C. Peng, G. Shen, Y. Zhang, Y. Li, and K. Tan. 2007. BeepBeep: a high accuracy acoustic ranging system using COTS mobile devices. In *Proceedings of the 5th international conference on Embedded networked sensor systems (SenSys '07)*. ACM, New York, NY, USA, 1–14.
- N. B. Priyantha. 2005. *The Cricket Indoor Location System*. Ph.D. Dissertation. Massachusetts Institute Of Technology.
- A. Savvides, C. Han, and M. B. Strivastava. 2001. Dynamic fine-grained localization in Ad-Hoc networks of sensors. In *Proceedings of the 7th annual international conference on Mobile computing and networking (MobiCom '01)*. ACM, New York, NY, USA, 166–179.
- A. Savvides, M. Srivastava, L. Girod, and D. Estrin. 2004. Wireless sensor networks. In *Localization in sensor networks*, C. S. Raghavendra, Krishna M. Sivalingam, and Taieb Znati (Eds.). Kluwer Academic Publishers, Norwell, MA, USA, 327–349.
- Sonic Studio 1999. Handbook of Acoustic Ecology. (1999). <http://www.sfu.ca/sonic-studio/handbook/Decibel.html>.
- M. Vorlander. 2001. Computer models in room acoustics - state of the art. *Nederlands Akoestisch Genootschap Journaal* 158 (2001), 39–50.
- E. Weinstein and A. Weiss. 1984. Fundamental limitations in passive time-delay estimation—Part II: Wide-band systems. *IEEE Transactions on Acoustics, Speech and Signal Processing* 32, 5 (1984), 1064–1078.
- A. Weiss and E. Weinstein. 1983. Fundamental limitations in passive time delay estimation—Part I: Narrow-band systems. *IEEE Transactions on Acoustics, Speech and Signal Processing* 31, 2 (1983), 472–486.
- M. Z. Win, A. Conti, S. Mazuelas, S. Yuan, W. M. Gifford, D. Dardari, and M. Chiani. 2007. Network localization and navigation via cooperation. *IEEE Communications Magazine* 49, 5 (2007), 56–62.
- B. Xu, R. Yu, G. Sun, and Z. Yang. 2011. Whistle: Synchronization-Free TDOA for Localization. In *Proceedings of the 2011 31st International Conference on Distributed Computing Systems (ICDCS '11)*. IEEE Computer Society, 760–769.
- J. Yang and T. K. Sarkar. 2006. Doppler-invariant property of hyperbolic frequency modulated waveforms. *Microwave and Optical Technology Letters* 48, 6 (2006), 1174–1179.
- J. Yang, S. Sidhom, G. Chandrasekaran, T. Vu, H. Liu, N. Cekan, Y. Chen, M. Gruteser, and R. P. Martin. 2011. Detecting driver phone use leveraging car speakers. In *Proceedings of the 17th annual international conference on Mobile computing and networking (MobiCom '11)*. ACM, New York, NY, USA, 97–108.
- J. Zhang, T. Yan, J. A. Stankovi, and S. H. Son. 2007. Thunder: towards practical, zero cost acoustic localization for outdoor wireless sensor networks. *SIGMOBILE Mobile Computing Communications Review* 11, 1 (2007), 15–28.

G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic. 2004. Impact of radio irregularity on wireless sensor networks. In *International conference on Mobile systems, applications, and services (MobiSys) (MobiSys '04)*. ACM, New York, NY, USA, 125–138.

Received August 2012; revised December 2012; accepted March 2013