SwingLoc: Acoustic Indoor Localization Leveraging Doppler Effects via Wearable Computing

Weiling Zheng
RMIT University, Australia
zwlcle@hotmail.com

ABSTRACT

Indoor localization using mobile sensing platforms has become a ubiquitous service that enables various smart building and healthcare related applications. As wearable device becomes an important player in the mobile market, an indoor localization system tailored specifically for it remains absent. In this paper, we present SwingLoc, an indoor positioning system aimed particularly for hand-wear devices. It takes use of the natural arm swinging when the user is walking, and the Doppler effects it triggers while receiving acoustic signals to locate the user. With the need of off-the-shelf speakers, which are already present in most public indoor areas, SwingLoc monitors the wearable device's direction toward the speakers in consecutive gait cycles, and solve a nonlinear least squares problem for the user's position. Our real-world tests involving 6 users at two different locations shows that SwingLoc can achieve overall 85% localization errors under $2m$, with extreme conditions where at most 3 speakers are present in the environment. Our experiments demonstrate SwingLoc to be robust and effective, and has great potential for providing fine-grained location-based services and improve human well-being.

CCS CONCEPTS

• **Human-centered computing → Smart sensing**; **Human behavior analysis**; **Indoor localization**.

ACM Reference Format:
Weiling Zheng. 2022. SwingLoc: Acoustic Indoor Localization Leveraging Doppler Effects via Wearable Computing. In *17th ACM Workshop on Mobility in the Evolving Internet Architecture (MobiArch'22), October 21, 2022, Sydney, NSW, Australia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3556548.3559634

1 INTRODUCTION

Recently, acoustic-based localization techniques have attracted many research efforts. One big advantage of these systems is that they can be directly deployed in buildings or outdoor plazas/markets which have audio infrastructures already in place [4]. In addition, the audio signals can be received by the mobile device with microphones, thus it is accessible to normal smartphone users. Prior approaches have explored angle of arrival (AOA) [7] and time ranging like time-of-arrival (TOA) and time-difference-of-arrival (TDOA) [9, 11] methods to deliver sub-meter location accuracy. Generally, they employ anchoring speakers to emit non-audible acoustic signals (around 20kHz) and use smartphones as receivers. The time-ranging based systems are less practical in large public arenas like airports or shopping malls, since they do not have enough channels to synchronize all of the transmitters precisely. A recent AOA based work, Swadloon [7] makes use of Doppler effects by phone shaking [15] to estimate the direction of the acoustic sources. Though the system achieves promising results in both direction finding and localization, there are several constraints in the real-world implementation. For instance, the user has to shake the phone gently and movements need to be done in a rough horizontal plane; in real-time tracking, the performance relies on number of nodes (speakers) and the initial location estimation accuracy, which would prolong the localization process and inevitably limit the system feasibility. Ideally, we expect a simple, widely accepted solution which can provide the desired accuracy without extra movements or deployments.

Moving along this direction, in this paper we present an indoor ultrasonic positioning system, SwingLoc, which leverages wearable devices to accurately estimate the direction from the target user to acoustic sources. Our solution is derived from a key insight: "arm movements during walking will cause Doppler effects on audio signals received by the wrist-worn smartwatch." Specifically, SwingLoc tracks natural user motion during (i.e., gait and swing) walking by using inertial sensors on smartwatches while at the same time leveraging its embedded microphones to record corresponding Doppler shifts on received ultrasonic signals. When the relative velocities (phone to speaker) and the frequency shifts are obtained, we are able to calculate the directions to the acoustic sources and estimate the indoor/outdoor location of the user via the AOA scheme.

While the idea of SwingLoc seems simple, many challenges arise in the practical implementation. Firstly, unlike previous localization scenarios [7], the sensory data collected from the smartwatch contains not only arm swing gestures but also gait movements, which could result in unpredictable patterns with noisy data. It is difficult to identify the motion cycles (e.g., start or end points) and estimate the wrist velocity. Secondly, the phase and frequency shift of the acoustic signal may fluctuate continuously due to the rapid change of the wrist motion (e.g., direction and speed). Since the smartwatch has limited processing power, the system has to face trade-offs associated with levels of accuracy, the cost of computation and the system delay. Thirdly, since a user may perform arbitrary motion during walking, the system should be able to distinguish swing gestures from other non-related hand movements.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
MobiArch'22, October 21, 2022, Sydney, NSW, Australia
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9518-2/22/10…$15.00
https://doi.org/10.1145/3556548.3559634

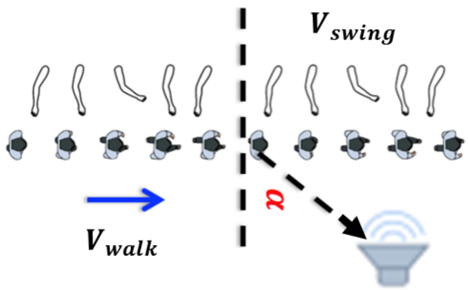

Figure 1: Normal walking with free arm swing.

To address above issues, we first utilize inertial sensors on smartwatch to distinguish user's different locomotion states such as walking, turning, as well as going upstairs/downstairs. Such locomotion information could aid to remove gait noise and infer direction changes during walking. A gesture spotting method is then deployed to track the posture of user's arm during walking and locate the key points of the swing motion (i.e., start, perpendicular to the ground, end). In addition, we use the linear regression method [7] to remove the error shifts, a step-wise model is presented for direction finding and localization. To balance the computation resources, we utilize the transmitting scheme on the smartwatch and run heavy tasks (e.g., fast Fourier transform (FFT)) on the smartphone. We refer to Section 4 for the detailed system design.

Specifically, we make the following contributions:

(1) We proposed a ultrasonic approach, SwingLoc, which leverages on off-the-shelf wrist-worn devices (e.g., smartwatch) for precise indoor localization. In order to estimate the directions to acoustic sources, the system utilizes embedded sensors (i.e., inertial sensors and microphones) on the device to capture Doppler effects on acoustic signals triggered by natural human motion (i.e., arm swing).
(2) We proposed the arm posture tracking method by analyzing the user's locomotion states to detect wrist positions at different time points during consecutive gait cycles. With the locomotion data and AOAs derived from frequency shifts, we are able to locate users by solving a nonlinear least squares problem based on our step-wise model.
(3) We conduct a set of real world experiments, involving 6 participants (3 females and 3 males) with different body sizes in 2 different indoor environments (an indoor parking garage and a lobby in the engineering building) to evaluate our system and empirically study the impact of various factors on the accuracy of acoustic-based indoor localization. Experimental results show that SwingLoc is robust to real world scenarios, and has great potential to provide desirable indoor positioning services.

The reminder of the paper is organized as follows. Section 2 discusses related work and how our approach differs. Section 3 gives a brief view of the design considerations of implementing the SwingLoc. Section 4 presents the SwingLoc architecture and detailed system design. Section 5 presents real world experiments and evaluation results. Section 6 concludes the paper and discusses the future work to improve the system performance.

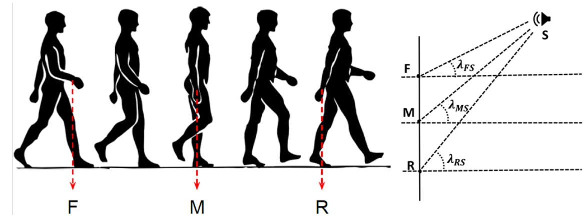

Figure 2: Three special wrist positions and their AOAs used by SwingLoc.

2 RELATED WORK

There has been extensive research work addressing indoor localization/positioning problems. In this section, we summarize most related ones to our proposed system. For more comprehensive surveys, we refer to [1] and [22].

2.1 Wi-Fi Based Approaches

Due to the wide deployment of wireless network infrastructures, RSSI based systems have been widely deployed and studied such as Radar [2] and Horus [23]. In addition to the coarse RSSI estimation, later work has leveraged the PHY layer WiFi channel state information (CSI) to achieve better localization precision [5, 16]. However, those systems usually become less accurate due to subtle changes in the environment and require a large amount of calibration effort [12, 22]. Though auto-calibration [18] and crowd-sourcing methods could help [14, 21], extra work is needed to guarantee the quality of data.

With the development of multi-antenna design to support MIMO communications, indoor localization using signal processing techniques have gained interest recently. A number of time ranging based work break the barrier of meter-level accuracy using channel hopping mechanisms [6, 19, 20]. The basic idea of these systems is to compute distances to *APs* (at least 3) via CSI and leverage geometrical models for location estimation [17]. The most recent work, SpotFi [8] can achieve decimeter-level accuracy by incorporating an advanced MUSIC-AoA algorithm to eliminate multipath components. However, those systems still have disadvantages such as communication disruption due to channel hopping, the performance significantly deteriorates when SNR decreases and slow adaptation on moving mobile targets (may need movements through multiple locations).

2.2 Acoustic Based Approaches

Recently, acoustic-based approaches have emerged, which could provide accurate indoor localization without special-purpose infrastructure [3]. For example, Lazik et al. [9] presents an indoor tracking system using modulated ultrasonic chirps on off-the-shelf audio speakers. It achieves sub-meter accuracy (1*m* at 95%) by applying a Time-Difference-of-Arrival (TDOA) pseudo-ranging technique. Nevertheless, these systems has to synchronize all of the transmitters/nodes precisely or require sufficient anchor network coverage, which is not possible in commodity deployments. Though a few past work combines WiFi measurements and acoustic ranging to improve indoor localization precision [10, 13].

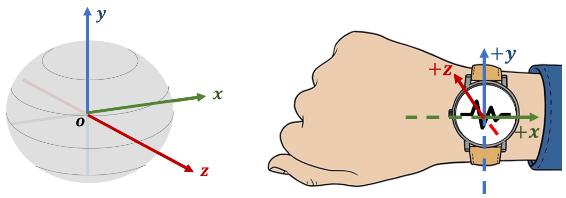

Figure 3: The earth coordinate system and smartwatch coordinate system

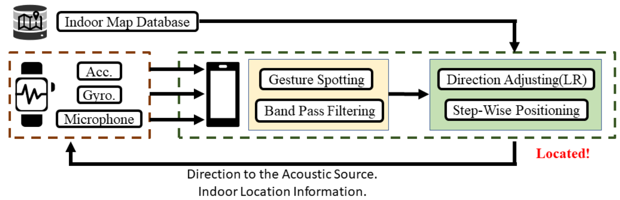

Figure 4: The architecture of SwingLoc.

Towards the most related work in acoustic-based indoor localization, Swadloon [7] utilizes multiple sensors (i.e., inertial sensors and microphone) on smartphones to measure signal Doppler effects caused by the user motion (shake the phone). By implementing a basic AoA based geometrical model, it achieves 90% accuracy of $0.92m$. However, the system is sensitive to the behavior of each individual such as estimation errors due to different shaking patterns and extra shaking gestures if a user fails to move the hand gently, which would make users feel exhausting and inevitably prolong the localization process. Correspondingly, our system tracks the displacement of the smartwatch (on hand) when a user swings the arm normally during natural human gait (walk), which provides a more effective solution to the problem.

3 DESIGN CONSIDERATIONS

The essential idea of SwingLoc is to measure the Doppler effects on ultrasonic signals that triggered by arm movements during walking. In this paper, we assume that the user will wear the smartwatch on one hand and swing the arms naturally. We use static acoustic sources to emit a continuous audio tone at a known frequency f_o,

When a user starts to walk and swing the arm, the value of frequency shift f_d can be calculated as:

$$f_d = f_r - f_o = \frac{v_r}{c} f_o = \frac{v_{wrist} \sin \alpha}{c} f_o, \tag{1}$$

where $f_r = \left(\frac{c+v_r}{c+v_s}\right) f_o$, here c, v_r, and v_s are the velocity of waves in the medium, of the receiver (smartwatch), and of the sender, respectively. In our scenarios, the smartwatches and the acoustic sources are considered at the same height (roughly) and the arm swing direction on the horizontal plane (parallel to the ground) can be approximated as the user moving direction as shown in Figure 1.

In addition, the direction α contains adjusting errors due to the height changing of the wrist when it swings roughly follows a circle. Thus we only select a period of the motion to ensure that the displacement of the wrist is relative small compare to the distance from user to the acoustic source.

As we will capture the frequency shifts and calculate the wrist velocity during walking, it is important to select proper sampling periods when a user swings. As shown in Figure 2, in our design the wrist positions corresponding to the arm swinging to the front-most / right beside the body trunk / rearmost as F, M and R. According to our observation, the replacements of these three positions are not significant (less than $30cm$) if users swing naturally. We consider the AOAs at these three positions to the audio source are roughly equal.

To estimate the velocity of the wrist, we have

$$\frac{v_{wrist_M} \frac{\sin \lambda_{MS}}{c} f_o}{f_M} = \frac{v_{wrist_F} \frac{\sin \lambda_{FS}}{c} f_o}{f_F}, \tag{2}$$

where the pair of v_{wrist_M} and v_{wrist_F} could be replace by v_{wrist_M} and v_{wrist_R}, or v_{wrist_R} and v_{wrist_F}. We only use v_{wrist_M} and v_{wrist_F} as an example. Since $v_{wrist} = v_{watch} + v_{body}$, where $v_{wrist_M} = v_{watch_M} + v_{body_M}$ and $v_{wrist_F} = v_{watch_F} + v_{body_F}$.

Therefore, how to identify F, M and R during the swing motion is a very important task. In our design, we utilize gesture spotting algorithms to detect user's arm postures at different points (as discussed in section 4.1.1). Then leverage the velocity information to facilitate direction finding and indoor localization.

4 SYSTEM DESIGN

In order to provide precise indoor positioning, the smartwatch gathers samples from the microphone and inertial sensors when the user swings the arm during walking. In this section, we present a detailed discussion of the SwingLoc architecture and algorithms. The primary goal of our proposed system is to examine the Doppler Shifts on acoustic signals (sensed by the smartwatch microphone during a swing) at different positions under indoor environments. The key components of SwingLoc is shown in Figure 4.

4.1 Walking Step Spotting and Data Filtering

4.1.1 Swing gesture spotting. As illustrated in Figure 2, we normally swing our arms follows a rough circle during walking. On the other hand, it has been proved that the angular velocity measured by smartwatch gyroscope could achieve reasonable accuracy. Thus we can leverage the rotation movements of the to generate soft hints for spotting the swing gesture.

In this paper, the measurement and calculation are conducted in a standard 3-axis global (Earth's) coordinate system as illustrated in Figure 3. The smartwatch sensor readings are aligned with the absolute coordinate system by using a rotation matrix R (provided by Android library), Z points towards the sky and is perpendicular to the ground, Y is tangential to the ground at the device's current location and points towards magnetic north and X is defined as the vector product Y· Z, as illustrated in Figure 3.

The first step for walking direction estimation is to determine whether a person is walking or not. As illustrated in Figure 6, the maximum angular speed (Z-axis) norm peak occurs almost periodically during walking. Based on this observation, we determine a person is walking if the peak times occur periodically. In addition, we can see the clear patterns of wrist angular speed (Z-axis) along

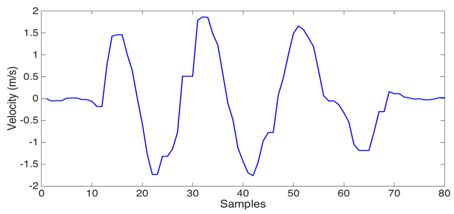

Figure 5: An example of acceleration integration for swing velocity.

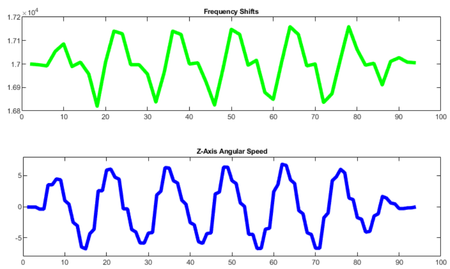

Figure 6: Doppler shifts along with the human swing gestures

with the frequency shifts on received signals. When the velocity as well as the angular speed reach the maximum value during the swing gesture, we can get the largest value of frequency shifts. Thus we use gesture spotting method to segment the time period of swing motion and then estimate the AOAs from the user to the audio resources.

We adopt a rule/experience-based solution that filters out the acceleration input if it does not belong to the task domain (e.g., anomalies, acceleration caused by turning) and only select a period of the swing motion for moving velocity calculation. And then we draw the idea from [7], equiripple FIR filter is chosen in SwingLoc and implemented on smartphone.

4.2 Derive the directions with Linear Regression

We apply the linear regression method in [7] to estimate the direction from the user's watch to the speaker and the direction of walking.

Within a short time period, the displacement of the watch is far less than its distance to the speaker. Thus the watch's direction to the speaker can be considered as a constant λ within this period. In the meantime, the accumulated error of integrating acceleration for velocity will be largely reduced if the interested time period is short. Bearing the two aspects in mind, we can build up the following group of equations and solve for λ with linear regression.

First, we assume the instrumental noise of the accelerometer is a constant within the interested short period and is denoted by (e_x, e_y). Let $(a_x[i], a_y[i])$ be the measured acceleration at time instant T_i, then the instantaneous velocity of the watch can be

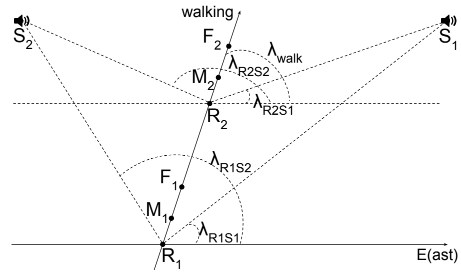

Figure 7: Relation between the speaker and the wrist positions in two consecutive gait cycles.

computed as,

$$\begin{cases} v_x[k] = v_x[0] + \sum_{i=0}^{k-1} a_x[i]T[i] + e_x \sum_{i=0}^{k-1} T[i] \\ v_y[k] = v_y[0] + \sum_{i=0}^{k-1} a_y[i]T[i] + e_y \sum_{i=0}^{k-1} T[i] \end{cases} \tag{3}$$

Here, $T[i] = T_{i+1} - T_i$, denotes the time interval between timestamps T_{i+1} and T_i.

From the Doppler's effect, we have,

$$\lambda_x v_x[k] + \lambda_x v_x[k] = \frac{v}{f} f[k], \tag{4}$$

where, v is the traveling speed of sound in the air, f is the frequency of the source ultrasound, $f[k]$ is the watch's received frequency at time instant T_k.

Finally, we can build the following equation system,

$$\begin{bmatrix} v_x[0] & v_y[0] & 1 & t[0] \\ v_x[1] & v_y[1] & 1 & t[1] \\ \vdots & \vdots & \vdots & \vdots \\ v_x[n] & v_y[n] & 1 & t[n] \end{bmatrix} \begin{bmatrix} \lambda_x \\ \lambda_y \\ \lambda_0 \\ \lambda_1 \end{bmatrix} = \frac{v}{f} \begin{bmatrix} f[0] \\ f[1] \\ \vdots \\ f[n] \end{bmatrix}, \tag{5}$$

Here, (λ_x, λ_y) is the unit direction vector λ pointing from the watch to the speaker; $\lambda_0 = \lambda_x v_x[0] + \lambda_y v_y[0]$; $\lambda_1 = \lambda_x e_x + \lambda_y e_y$.

4.3 Step-wise localization

Human walking expresses a highly repetitive pattern, which has been widely employed for pedometers and fitness tracking on mobile phones. A human gait cycle consists of two steps, one by each leg. In an ideal case, the arm swinging in each gait cycle draws the same arc in the air. Therefore, if we take a pair of wrist positions each from the same stage of two consecutive gait cycles, the vector that connects the two positions actually represents the displacement of the human body. Since we denote the wrist positions corresponding to the arm swinging to the front most / right beside the body trunk / rearmost as F, M and R respectively (as shown in Fig. 2), then we'll have:

$$\overrightarrow{F_1F_2} = \overrightarrow{M_1M_2} = \overrightarrow{R_1R_2}, \tag{6}$$

where F_1, M_1, R_1 are wrist positions of the first gait cycle, and F_2, M_2, R_2 of the second.

As discussed in the previous subsection, we can solve the directions of the points F, M, R towards the sound source, as well as the user's walking direction. Based on them, we can easily compute

Figure 8: Indoor environments for evaluation

the angles related to the east direction, i.e. $\lambda_{R_1S_1}$, $\lambda_{R_1S_2}$, $\lambda_{R_2S_1}$, $\lambda_{R_2S_2}$, λ_{walk} in Fig. 7. Then we can have,

$$\angle S_1R_1S_2 = |\lambda_{R_1S_2} - \lambda_{R_1S_1}|,$$

Angles for points M and F can be computed similarly.

In an indoor environment, the compass on mobile devices suffers from interference of electromagnetic fields from all kinds of electronic products, and the world coordinate system reported by the compass is actually an rotated version of the real one. However, we can assume that within a small space, e.g. a radius of two steps, the rotation from the reported world coordinate system to the real one is a constant, and will be cancelled out when we compute angles like $\angle S_1R_1S_2$.

Then, within each gait cycle i, we can have the following equations using the cosine rule,

$$\begin{cases} |S_1S_2|^2 = |R_iS_1|^2 + |R_iS_2|^2 - 2|R_iS_1||R_iS_2|\cos\angle S_1R_iS_2 \\ |S_1S_2|^2 = |M_iS_1|^2 + |M_iS_2|^2 - 2|M_iS_1||M_iS_2|\cos\angle S_1M_iS_2 \\ |S_1S_2|^2 = |F_iS_1|^2 + |F_iS_2|^2 - 2|F_iS_1||F_iS_2|\cos\angle S_1F_iS_2 \end{cases}. \quad (7)$$

In conclusion, for each two consecutive gait cycles, we have 12 unknowns, i.e. the coordinates (F_{1x}, F_{1y}), (M_{1x}, M_{1y}), (R_{1x}, R_{1y}), (F_{2x}, F_{2y}), (M_{2x}, M_{2y}), and (R_{2x}, R_{2y}); we'll have $3*2 = 6$ equations from Eq. 7, and $2*2 = 4$ equations from Eq. 6, which are 22 equations in total that can be solved as a nonlinear least-squares problem. By solving this problem, we can get the position of M_1 and M_2, which are used as the estimation of the user's locations.

In this paper, we do not consider 2-speaker case (discussed above) in the evaluation, for a 3-speaker case, an AoA based method like [7] could be employed to solve the problem.

5 EVALUATION

In this section, we conduct experiments for the proposed ultrasonic positioning system, SwingLoc, in real world environments. Specifically, we first describe the experiment setup, then we examined the performance of the real time positioning method, we also analyzed system latency in the experiments.

5.1 Experimental Setup

5.1.1 Data Collection. We implemented SwingLoc on a Samsung Gear Live smartwatch with quad-core 1.2 GHz Cortex-A7 CPU, 512 MB RAM and a Google Nexus 6 Android smartphone which is equipped with quad-core 2.7 GHz Krait 450 CPU, 3G RAM. As we discussed in section 3, the smartwatch is used for wrist motion sensing and audio signal recording, the audio sample rate is 44100Hz, and sample rate of the inertial sensors is 20Hz. All other components, including gesture spotting, BPF and localization algorithm were running on the smartphone.

Table 1: Angular Error Garage

Angular Error (degree)	<5	<10	<15	other
10m - 20m	126	155	167	13
20m - 30m	137	156	164	16

As illustrated in Figure 8, we choose a lobby and an parking garage for real world indoor localization. More specifically, at each place we placed 3 speakers generating sinusoidal signals at the frequency from $17900Hz$ to $20500Hz$ and manually labeled 20 spots (10 each) as positioning ground-truth. We will only use three speaker at the same time in the localization to test our system performance under extreme conditions (now more acoustic source available).

We implement our experiments in two Non-line of sight (NLOS) scenarios, 6 participants (3 females and 3 males) walked through these spots 5 times at normal speed. We assumed that that all participants do swing the arm during walking. During the evaluation, participants were asked to firstly wear the smartwatch on either side of their hands.

5.2 Indoor Direction Finding

As we discussed in Sec. 4.2, the device's direction towards the speaker can be computed using linear regression, under the assumption that the device's displacement in a short period is much smaller than its distance to the speaker. To evaluate the performance of direction finding, we carried out experiments in the indoor garage with the user being 10m - 20m and 20m - 30m away from the speaker, 3 spots are manually selected for each range, respectively. Each participant will repeat the test 5 times, thus we have 180 tests in total. Since the lobby is narrow for direction finding, we only test position tracking there. The estimation errors are shown in Table 1, at both distances from the speaker, we can see the system could narrow the error to 5 degrees most of the time in both two cases. And we find that the direction finding accuracy is slightly worse if the audio source (speaker) is closer to the receiver (smartwatch). As the user approaches the source, the assumption eventually becomes too strong, and the corresponding angular error increases.

5.3 Non-line of Sight Cases

We presented two NLOS cases in the evaluation, the first one is the study of moving human body effects on the acoustic signals. We let one person stand between the user and the speaker. The other one, we evaluate how user body affects the system performance. Since when we swing the arm, the line between the smartwatch and the acoustic source may be blocked by body trunk, back or legs. Though it happens just in a short period, the events frequency is large (a user may keep swinging the arm).

In the test, we verify the amplitude of the ultrasonic signal. In case 1, when a moving person is close to the user (around $2m$), the signal attenuated significantly. It happens a gain when the person moving to the speaker. In case 2, though the period is very shot, we still can catch the weak signal one or two times per swing. Thus, we should improve the motion spotting scheme to overcome this limitation. For other conditions like walls or on-road objects, we leave the study as the future work. To handle above issues, we may use mobile crowd-sensing techniques to get the map information

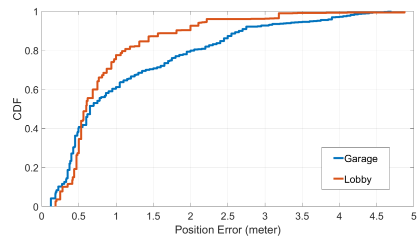

Figure 9: CDF of Positioning Errors

from other users at the same place, which helps build a more accurate model and estimate real-time distance. The detailed plan is discussed in Section 6.

5.4 Step-Wise Localization

For SwingLoc, each time only three speakers are needed to locate the user. In reality, there are usually more than three speakers in the ambient environment that the user's device can hear from, thus more information to utilize. In general, by taking average over localization results computed from different pairs of speakers, higher localization accuracy can be achieved, which means the more speakers in the environment, the better the localization performance will be. In this experiment, we test with poor conditions where only 3 speakers in total are present in the environment.

The CDF of the localization errors are shown in Fig. 9. When there are only 3 speakers, the system could achieve up to 80% percentile of the localization error is under $2m$, and 89% under $2.5m$. For the lobby environment, since there is no obstacle between the user and speakers (may have people passing through sometimes), the performance of the system is better than it in the garage. Since there are several pillars (as shown in Figure 8) in the space, and both the direction finding and localization results are definitely affected by such obstacles. But it also demonstrate that our system could run in complex situations, and the accuracy still could be improved.

6 CONCLUSION AND FUTURE WORK

In this paper, we propose SwingLoc, an ultrasonic system to study indoor positioning problem. We leverage wearable sensing techniques and make use of Doppler effects on acoustic signals. We improve the system efficiency by piggybacking on user's smartphone and decrease the estimation error caused by considering the human natural swing motion and walking conditions. Experimental results show that SwingLoc is robust to real world scenarios, and has great potential to provide desirable indoor positioning services.

Though our proposed indoor positioning solution achieved acceptable performance in the evaluation, it can still be further improved by enhancing the accuracy. For example, we discussed solution for 2-speaker localization case in section 4.3 but do not carry out the real-world evaluation. In the future, we would like to collect more digital information of the building and apply building information modeling (BIM) method to test 2-speaker cases.

REFERENCES

[1] Isaac Amundson and Xenofon Koutsoukos. A survey on localization for mobile wireless sensor networks. *Mobile entity localization and tracking in GPS-less environnments*, pages 235–254, 2009.

[2] Paramvir Bahl and Venkata N Padmanabhan. Radar: An in-building rf-based user location and tracking system. In *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 2, pages 775–784. Ieee, 2000.

[3] Chao Cai, Rong Zheng, and Menglan Hu. A survey on acoustic sensing. *arXiv preprint arXiv:1901.03450*, 2019.

[4] F Donovan. Indoor location market to reach $4 billion in 2018. *Online. FierceMobileIT.[Online]. Available: http://www. fiercemobileit. com/story/indoorlocation-market-reach-4-billion-2018-predicts-abi/2013-10-18, Accessed November*, 28, 2013.

[5] Negar Ghourchian, Michel Allegue-Martinez, and Doina Precup. Real-time indoor localization in smart homes using semi-supervised learning. In *AAAI*, pages 4670–4677, 2017.

[6] Wei Gong and Jiangchuan Liu. Robust indoor wireless localization using sparse recovery.

[7] Wenchao Huang, Yan Xiong, Xiang-Yang Li, Hao Lin, Xufei Mao, Panlong Yang, and Yunhao Liu. Shake and walk: Acoustic direction finding and fine-grained indoor localization using smartphones. In *INFOCOM, 2014 Proceedings IEEE*, pages 370–378. IEEE, 2014.

[8] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. Spotfi: Decimeter level localization using wifi. In *ACM SIGCOMM Computer Communication Review*, volume 45, pages 269–282. ACM, 2015.

[9] Patrick Lazik and Anthony Rowe. Indoor pseudo-ranging of mobile devices using ultrasonic chirps. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, pages 99–112. ACM, 2012.

[10] Hongbo Liu, Yu Gan, Jie Yang, Simon Sidhom, Yan Wang, Yingying Chen, and Fan Ye. Push the limit of wifi based localization for smartphones. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 305–316. ACM, 2012.

[11] Kaikai Liu, Xinxin Liu, and Xiaolin Li. Guoguo: Enabling fine-grained indoor localization via smartphone. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pages 235–248. ACM, 2013.

[12] Dimitrios Lymberopoulos, Jie Liu, Xue Yang, Romit Roy Choudhury, Vlado Handziski, and Souvik Sen. A realistic evaluation and comparison of indoor location technologies: Experiences and lessons learned. In *Proceedings of the 14th international conference on information processing in sensor networks*, pages 178–189. ACM, 2015.

[13] Rajalakshmi Nandakumar, Krishna Kant Chintalapudi, and Venkata N Padmanabhan. Centaur: locating devices in an office environment. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 281–292. ACM, 2012.

[14] Anshul Rai, Krishna Kant Chintalapudi, Venkata N Padmanabhan, and Rijurekha Sen. Zee: Zero-effort crowdsourcing for indoor localization. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 293–304. ACM, 2012.

[15] Zheng Sun, Aveek Purohit, Raja Bose, and Pei Zhang. Spartacus: spatially-aware interaction for mobile devices through energy-efficient audio sensing. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pages 263–276. ACM, 2013.

[16] Xuyu Wang, Lingjun Gao, Shiwen Mao, and Santosh Pandey. Csi-based fingerprinting for indoor localization: A deep learning approach. *IEEE Transactions on Vehicular Technology*, 66(1):763–776, 2017.

[17] Kaishun Wu, Jiang Xiao, Youwen Yi, Min Gao, and Lionel M Ni. Fila: Fine-grained indoor localization. In *INFOCOM, 2012 Proceedings IEEE*, pages 2210–2218. IEEE, 2012.

[18] Michel Allegue Xi Chen, Chen Ma and Xue Liu. Taming the inconsistency of wi-fi fingerprints for device-free passive indoor localization. In *INFOCOM, 2017 Proceedings IEEE*. IEEE, 2017.

[19] Yaxiong Xie, Zhenjiang Li, and Mo Li. Precise power delay profiling with commodity wifi. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 53–64. ACM, 2015.

[20] Jie Xiong, Karthikeyan Sundaresan, and Kyle Jamieson. Tonetrack: Leveraging frequency-agile radios for time-based indoor wireless localization. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pages 537–549. ACM, 2015.

[21] Zheng Yang, Chenshu Wu, and Yunhao Liu. Locating in fingerprint space: wireless indoor localization with little human intervention. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 269–280. ACM, 2012.

[22] Zheng Yang, Zimu Zhou, and Yunhao Liu. From rssi to csi: Indoor localization via channel response. *ACM Computing Surveys (CSUR)*, 46(2):25, 2013.

[23] Moustafa Youssef and Ashok Agrawala. The horus wlan location determination system. In *Proceedings of the 3rd international conference on Mobile systems, applications, and services*, pages 205–218. ACM, 2005.