Enhancing Localization Scalability and Accuracy via Opportunistic Sensing

Kaikai Liu*, and Xiaolin Li

***Abstract*—Using a mobile phone for fine-grained indoor localization remains an open problem. Low-complexity approaches without infrastructure have not achieved accurate and reliable results due to various restrictions. Existing accurate solutions rely on dense anchor nodes for infrastructure and are therefore inconvenient and cumbersome. The problem of beacon signal blockage further reduces the effective coverage. In this paper, we investigate the problems associated with improving localization scalability and accuracy of a mobile phone via opportunistic anchor sensing, a new sensing paradigm which leverages opportunistically connected anchors. One key motivation is that the scalability of the infrastructure-based localization system can be improved by lifting the minimum requirement for anchor numbers or constellations in trilateration. At the same time, location accuracy under insufficient anchor coverage will be improved by exploring the opportunity of diverse data types rather than deploying more anchor nodes. To enable this highly scalable and accurate design, we leverage low-coupling hybrid ranging using our low-cost anchor nodes with centimeter-level relative distance estimation. Activity patterns extracted in users' smartphones are utilized for displacement compensation and direction estimation. The system also scales to finer location resolution when anchor access is improved. We introduce robust delay-constraint semidefinite programming in location estimation to realize optimized system scalability and resolution flexibility. We conduct extensive experiments in various scenarios. Compared with existing approaches, opportunistic sensing could improve the location accuracy and scalability, as well as robustness, under various anchor accessibilities.**

***Index Terms*—Localization, Smartphone, Anchor network, Motion, Acoustic.**

I. Introduction

Ubiquitous smartphone and location information are enabling new features of location-based services (LBS) around local navigation, retail recommendations, proximity social networking, and location-aware advertising. Recently, the focus has also been shifting geographically from outdoors to indoors, where we spend the most money, meet friends, work, and do business.

Location information has infiltrated our everyday life in ways that we had not imagined before. The indoor location market will be more enormous than the outdoor, since we spend more than 80% of our time indoors on our daily activities, e.g., working, shopping, eating, etc. Technologically speaking, outdoor localization techniques cannot be directly applied indoors. Satellites-based localization, i.e., GPS, has been one of the most important technological advances of the last half century. However, no matter how effective these systems are outdoors, their accuracy, coverage, and quality deteriorate significantly in small-scale indoor spaces. Over the past few years, a broad variety of location services has been targeted to revolutionize how people sense and interact with everyday objects and locations. For example, sensor networks help firefighters find the best route for search and rescue; GPS and WLAN systems provide coarse-grained way-finding and navigation services [1], [2], [3], [4], [5], [6], [7]; RFID and short range communication devices provide proximity detection and awareness [8], [9], [10], [11], [12], [13]. As the market needs keep escalating, customers need new and fine-grained location-based services and applications, for example, offering in-store deals, delivering biographies or art commentaries on a specific objects when wandering through a museum, reminding users of social events in close proximity, dispatching police officers to indoor/outdoor places in need, navigating a blind or visually impaired person visiting public places with step-by-step navigation, or enabling augmented-reality for city visitors with fine-grained location beacon support of the VR headset or smartphone. However, different location services have dramatically different application needs. **Different Accuracy Requirements:** Meter-level (e.g., GPS with five-meter accuracy) localization is sufficient to navigate a car (meter-level footprint) on a street (several-meter footprint), but it is far from sufficient to navigate a user (foot-level footprint) in a library (with half-meter-wide aisles and inch-level books). **Speed in terms of Latency or Refresh Rate:** Tens of seconds delay for outdoor navigation could lead to a wrong turn decision for drivers; with second-level delay it is still hard to get the complete moving trace of a basketball player for performance analysis and evaluation. **Various Costs:** With infrastructure already installed throughout the great indoors, the easiest solution may be Wi-Fi fingerprinting approaches. While the cost may be attractive, the accuracy is barely sufficient for differentiating different stores, e.g., WiFi access point and cellular towers, could only achieve limited accuracy, or need extensive war-driving and calibration [3], [7]. Other accurate approaches rely on the deployment of additional infrastructure [12], [14], [15], [16], [17], e.g., dense anchor nodes. These approaches have a high requirement for the minimum anchor number, e.g., at least three anchors for 2-D trilateration.

In this paper, we propose a highly accurate and scalable mobile phone localization system via opportunistic anchor access. This work narrows down to fine-grained localization techniques for future location-aware applications, for example,

K. Liu is with the Computer Engineering Department, San Jose State University, San Jose, CA, 95192. X. Li is with the Scalable Software Systems Laboratory, University of Florida, Gainesville, Florida, 32611.
E-mail: kaikai.liu@sjsu.edu, andyli@ece.ufl.edu
* Corresponding author. The work presented in this paper is supported in part by National Science Foundation under Grant No. CNS 1637371. The conference version of the paper was published in IEEE SECON 2016 and got the Best Paper Award.

step-by-step indoor navigation for the blind, locating Virtual Reality (VR) headsets in gaming, and robotic indoor navigation, mapping, and autonomous driving. The key motivation of our design is originated from the problem associated with acoustic anchor-based solution [14] with centimeter-level accuracy, i.e., the acoustic anchor can be easily blocked, and it is hard to access the minimum *three* anchors for trilateration in real environments. Deploying more anchor nodes is not economically practical. Furthermore, most fine-grained solutions rely on high timing accuracy of entire anchor networks. Increasing the anchor number also increases the management burden of the timing accuracy. Leveraging multi-modal opportunistic sensor data instead of deploying more nodes is a good step toward improving the efficiency and scalability. However, existing multi-modal solutions mainly focused on meter-level coarse-grained applications, for example, using motion data to improve the room-level or meter-level WiFi localization accuracy. For fine-grained applications, such approaches are not well suited. One particular example is that combining meter-level results with centimeter-level results would downgrade the overall performance.

Transforming this high-level design goal into a practical working system poses significant challenges: (1) How can we improve the location accuracy even with limited anchor nodes? (2) Will our system adapt to higher accuracy with more data types or better anchor accessibility? This paper addresses these challenges, and prototypes the system via *opportunistic anchor sensing*. Testbed results confirm that our design goals could adapt to different anchor coverage and service quality requirement with high scalability, e.g., from only one node to multiple nodes with multi-modal data. We believe this could be a practical approach to achieve fine-grained localization results with very low hardware requirement and deployment costs, while also being scalable for applications with fine-grained resolution demand, e.g., indoor accessibility assistance for the visually impaired, and robotic navigation.

II. System Overview

A. Motivation

The major metrics of existing spatial enabling technologies, most notably **accuracy**, and **deployability** performance, are far from satisfactory. Existing trilateration approaches require proper GDOP of the infrastructure coverage, for example, deploying at least three nodes around the area of interest for trilateration. However, many real world scenarios pose significant constraints in terms of the number of nodes deployed and the geometric distribution of the nodes. For example, existing infrastructure-based fine-grained localization solutions require a minimum of *three* anchors [14]. The achieved localization accuracy drops significantly when the number of accessed anchor nodes is less than four; it is in fact unlocalizable with less than three nodes. Even if there are 3 anchors, the achieved accuracy is very low compared with scenarios involving redundant anchor coverage (larger than 3). There exist inherent tradeoffs between the localization accuracy and the deployment complexity. Existing low accuracy or high-complexity indoor localization solutions in a mobile phone call for significant innovations in balancing both accuracy and complexity.

Highly dynamic and mobile settings, where humans are essentially moving, present further challenges for current solutions either using existing infrastructure or self-deployed anchor networks. Fig. 1 shows a real case of a person moving across multiple local anchor networks. The nodes deployed along the way can not provide sufficient coverage, i.e., the GDOP is very high. When the target moves from one location to another, i.e., t_1 to t_5 in Fig. 1, the number of accessible anchor nodes changes over time. In anchor coverage area 1, the target can perform trilateration to get its location; however, the location of the target is unsolvable using traditional trilateration when it moves into the coverage area 2 (only one or two anchor nodes are accessed). The only obtainable error surface forms a ring shape, with the distance as the radius and the RSS ranging accuracy (meter-level) as the width. Thus, the accuracy, coverage, and quality of location services deteriorate significantly and even stop working in real-world environments.

B. System Challenges

Lowering the minimum anchor number requirement is essential for system scalability and low-complexity. The approach proposed in this paper works under different "anchor" coverage environments, i.e., scales from only one "anchor" rather than the minimum *three* anchors. For applications with a limited budget for anchor deployment, deploying several anchor nodes (less than three) could still achieve significantly higher accuracy than an anchor-free system. When the number of deployed anchors increases (larger than three), our proposed algorithms can adapt to reach highly accurate results as a result of the additional anchors. With this flexibility, location service operators could select configurations that suit various service quality requirements.

In this project, we propose a highly accurate and scalable location sensing system via opportunistic anchor access. The key motivation of our design originated from the problems associated with insufficient anchor coverage for trilateration in real environments. Deploying more anchor nodes is economically impractical and hard to manage, especially in maintaining high timing accuracy of the complex network. Leveraging more data types instead of more nodes is a better option. However, existing hybrid solutions mainly focuse on meter-level coarse-grained applications [18], and these approaches do not match well with fine-grained approaches. For example, combining meter-level results with centimeter-level results would downgrade overall performance instead of improve it. Moreover, transforming the high-level system design goal into a practical working system poses significant challenges in 1) adaptively fusing sensor results according to application demand and 2) making the solution robust at various levels of infrastructure coverage. The **major research problem** is how to achieve the desired accuracy when only a few nodes are accessed (i.e., trilateration cannot be used), and how to scale to higher accuracy when the coverage of the infrastructure is sufficient. Specifically, we utilize three *opportunistic* sensing

sources: 1) ranging information for the anchor node and smartphone pair; 2) relative angle information; and 3) displacement and moving direction of the target. If multiple anchor nodes could be sensed by the smartphone, i.e., multiple ranging and relative direction results, the accuracy of the location estimation result could be further improved and optimized via multiple constraints.

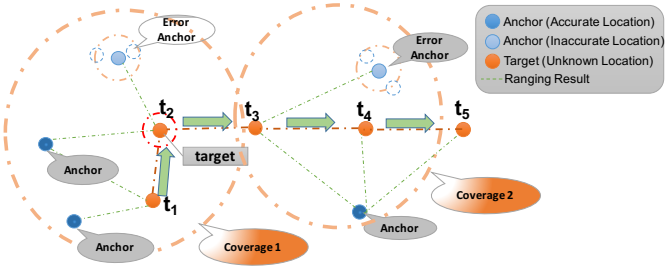


Fig. 1. System model.

C. System Design and Implementation

In this paper, we present an *opportunistic sensing* scheme that involves Bluetooth-Low-Energy (BLE) based received-signal-strength (RSS) ranging, acoustic-based angle measurement, relative Time-of-Arrival (TOA) ranging, and displacement and moving direction estimation. We further fuse these hybrid data, and develop an algorithm which achieves centimeter-level accuracy with sufficient measurements. The overall system architecture is shown in Fig. 2.

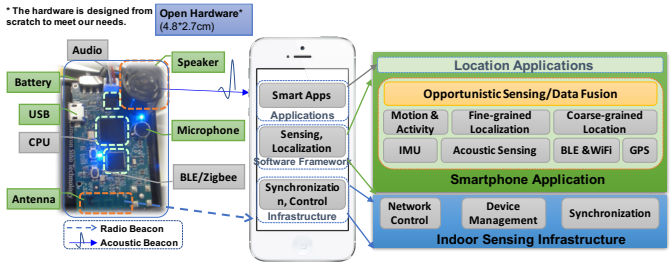


Fig. 2. System Implementation.

We propose a smartphone-based indoor localization system using our newly designed low-complexity anchor node, because opportunistic sensing will never succeed without adequate anchor nodes as sensing sources. We select the BLE and acoustic signal as the two main beacon signal types. The BLE signal is low-power and high efficiency in terms of RSS scanning compared to WiFi; it is also the reason that Apple and Qualcomm both select BLE as the main beacon source [19], [20]. Due to the strong promotion force by these industry leaders, and the convenient iBeacon APIs introduced in iOS7 [19], we propose to design an anchor node that could support Apple's iBeacon specifications.

We design this low-complexity anchor node from scratch to lower the overall cost and meet our long-term objective. The overall bill of material (BOM) price is kept lower than $20 to allow for low cost and thus potential for wider availability. The system architecture is shown in Fig. 2 with three important parts: BLE radio, microcontroller, and audio codec. The power system is designed to adapt power from multiple sources, i.e., micro-USB, lithium battery and solar-panel.

We implement the signal processing, ranging, and trilateration-based localization algorithm inside the smartphone. To fast deploy the developed algorithm, we design a back-end server for the smartphone to optionally offloading complex processing tasks. This approach balances computation and network consumption in smartphones, and the introduced delay is less than $100ms$, which is negligibe for the sub-second level location update rate. It also helps the debugging process, where the data will be transmitted to the backend server in real-time. A Redis server [21] is used as a cloud key-value store based on Pub/Sub for the smartphone; A Java server performs computation for the structured data in the Redis server and sends the final results back to the smartphone.

III. Multimodal Ranging and Measurements

In the simplest case, one anchor node provides basic check-in services, and coarse-grained location estimation. To improve accuracy, we utilize the acoustic ranging for distance measurement, inertial navigation system (INS) for displacement and direction estimation, and leverage activity measurement for error mitigation. To support multiple users simultaneously, the ranging process occurs entirely in one-way passive mode, whereby only the smartphone needs to receive beacons. Thus, the highly accurate TOA result is in the pseudorange between the smartphone and anchor pair. The inertial sensors, i.e., accelerometer and gyroscope, are utilized for accurate displacement and direction estimation.

When utilizing the omnidirectional ranging information with the single anchor node, the performance improvement is limited if the rough direction is unknown. To estimate the direction, BLE signal needs an expensive directional antenna; while the speaker of acoustic signal is mostly directional. Thus, combining BLE and acoustic for ranging could be an economical way to obtain coarse-grained distance and angle. Another benefit comes from the acoustic TOA estimation result, which is beneficial if the user start moving under single anchor coverage.

A. The Model of Signal Detection and Ranging

Signal Detection. The total number of anchor nodes is M_A, where m is the index with $m = 1, \ldots, M_A$. Each anchor node broadcasts its own unique pseudocode sequence $\mathbf{p_m} = [p_j]$ with length L for the m-th anchor node ($j = 1, \ldots, L$). The symbol duration of the acoustic beacon is T_s, i.e., the total time for each beacon is LT_s.

Assume the sampling rate of the smartphone is F_s, and the received acoustic signal sample is $g(k)$. We decode the $\hat{p}_j$ associated with the current symbol ($\hat{p}_j$ for the estimated version of p_j, with the vector term as $\hat{\mathbf{p}}$). By performing code matching with the pre-stored pseudocode sequence $\mathbf{p_m}$, we could obtain the station id m for the j-th symbol.

Ranging. The basic process of ranging is to measure the flight delay (t_j) of the first one in all multipaths, i.e., $r = v_s \times t_j$,

where v_s is the acoustic sound speed. In the discrete sample domain, we estimate the sampling point of TOA path $\hat{k}_j$ for the j-th symbol as a TOA value of $\hat{t}_j = \hat{k}_j / F_s$. In this process, we associate the TOA measurement $\hat{k}_j$ to the m-th anchor node and l-th index of the pseudocode $\mathbf{p_m}$. We also convert $\hat{t}_j$ into the base symbol time ($j = 1$) and add it into the vector of TOA measurement $\mathbf{k}_m = [\hat{k}_j - jT_sF_s, \ldots]$, and obtain ranging measurements as $\hat{\mathbf{r}}_m = \mathbf{k}_m v_s / F_s$.

Due to the one-way passive ranging mode utilized for multi-user simultaneous access, the distance measured by TOA estimation is *pseudorange*, with unknown bias δ_r. To solve this unknown bias, we synchronize all the M anchor nodes, and make the δ_r fixed for every node.

B. BLE and Acoustic Ranging

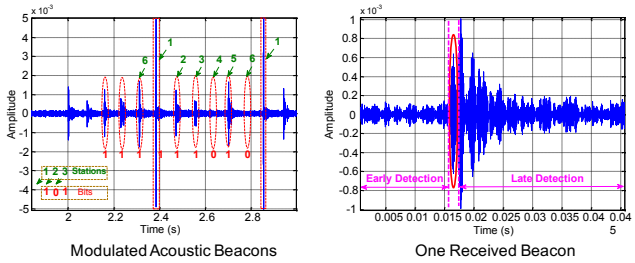


Fig. 3. (a) Modulated acoustic beacons; (b) One received beacon.

1) Acoustic TOA Based Ranging: TOA based ranging approaches rely on delay estimation of the beacon signal from the anchor node to the mobile phone. Based on [14], we design acoustic one-way TOA ranging approaches to support multiple users simultaneously, and make the beacon signal unnoticeable to the human ear. Leveraging the DAC and miniDSP module in our audio chip, we generate modulated signal beacons based on the pseudocode $\mathbf{p_m}$. As shown in Fig. 3 (a), the transmitted acoustic beacons are modulated based pulse amplitude modulation (PAM). The smartphone samples the acoustic beacon in full resolution, and obtain the channel response in full detail as shown in Fig. 3 (b).

The basic process of ranging is to measure the flight delay (t_j) of the first one in all multipaths as shown in Fig. 3 (b), i.e., $r = v_s \times t_j$, where v_s is the acoustic sound speed. In discrete sample domain, we estimate the sampling point of TOA path $\hat{k}_j$ for the j-th symbol as TOA value of $\hat{t}_j = \hat{k}_j / F_s$. In this process, we associate the TOA measurement $\hat{k}_j$ to the m-th anchor node and l-th index of the pseudocode $\mathbf{p_m}$ demodulated from the beacon as shown in Fig. 3 (a). We also convert $\hat{t}_j$ into the base symbol time ($j = 1$) and add it into the vector of TOA measurement $\mathbf{k}_m = [\hat{k}_j - jT_sF_s, \ldots]$, and obtain ranging measurements as $\hat{\mathbf{r}}_m = \mathbf{k}_m v_s / F_s$. After detecting the first path in the channel response, the delay and energy of the first path is the TOA (t_m) and RSS value (ε_m^r). The obtained TOA ranging measurement is $\tilde{r}_{n,m}$ for the n-th user from the m-th anchor.

To further improve the robustness, we differentiate multi-paths to get the highly accurate ranging results as shown in Fig. 4. The *first path*, i.e., the TOA path, in differentiated multi-paths may not be the strongest (NLOS case), but it provides an accurate representation of the real physical distance. We utilized the iterative optimization approaches to predict the *first path* region via the previous location estimation results. These two approaches combined together can help significantly reduce the rate of outliers and large errors. As shown in Fig. 4, we can extract fine-grained multi-path beacon components from the initial noisy signals.

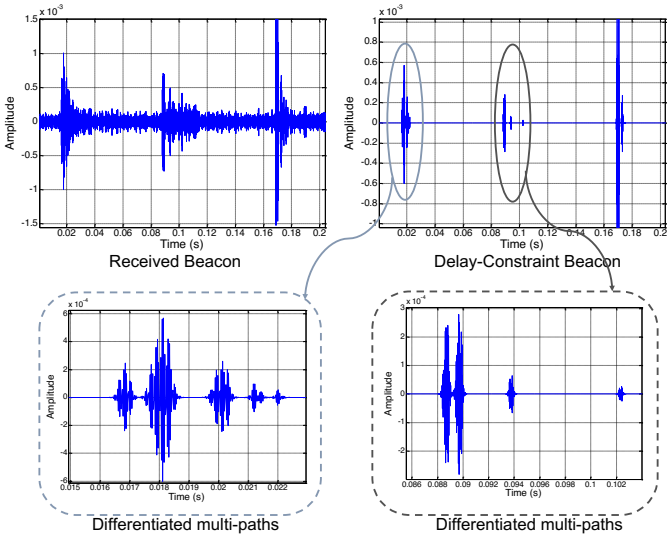


Fig. 4. Differentiated multi-path signals via delay constraints.

Due to the one-way passive ranging mode utilized for multi-user simultaneous access, the distance measured by TOA estimation is *pseudorange*, with unknown bias δ_r. To solve this unknown bias, we need to synchronize all the M anchor nodes, and make the δ_r fixed for every nodes.

2) RSS Based Ranging: One drawback of one-way TOA-based ranging lies in its relative measurement feature, i.e., the bias caused by clock drift in (5) is unknown. To obtain an absolute distance measurement, we can directly utilize the acoustic and BLE RSS for coarse-grained ranging, and then apply the fine-grained relative TOA distance as a constraint when the user is moving.

The RSS from BLE is calculated at the packet level, where the energy is averaged over all multi-path components. Using the BLE RSS for ranging, we can obtain a fast energy scan at low cost, e.g., micro-second level scanning time for the BLE signal in iOS platform, compared to a WiFi scan, which takes nearly one second and is even not allowed by Apple.

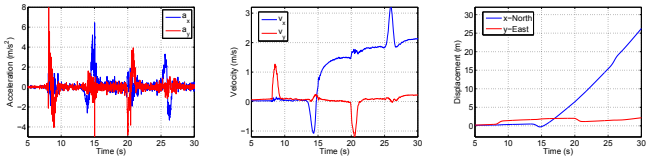
Compared with the RSS value obtained from the BLE signal, the RSS from the acoustic beacon of the anchor node is more accurate and sensitive, especially in short range and LOS conditions. Moreover, as the full channel impulse response is available in our acoustic ranging system, we can obtain the signal energy (ε_m^r) of the TOA path as the RSS, which is more precise and shows better robustness on multi-paths in indoor complex environments.

Although the RSS ranging result from acoustic beacon is more accurate, the coverage and the NLOS robustness are still worse than BLE beacon. If the received TOA RSS is unavailable, the RSS ranging result should be replaced by the

BLE RSS ranging value. Thus, the RSS ranging result $\hat{r}_{n,m}^{(k)}$ is a combined value from BLE and acoustic signals.

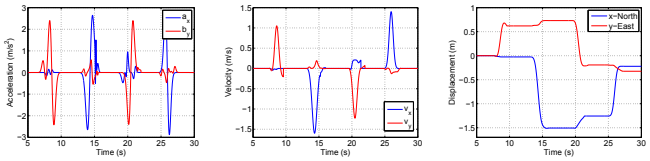
C. Displacement Estimation

To estimate the relative translation ($\mathbf{t}$) of the smartphone, we can utilize the Accelerometer on board by measuring its force, acceleration, and infer the displacement by double integrating the acceleration. The quantity resulting from the Accelerometer is the acceleration rate in m/s^2, and can be denoted as $\mathbf{f}_t^b = (f_t^{b_x}, f_t^{b_y}, f_t^{b_z})^T$, where $f_t^{b_{x,y,z}}$ is the measured force in 3-axis directions in the *body* frame.



(a) Acceleration (b) Velocity (c) Displacement

Fig. 5. Motion estimation result via conventional method: (a) acceleration, (b) velocity, (c) displacement.



(a) Acceleration (b) Velocity (c) Displacement

Fig. 6. Motion estimation result via proposed method: (a) acceleration, (b) velocity, (c) displacement.

Direct integration $\mathbf{f}_t^b$ obtains the displacement in the *body* coordinate, which is not related to the real geodesic displacement. To convert the obtained acceleration of the smartphone to the local *navigation* coordinate, we could apply rotation and translation over $\mathbf{f}_t^b$ by

$$\mathbf{f}_t^n = \mathbf{R}_b^n \mathbf{f}_t^b + \mathbf{e}^n \tag{1}$$

where $\mathbf{e}^n$ is the error of the force that is applied to the smartphone. $\mathbf{R}_b^n$ is the rotation matrix used to convert the coordinate, which can be directly accessed from mobile OS [22]. To obtain the acceleration caused by the applied forces, gravity should be subtracted by $\mathbf{a}_t^n = \mathbf{f}_t^n - \mathbf{g}$, where $\mathbf{g} = [0, 0, g]$ is the gravity vector.

We use a rectangular shaped movement to quantify the displacement estimation accuracy. The user walks in a rectangular path and go back to the original point. The estimated displacement trace should be close to the ground truth with the ending point overlaping with the starting point. The measured acceleration results after gravity subtraction are shown in Fig. 5a. After the denoising process, the velocity of the smartphone can be obtained by $\mathbf{v}_t^n = \mathbf{v}_0^n + \int_0^t \mathbf{a}_t^n$ as shown in Fig. 5b. From Fig. 5b, we know that the velocity drifts even when the user is stationary. The displacement can be calculated by $\mathbf{s}_t^n = \mathbf{s}_0^n + \int_0^t \mathbf{v}_t^n$, where $\mathbf{v}_0^n$, and $\mathbf{s}_0^n$ are the initial velocity and displacement. The results of estimated displacement in the x and y directions are shown in Fig. 5b. As shown in Fig. 5b, the starting point (5 second) and the ending point (30 second) do not overlap with each other, which means the estimated displacement trace has a significant drift (> 20 meters) when compared to the group truth.

The process of obtaining relative displacement $\mathbf{s}_t^n$ involves double integration, whereby the measurement noise, i.e., $\int\int \mathbf{e}^n$, is also integrated and amplified. Here, the white noise in acceleration measurements is integrated *twice* and causes a second-order random walk in displacement of the smartphone. As a result, bias errors lead to errors in position that grow proportionally to t^2. The error of the accelerometer measurement can be modeled as

$$\dot{\delta} = -\frac{1}{T_a}\delta + \mathbf{w_a} \tag{2}$$

where T_a is the correlation time of the accelerometer. The value of the T_a differs across devices, and needs to be estimated prior to calibration. $\mathbf{w_a}$ is the modeled Gaussian white noise.

D. Mitigating Displacement Estimation Error

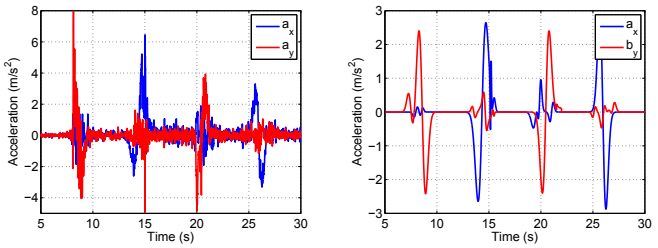
Performing moving pattern extraction before integration can be a feasible approach to calibrating drift. The normal movement model cannot be applied to estimating human's movement. That is, we cannot walk at constant velocity (CV) or constant acceleration (CA) like a vehicle or a plane. Human walking or movement has its own pattern, and we need to "accelerate" and "decelerate", and then "accelerate" again for another footstep. Here we use the "start-moving-stop" (SMS) movement model.

Performing "start-moving-stop" pattern decomposition could help us estimate the displacement in a more meaningful way. Using start, acceleration, deceleration, and stop as one basic step, the velocity changes (from zero to top to zero) can be modeled as a Gaussian shape $g_v(\mu, \sigma) = \pm \exp\{-(x-\mu)^2/(2\sigma^2)\}$, where + means moving forward and − means backward. In this case, acceleration is the derivative of the velocity, i.e., $g_a(\mu, \sigma) = \pm(x-\mu)/\sigma^2 \exp\{-(x-\mu)^2/(2\sigma^2)\}$. Using $g_a(\mu, \sigma)$ as the *kernel* function, we can decompose the acceleration measurement $\mathbf{a}_t^n$ into a series of $\eta g_a(\mu, \sigma)$ with different parameters η, μ, and σ. Thus, the decomposed series of acceleration is $\sum_{i=1}^{n} \eta_i g_a(\mu_i, \sigma_i)$, where η_i is the amplitude of each Gaussian derivative pulse. The fitting process can be modeled as

$$\{\eta_i, \mu_i, \sigma_i\} = \min_{\{\eta_i, \mu_i, \sigma_i\}} ||\mathbf{a}_t^n - \eta_i g_a(\mu_i, \sigma_i)|| \tag{3}$$

To reduce the number of parameters during the fitting process, we extract the feature points of $\mathbf{a}_t^n$, e.g., peak position and width, by thresholding the peak maximum and rising edges. We then use the number of peaks found and the peak positions and widths to fit the specified peak model. This combination yields better and faster computation, and deals with overlapping peaks as well. During the decomposition and fitting process, the sign of η_i is determined by comparing the remaining error of using positive and negative results.

Fig. 7a shows the measured acceleration of (1). Instead of directly integrating the noisy results, we use the "start-moving-stop" (SMS) movement model to extract the key moving features, and reconstruct the acceleration as shown in Fig. 7b, which is significantly clean and represents most of the original moving details.



(a) Measured acceleration (b) Modeled acceleration

Fig. 7. (a) Measured acceleration and (b) the modeled acceleration.

After applying the SMS approach for the results in Fig. 5, we could obtain better acceleration, velocity, and displacement results as shown in Fig. 6. The ground truth of the moving path is like a rectangle, where the starting point and ending point overlap. The velocity is calibrated as multiple start-moving-stop patterns. Comparing with Fig. 5c, the drift of the estimated displacement in Fig. 6c is significantly reduced (<25cm), and the starting point and end point of the walking trace are very close.

IV. Opportunistic Sensing and Multi-modal Constraints

With the acoustic ranging results and the displacement and direction estimation, we perform location optimization to the initial location even with a single anchor.

A. Localization Model

The signal detection, ranging, localization and tracking are the four basic steps involved in the whole system. For the localization and tracking process, we define $\mathbf{y}^n$ and $\mathbf{x}_m^n$ as the positions of the smartphone and m-th anchor in the n-frame. Using M pseudoranges $\hat{\mathbf{r}}_m$ and the preconfigured coordinates of anchor nodes $\mathbf{x}_m^n$, we can estimate the 3D position of the smartphone $\mathbf{y}^n$ by minimizing the quadric term of the remaining error

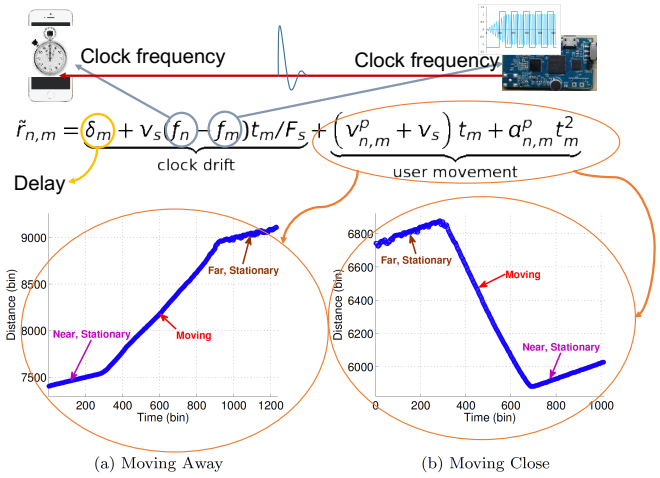
$$\varepsilon_m = ||\hat{\mathbf{r}}_m - (||\mathbf{y}^n - \mathbf{x}_m^n||_2 + \delta_r)||_2 \quad (4)$$

where δ_r is the unknown delay that compensates for the difference between the pseudorange and real distance. The unknown bias (δ_r) can be estimated during the localization process with sufficient anchor numbers, e.g., solving a 3D location (x, y, z) needs four equations (anchor nodes) instead of three.

B. Initial Location Constrains

The initial location of the smartphone could be accessed by using the API provided by the mobile operating system via conventional GPS and WiFi localization approaches. The obtained initial location is in geodetic coordinates (WGS 84 datum latitude ϕ, longitude λ, height h) with room-level accuracy. To convert the geodetic coordinates to the *navigation* coordinate, we first convert it to the earth-centered, earth-fixed (ECEF) coordinate, and then convert the ECEF to the ENU frame. By subtracting the reference point O_R, the GPS location is mapped to the *navigation* coordinate (n-frame) for more intuitive and practical analysis. We define the POI's 3D position as $\mathbf{x}_i^n = [x_i^n, y_i^n, z_i^n]^T$, where the superscript n denotes the position value in the *navigation* coordinate and i denotes the i-th POI in M POIs ($i = 0, M-1$). The current location of the smartphone is defined as $\mathbf{p}^n = [x^n, y^n, z^n]^T$. In most location applications, the z-axis z^n is not required, and we can convert the 3D location space into the 2D coordinate as $\mathbf{p}^n = [x^n, y^n]^T$.

C. Delay Constraints



(a) Moving Away (b) Moving Close

Fig. 8. Delay constraints.

Most of the existing literature on mobile phone localization focuses on the case where the ranging measurements are known with some slight perturbations, i.e., using zero-mean Gaussian noise to represent ranging error. This assumption is only effective when all the ranging results are in line-of-sight (LOS) conditions with no bias or large error. However, in practice, a significant portion of the ranging results contain outliers. For example, the non-line-of-sight (NLOS) propagation of the beacon introduces outliers or large errors.

One possible way to minimize these outliers is to have an accurate model of the delay, i.e., the distance. By predicting the region of the next beacon in the detection process based on the previous measurements, the outliers and large errors can be significantly reduced. For example, we can directly search the narrow region around the truth TOA path in Fig. 3 by leveraging the iterative feedback based on the previous location results.

As shown in Fig. 3, the TOA ranging process is designed to detect the first path in the channel response, i.e., the delay and energy of the first path is the TOA (t_m) and RSS values (ε_m^r). This one-way ranging mode does not require

peer synchronization, and is scalable for multiple users' cases. However, there is one unknown delay between the transmitter and receiver. When the user moves, the delay of the received beacon becomes more complicated due to the unknown movement of the users. The delay model can be formulated as the obtained TOA ranging measurement $\tilde{r}_{n,m}$ for the n-th user from the m-th anchor as

$$\tilde{r}_{n,m} = \underbrace{\delta_m + v_s(f_n - f_m)t_m/F_s}_{\text{clock drift}} + \underbrace{\left(v^p_{n,m} + v_s\right)t_m + a^p_{n,m}t^2_m}_{\text{user movement}} \tag{5}$$

where δ_m is the unknown bias for the m-th anchor node that maps the $\tilde{r}_{n,m}$ to the real distance; t_m is the TOA time for the m-th anchor; $(f_n - f_m)$ is the clock frequency drift between the n-th mobile phone (f_n) and m-th anchor node (f_m); $v^p_{n,m}$ and $a^p_{n,m}$ are the projected moving speed and acceleration of the n-th user in the direction to the m-th anchor. v_s is the speed of the acoustic signal and F_s is the sampling rate of the mobile phone. The first part of (5) is contributed to the clock drift, and can be approximated as a linear function to the time t_m, while the second part is contributed to the user movement, and becomes a curve (second or third order) when the user is moving.

After the localization and tracking process through which we obtain the estimated value of the users' movement, we can input the estimated results into (5) to get the predicted region of the next beacon signal. We model the clock drift as one unknown parameter, and estimate this parameter together with an unknown location. The detailed process is shown in Section. IV-F. Through this interactive process, we will eliminate large ranging errors by enforcing the delay constraints.

D. Geometric Movement Constraints

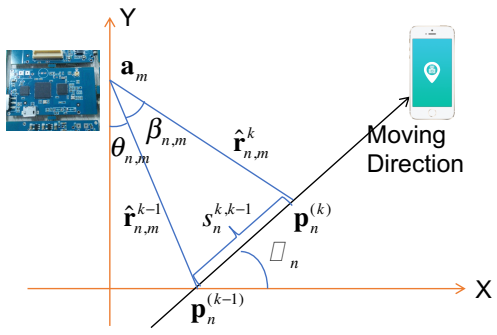


Fig. 9. Geometric movement constraint for one anchor node.

Assume the position coordinate of the anchor node is $\mathbf{a}_m \in \mathbb{R}^d$, where m is the index of total M anchor nodes. For 2-D coordinates $(d = 2)$, $\mathbf{a}_m$ is $[x_m, y_m]^T$, $m = 1, \ldots, M$. Denote the location coordinate of the n-th user as $\mathbf{p}_n$, $n = 1, \ldots, N$.

To refine the user's location, ranging information is utilized as a constraint. Assume the initial position coordinate of a user obtained by smartphone is $\hat{\mathbf{p}}_n$, which is direct from location API and low-accurate compared to the location of anchor nodes $(\mathbf{a}_m)$. Defining the RSS ranging measurement between the user and anchor pair is $\hat{r}_{n,m}$ and the estimated relative TOA distance is $\tilde{r}_{n,m}$. The real distance $r_{n,m}$ from the n-th mobile phone to the m-th anchor node is written as $r_{n,m} = ||\mathbf{p}_n - \mathbf{a}_m||_2$, where $||\cdot||_2$ calculates the 2-norm and obtains the Euclidean distance. The vector form of the RSS ranging observation from m-th anchor to n-th mobile phone can be written as

$$\hat{\mathbf{r}}_{n,m} = ||\mathbf{p}_n - \mathbf{a}_m||_2 + \hat{\mathbf{n}}_{n,m} \tag{6}$$

where $\hat{\mathbf{n}}_{n,m}$ is the measurement noise; $m = 1, \cdots, N^B$. Then, the TOA distance measurement from m-th anchor to n-th mobile phone is

$$\tilde{\mathbf{r}}_{n,m} = ||\mathbf{p}_n - \mathbf{a}_m||_2 + \delta_{n,m} + \mathbf{n}_{n,m} \tag{7}$$

where $\mathbf{n}_{n,m}$ is the TOA measurement noise, which is lower than $\hat{\mathbf{n}}_{n,m}$ in (6) $m = 1, \cdots, N^A$, where $N^A \leq N^B$. $\delta_{n,m}$ is the unknown bias between the m-th anchor and n-th mobile phone pair due to the unsynchronized clock. Thus, TOA result (7) shows the relative distance measured between the smartphone and anchor.

The obtained displacement could be another measurement that contributes to the location optimization. For k and $k+1$ measurements, the displacement can be written as

$$s^{k,k-1}_n = ||\mathbf{p}^{(k)}_n - \mathbf{p}^{(k-1)}_n||_2 + \mathbf{n}_s \tag{8}$$

The direction of the motion traces obtained from the altitude value is assumed as α_n as shown in Fig. 9. With the RSS ranging results and TOA relative ranging results from the anchor to the smartphone, the anchor-related measurement could be written as

$$\begin{aligned} \hat{\mathbf{r}}^{k-1}_{n,m}\cos(\theta_{n,m}) - \hat{\mathbf{r}}^{k}_{n,m}\cos(\theta_{n,m} + \beta_{n,m}) \\ = s^{k,k-1}_n \sin(\alpha_n) + \mathbf{n}_A \end{aligned} \tag{9}$$

where the angle $\beta_{n,m}$ could be calculated by the law of cosines. The geometric relation is shown in Fig. 9.

$\theta_{n,m}$ is the *angle* measurement information. When the user moves into the acoustic beacon coverage, the rough value of $\theta_{n,m}$ is related to the anchor installation direction, which is known. $\theta_{n,m}$ is also related to the location difference of $\mathbf{p}_n = [p_x, p_y]$ and $\mathbf{a}_m = [a_x, a_y]$, with its x and y coordinates related by $p_x = a_x + \hat{\mathbf{r}}^{k-1}_{n,m}\sin(\theta_{n,m})$ and $p_y = a_y - \hat{\mathbf{r}}^{k-1}_{n,m}\cos(\theta_{n,m})$. Due to the wide beam of the acoustic speakers, the accuracy of $\theta_{n,m}$ is not high. This is the inherent limitation of (9). However, the error surface reduction is significant when compared with the case in which the angle is not utilized.

E. Anchor Coverage Constraints

1) Cramer-Rao Low Bound and Fisher Information Matrix: To evaluate the position accuracy, the Cramér-Rao low bound (CRLB) is often used as a theoretical optimal value from any unbiased estimator. The CRLB can be written as the reciprocal of the fisher information. For an estimate of $\hat{\theta}$ obtained, we have the CRLB as

$$\mathrm{E}_r\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T\} \geq \mathbf{J}^{-1}_\theta \tag{10}$$

where (10) is in the form of covariance matrix, and the right part of (10) is $CRLB = \mathbf{J}^{-1}_\theta$, which means the variance of the estimated parameter could not be lower than the CRLB

under the given estimator. The term $\mathbf{A} \geq \mathbf{B}$ expresses that the difference $\mathbf{A} - \mathbf{B}$ of the real symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ is positive semidefinite. $\geq$ will be reduced to $\geq$ if θ is a scalar [23]. The $\mathbf{J}_\theta^{-1}$ in (10) is the Fisher information.

For 2-D coordinates, denote the unit vector from the m-th anchor node to target as $\alpha_m = (x - x_m)/r_m$ in x domain; $\beta_m = (y - y_m)/r_m$ in y domain, where $r_m = ||\mathbf{y} - \mathbf{x}_m||$. $\mathbf{J}_{\theta|(2\times2)}$ can be shown as

$$\mathbf{J}_{\theta|(2\times2)} = \begin{bmatrix} \sum_{m=1}^{M} {\alpha_m}^2/\sigma_r^2 & \sum_{m=1}^{M} \alpha_m\beta_m/\sigma_r^2 \\ \sum_{m=1}^{M} \alpha_m\beta_m/\sigma_r^2 & \sum_{m=1}^{M} {\beta_m}^2/\sigma_r^2 \end{bmatrix} \quad (11)$$

2) Anchor Network Coverage: Relying on the anchor network for mobile phone localization, the placement of anchor nodes is very important in achieving high resolution results. To evaluate the effects caused by the geometric layout of anchor nodes, the term of geometric dilution of precision (GDOP) is often used. GDOP can be defined as $GDOP = \sigma_p/\sigma_r$, where σ_p and σ_r indicates the variance of position and ranging results, respectively. σ_p is the variance of localization resulting in the previous iteration. The GDOP quantifies the amplification of the ranging error in the position result when passing through the position calculating unit.

For an unbiased estimator, GDOP can be written as

$$GDOP = \sqrt{\text{trace}(\mathbf{J}_{\theta|(2\times2)}^{-1})}/\sigma_r = \sqrt{\text{trace}(\tilde{\mathbf{J}}_{\theta|(2\times2)}^{-1})} \quad (12)$$
$$= \sqrt{\text{trace}\left(\begin{bmatrix} \sum_{m=1}^{M} {\alpha_m}^2 & \sum_{m=1}^{M} \alpha_m\beta_m \\ \sum_{m=1}^{M} \alpha_m\beta_m & \sum_{m=1}^{M} {\beta_m}^2 \end{bmatrix}^{-1}\right)}$$

where α_m and β_m comprise the unit vector in the x and y directions.

As (12) only depends on the positions of the anchor node $\mathbf{x}_m$ and mobile phone $\mathbf{y}$, it is independent of the ranging noise variance. The difference between GDOP (12) and CRLB lies in the ranging variance σ_r.

The GDOP calculated by (12) can be used as a metric for evaluating the anchor network coverage. The value of GDOP is determined by the relation between the position of mobile phone ($\mathbf{y}$) and anchor networks ($\mathbf{x}_m$). A smaller value of GDOP indicates solid coverage. If the location estimation result ($\hat{\mathbf{y}}$) is available, the corresponding GDOP value of (12) can be calculated. If the GDOP value for $\hat{\mathbf{y}}$ is higher than the threshold, then the confidence level of $\hat{\mathbf{y}}$ should be lowered. Such a scheme could be used as a post-position constraint to filter out the localization results. If the ranging variance is available or set as a constant, using CRLB in (10) as the metric for evaluating the coverage is equivalent to GDOP.

The GDOP value of (12) can be calculated by inserting the estimated location results from the previous iteration. If the GDOP value is higher than a threshold, then the result $\hat{\mathbf{y}}$ should be mitigated. Specifically, if the calculated position result falls outside of the anchor network coverage, the confidence of the result is low (high GDOP value). The threshold is determined by the minimum acceptable resolution and is calculated by experiment. Using the GDOP as the coverage constraint, the problematic localization results could be filtered out. If the ranging variance is available, using CRLB of (10) instead of GDOP is equivalent. The filtered result of $\hat{\mathbf{y}}$ is the final calculated position of mobile phones, and can be delivered to the user for other location-based-services (LBS).

F. Location Optimization

Thus, the location optimization problem when the anchor node number is insufficient for trilateration can be defined by using (6), (7), (8), and (9) to obtain a refined result of $\mathbf{p}_n$.

For the k-th iteration, the position refinement process is achieved by minimizing the error term of adjacent measurements $\mathbf{p}_{n,m}^{(k)}$ and $\mathbf{p}_{n,m}^{(k-1)}$ for all the received anchor nodes as

$$\mathbf{p}_n^{(k)} := \arg \min_{\mathbf{p}_n^{(k)} \in \mathbb{R}} \sum_{m \in \xi_M} \mathbf{e}\left(\mathbf{p}_n^{(k)}, \mathbf{p}_n^{(k-1)}\right) \quad (13)$$

where ξ_M is the set of all the received anchor nodes; when there is only one anchor node in the coverage area, then $m = 1$. The error term $\mathbf{e}(\mathbf{p}_n^{(k)}, \mathbf{p}_n^{(k-1)})$ illustrates the residual error between the measured distance and the calculated distance of the position coordinates (anchor and mobile phone). The introduced term $\mathbf{p}_n^{(k-1)}$ is to improve accuracy by leveraging the highly accurate relative TOA measurements and estimated moving direction. By refining the search region of $\mathbf{p}_n^k$ within the region $\mathbb{R}$, some local optimum values outside the real region could be avoided. Such refinement could significantly minimize large errors caused by insufficient and inaccurate measurements.

Considering all the available measurements, the error constraints between the current and previous $(k-1)$-th term $\mathbf{e}(\mathbf{p}_n^{(k)}, \mathbf{p}_n^{(k-1)})$ can be written as

$$\mathbf{e}(\mathbf{p}_n^{(k)}, \mathbf{p}_n^{(k-1)}) = (\gamma_1\mathbf{e}_S + \gamma_2\mathbf{e}_D + \gamma_3\mathbf{e}_M + \gamma_4\mathbf{e}_A) \quad (14)$$

where $\mathbf{e}_S$ and $\mathbf{e}_D$ show the remaining error term of the RSS and relative TOA in (6) and (7). $\mathbf{e}_M$ and $\mathbf{e}_A$ indicate the remaining error of (8) and (9), respectively. γ_1, γ_2, γ_3 and γ_4 are weighting coefficients that control the contribution of different measurements.

By performing the gradient operation ∇ to the error residues $\mathbf{e}(\mathbf{p}_{n,m}^{(k)}, \mathbf{p}_{n,m}^{(k-1)})$ with respect to the anchor node m, the refined position can be updated via the steepest descent approach by

$$\mathbf{p}_n^{(k)} : = \hat{\mathbf{p}}_n^{(k)} + \alpha\nabla \sum_{m \in \xi_M} (\gamma_1\mathbf{e}_S + \gamma_2\mathbf{e}_D + \gamma_3\mathbf{e}_M + \gamma_4\mathbf{e}_A) \quad (15)$$

where $\alpha \in (0, 1]$ is the update step size to control the convergence rate and ξ_M is the total received anchor nodes, where the number is insufficient for trilateration. Substituting measurements into (15), $\mathbf{p}_n^{(k)}$ can be optimized and updated by leveraging the RSS and TOA ranging measurements. (15) starts with the initial coarse-grained location result $\hat{\mathbf{p}}_n^{(k)}$, and optimizes the location result by substituting the initial value of $\mathbf{p}_n^{(k)}$ by $\hat{\mathbf{p}}_n^{(k)}$.

The form of the (15) is very flexible, so all these measurements ($\mathbf{e}_S$, $\mathbf{e}_D$, $\mathbf{e}_M$, $\mathbf{e}_A$) can be opportunistically combined depending on the availability. For the ranging results, the refined position can be updated by

$$\hat{\mathbf{p}}_n := \mathbf{p}_n + \alpha\nabla e(\mathbf{p}_n) \quad (16)$$

where $\alpha \in (0, 1]$ is the update step size. The $\nabla e(\mathbf{p}_n)$ in (16) can be written as

$$\nabla e(\mathbf{p}) = 2 \sum_{m \in \xi_M} \left(||\mathbf{p}_n - \mathbf{a}_m||_2 - (\hat{\mathbf{r}}_m - \bar{\delta}_r) \right) \frac{\mathbf{p}_n - \mathbf{x}_m}{||\mathbf{p}_n - \mathbf{x}_m||_2} \tag{17}$$

$$= 2 \sum_{m \in \xi_M} \left(1 - \frac{\hat{\mathbf{r}}_m - \bar{\delta}_r}{||\mathbf{p}_n - \mathbf{x}_m||_2} \right) (\mathbf{p}_n - \mathbf{x}_m)$$

V. Location Optimization via Semidefinite Programming

When there are multiple measurements from different anchor nodes, a fine-grained location result could be obtained by aggregating all the available information together. However, we need to carefully prevent adding bad measurements into our data pool, as this would downgrade the overall performance, and may cause the estimation algorithm to diverge if not handled appropriately.

To prevent the location estimation algorithm from diverging or converging to the local optimality, the concept of relaxation onto convex sets has been proposed [24]. Without the requirement of performing the inverse operation on the Jacobian matrix in LS-based approaches, an SDP-based approach achieves better computational efficiency by leveraging existing SDP packages, which is especially important when the Jacobian matrix in LS problems is badly scaled or close to singular. Existing SDP algorithms are mainly used for sensor network localization applications, where the key principle is leveraging multiple nodes and their relative distance measurements to optimize overall location accuracy. However, our application demand is different. We need to leverage multiple data sources rather than multiple nodes. In this section, we propose an optimized real-time SDP algorithm in mobile phone location estimation. The proposed algorithm is made robust in the presence of outliers by leveraging the delay-constraint, and Huber M-estimator [25].

A. Min-max Criterion

The location estimation process has a *nonconvex* optimization problem. While semidefinite programming (SDP) techniques can be used to relax the initial *nonconvex* problem into a convex one, among existing relaxation criteria, *min-max* approximation and semidefinite relaxation can find the global minimum value without the "inside convex hull" requirement [24]. To utilize the SDP relaxation, we modify the problem formulation into $\hat{\mathbf{r}} - \delta_r = ||\mathbf{p} - \mathbf{a}_m||_2 + \mathbf{n}$. Performing a square operation on both sides leads to

$$(\hat{\mathbf{r}} - \delta_r)^T \Sigma^{-1} (\hat{\mathbf{r}} - \delta_r) = (||\mathbf{p} - \mathbf{a}_m||_2 + \mathbf{n})^2 \tag{18}$$

where the right side of (18) is $||\mathbf{p}-\mathbf{a}_m||_2^2 + 2\mathbf{n}^T||\mathbf{p}-\mathbf{a}_m||_2 + \mathbf{n}^T\mathbf{n}$. By adopting the *min-max* criterion [24], (18) can be formulated as

$$\mathbf{p} = \arg\min_{\mathbf{p}} \max_{m=1,\ldots,M} \underbrace{\left| ||\mathbf{p} - \mathbf{a}_m||_2^2 - (\hat{\mathbf{r}} - \delta_r)^T \Sigma^{-1} (\hat{\mathbf{r}} - \delta_r) \right|}_{\xi} \tag{19}$$

where the term ξ can be viewed as the residual error. (19) calculates $\mathbf{y}$ that corresponds to the minimum value of the maximum residual error. (19) remains nonconvex, but it is comfortable for the following semidefinite relaxations.

The first term in ξ can be written into a matrix form of

$$\begin{aligned} ||\mathbf{p} - \mathbf{a}_m||_2^2 &= \begin{bmatrix} \mathbf{p}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_d & -\mathbf{a}_m \\ -\mathbf{a}_m^T & \mathbf{a}_m^T \mathbf{a}_m \end{bmatrix} \begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix} \\ &= \text{trace} \left\{ \begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix} \begin{bmatrix} \mathbf{y}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_d & -\mathbf{a}_m \\ -\mathbf{a}_m^T & \mathbf{a}_m^T \mathbf{a}_m \end{bmatrix} \right\} \\ &= \text{trace} \left\{ \begin{bmatrix} \mathbf{P} & \mathbf{p} \\ \mathbf{p}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_d & -\mathbf{a}_m \\ -\mathbf{x}_m^T & \mathbf{a}_m^T \mathbf{a}_m \end{bmatrix} \right\} \end{aligned} \tag{20}$$

where $\mathbf{P} = \mathbf{p}\mathbf{p}^T$, trace$\{\cdot\}$ calculates the trace of the matrix and $\mathbf{I}_d$ is an identity matrix of order d. Following the same process in (20), the second term in ξ can be written into

$$\begin{aligned} &(\hat{\mathbf{r}} - \delta_r)^T \Sigma^{-1} (\hat{\mathbf{r}} - \delta_r) \\ &= \text{trace} \left\{ \begin{bmatrix} \mathbf{\Delta} & \delta_r \\ \delta_r^T & 1 \end{bmatrix} \begin{bmatrix} \Sigma^{-1} \mathbf{I}_d & -\Sigma^{-1} \hat{\mathbf{r}} \\ -\hat{\mathbf{r}}^T \Sigma^{-1} & \hat{\mathbf{r}}^T \Sigma^{-1} \hat{\mathbf{r}} \end{bmatrix} \right\} \end{aligned} \tag{21}$$

where $\mathbf{\Delta} = \delta_r \delta_r^T$.

B. Delay-Constraint Robust Semidefinite Programming

From (5), we know that the unknown parameter δ_r incorporates the unknown clock drift. The trend of δ_r is known as a line and the future value can be directly estimated, e.g., by linear fitting. Using such prior information, the location estimation accuracy can be further improved by substituting this pre-estimated delay δ_r as a constraint; we name this approach as delay constraint (DC).

The objective function of ξ can be converted to minimize ϵ at the constraint of an inequality expression $-\epsilon < \xi < \epsilon$, while ξ can be written as the form of (20) and (21). However, the outliers could not be ignored during location estimation. One possible solution is to apply a penalty function to the residual error rather than only using the quartic term (l_2-norm). Specifically, we still apply l_2-norm on any residual smaller than a preset threshold σ_{th}, but put a linear weight (reverts to l_1-like linear growth) on any residual larger than σ_{th}. Using l_1-norm for large errors would lower the weight for outliers and improve the robustness. We choose Huber function $\theta_{hub}(\varepsilon)$ as the penalty function [23]. This penalty function can be considered as a convex approximation of other outlier penalty functions. The constraint forms of (20) and (21) are convex, but the equality constraints of $\mathbf{P} = \mathbf{p}\mathbf{p}^T$ and $\mathbf{\Delta} = \delta_r \delta_r^T$ are nonconvex. Using semidefinite relaxation, these two equalities can be relaxed to inequality constraints of $\mathbf{P} \succeq \mathbf{p}\mathbf{p}^T$ and $\mathbf{\Delta} \succeq \delta_r \delta_r^T$, respectively. The matrix form of these two equalities is

$$\begin{bmatrix} \mathbf{P} & \mathbf{p} \\ \mathbf{p}^T & 1 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} \mathbf{\Delta} & \delta_r \\ \delta_r^T & 1 \end{bmatrix} \succeq 0 \tag{22}$$

where $\succeq$ means a positive definite (semidefinite) matrix, which is different from $\geq$.

Accordingly, the initial localization problem can be relaxed to a semidefinite programming form as

$$\begin{aligned} &\min_{\{\mathbf{p},\mathbf{P},\delta_r,\mathbf{\Delta}\}} \theta_{hub}(\epsilon) \\ \text{s.t.} \quad &-\theta_{hub}(\epsilon) < \text{trace}\left\{\begin{bmatrix}\mathbf{P} & \mathbf{p}\\ \mathbf{p}^T & 1\end{bmatrix}\begin{bmatrix}\mathbf{I}_d & -\mathbf{a}_m\\ -\mathbf{a}_m^T & \mathbf{a}_m^T\mathbf{a}_m\end{bmatrix}\right\} - \\ &\text{trace}\left\{\begin{bmatrix}\mathbf{\Delta} & \delta_r\\ \delta_r^T & 1\end{bmatrix}\begin{bmatrix}\Sigma^{-1}\mathbf{I}_d & -\Sigma^{-1}\hat{\mathbf{r}}\\ -\hat{\mathbf{r}}^T\Sigma^{-1} & \hat{\mathbf{r}}^T\Sigma^{-1}\hat{\mathbf{r}}\end{bmatrix}\right\} < \theta_{hub}(\epsilon), \\ &m = 1,\ldots,M, \\ &\begin{bmatrix}\mathbf{P} & \mathbf{p}\\ \mathbf{p}^T & 1\end{bmatrix} \succeq 0, \quad \begin{bmatrix}\mathbf{\Delta} & \delta_r\\ \delta_r^T & 1\end{bmatrix} \succeq 0 \\ &\hat{\delta}_r(1-\alpha) < \delta_r < \hat{\delta}_r(1+\alpha) \end{aligned}$$

where $\hat{\delta}_r$ is the estimated delay value based on historical data of δ_r and α is predefined and used to relax the delay-constraint (DC). The n-th mobile phone position $\mathbf{p}$ can be extracted from the optimal solution of $\{\mathbf{p}, \mathbf{P}, \delta_r, \mathbf{\Delta}\}$. This delay-constraint robust SDP problem can be solved by some standard convex optimization packages, e.g., SeDuMi. By using the steepest descent approach in (15), the estimation error can be further reduced by performing a local search above the global optimized value obtained by SDP.

VI. Evaluation

A. System Design and Evaluation

We implemented the basic beacon detection, ranging, and localization algorithm in a smartphone based on Apple's iOS 10 system. The SDP optimization part was offloaded to the server. To evaluate the power consumption of our app, we utilized the Xcode Energy Profile to get the the Graphics, Audio, Foreground App, and Total CPU activity (all of which are defined by Apple in iOS 10). As shown in Fig. 10, our app has less CPU utilization than the Google Maps. The reason we used the Google Maps iOS App as the baseline for comparison is to demonstrate: the CPU utilization of our App is acceptable for most users.

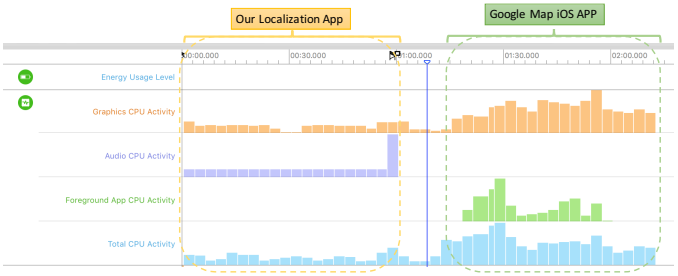


Fig. 10. The screenshot of the Xcode Energy Profile of our localization app vs. Google Map iOS App.

Besides the implementation in smartphones, we designed two versions of the anchor node for indoor localization purposes as shown in Fig. 11. As shown in Fig. 12, the power consumption of the anchor node in two different modes (network-only and network plus the acoustic beacon) is 0.071W and 0.143W under USB power, respectively. Using a 2000mAh battery as the power source, the life time of the anchor network (relying on the node with the shortest battery life) is can be calculated by $t_f = 2000mAh \times 5V/(1000 \times 0.143W \approx 70hours$.

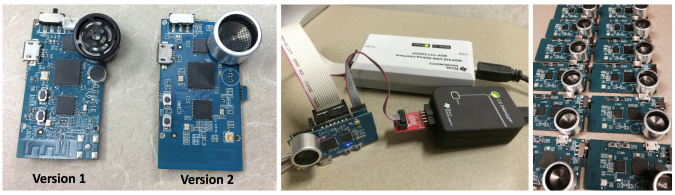


Fig. 11. Our developed hardware: (left) version 1 vs. version 2; (center) debuging via the extensition board; (right) multiple pieces.

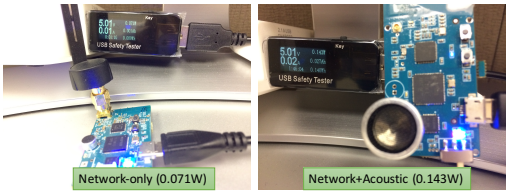


Fig. 12. The energy consumption of our anchor node in two different modes: network-only and network plus the acoustic beacon.

B. Experiment Setting

To evaluate the effectiveness of the system, we deployed 8 anchor nodes via tripod in multiple places, for example, office environment, two museum environments, and one classroom environment, as shown in Fig. 13. The maximum operating distance for one anchor node is near 20 meters. To evaluate the maximum operating distance, we utilized two large environments (aisle and hall) for the ranging test as shown in Fig. 13.

Fig. 13. Anchor deployment and experiment environment.

C. Displacement Estimation

To evaluate the performance of our proposed INS displacement estimation algorithm, we compared our proposed solution (SMS) to other approaches (e.g., AD, ATAD, WAV in [26]) in terms of *drift*. As shown in Fig. 14, we tested over 15 cases with different configurations: we put smartphones in hand (case 1-7); smartphones in pocket (case 8-11), and

smartphones in handbag (case 12-15). From the drift results, our proposed solution achieves the smallest drift for most cases.

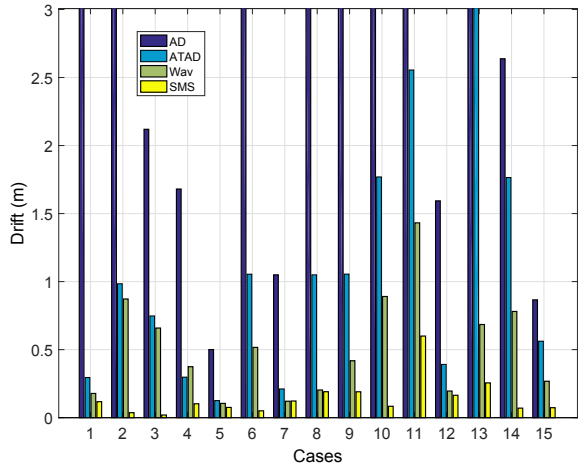
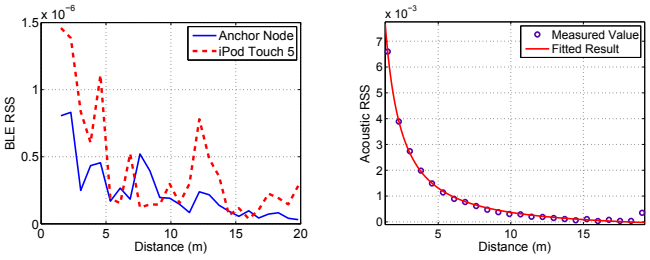
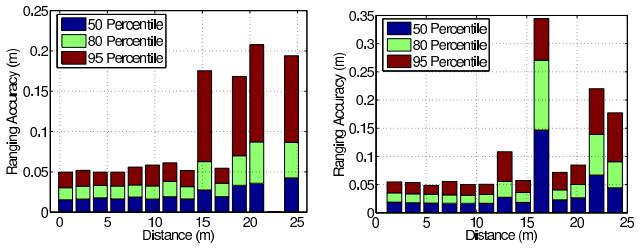


Fig. 14. Displacement test cases



(a) BLE Ranging (b) Acoustic Ranging

Fig. 15. The RSS attenuation model for ranging when using: (a) Bluetooth-Low-Energy (BLE) signal and (b) the TOA path of acoustic signal.



(a) Scenario 1 (b) Scenario 2

Fig. 16. The ranging result in two different scenarios.

D. Ranging Results

To evaluate and quantify ranging performance, we conducted a series of ranging measurements by changing the relative distance of smartphone-anchor pairs from 1.5m to 20m with the step size of 0.762m as shown in Fig. 13. Every measurement lasted more than 100 seconds to minimize randomness.

To model the RSS attenuation curve as the distance changes, we conducted experiments to measure the RSS of the acoustic signal and BLE signal by changing the distance of the anchor and the mobile phone pair. The obtained RSS attenuation curve from BLE signal and acoustic signal are shown in Fig. 15. The RSS value of the BLE signal in Fig. 15a was obtained by using our anchor node and Apple iOS device as the beacon, respectively. The acoustic RSS result in Fig. 15b shows strong monotonicity due to the fine-grained channel measurement of the TOA path, which contributes to better ranging accuracy. Using the power law function to fit the RSS attenuation data, the fitted curve can be modeled as $rss(d) = 0.0115d^{-1.199}$, where d is the distance. In real cases, the RSS ranging could be realized by measuring the current RSS value rss, and inputting the rss into the inverse function of $rss(d)$, where the distance can be calculated by $d = (rss/0.0115)^{-0.834}$.

The TOA ranging errors in different environments are shown in Fig. 16 and Fig. 17a. The 20 meter coverage with less than 10 cm ranging error is a very attractive feature for the indoor anchor node.

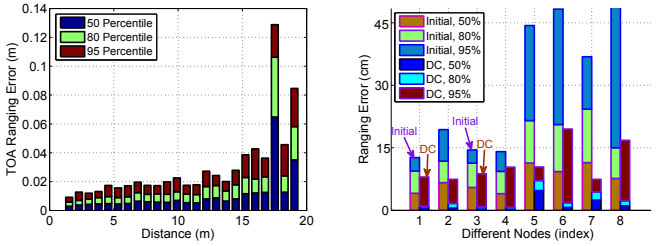
To evaluate the performance improvement achieved by the delay-constraints (DC), we compared the ranging results in Fig. 17b. Using 50%, 80% and 95% probability of the CDF results, the performance improvement is more than 3 times for most cases, especially in the large error part.

E. Angle Measurement Results

To test the directional coverage or the angle of the BLE and acoustic RSS signals, we conducted experiments to measure the radiation beam of the BLE antenna and acoustic speaker as shown in Fig. 18a and Fig. 18b, respectively. The measured RSS values in different angles (0^o to 360^o) are normalized. Compared with the almost omnidirectional radiation of the BLE beacon, acoustic beacon has high directionality, which contributes to its angle estimation features.

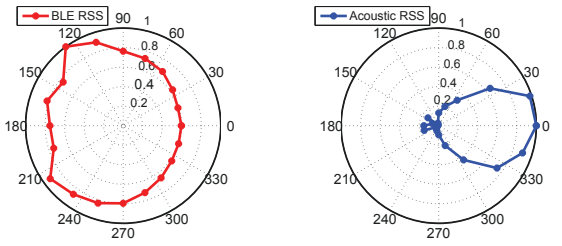
F. Improving Location Accuracy with Fewer Anchors

To evaluate the performance improvement of our proposed algorithm when the anchor nodes are insufficient for trilateration, we conducted experiments using Apple iPhone in the museum environment. We randomly selected 1 to 2 anchors from the total 8 anchors in Fig. 13 as 7 different configurations in the x-axis of Fig. 19. Fig. 19a was conducted when no anchor is blocked, i.e., the line-of-sight (LOS) case; Fig. 19b was



(a) TOA Ranging Error (b) Error Comparison

Fig. 17. (a) TOA ranging error and (b) the ranging error for two different methods.

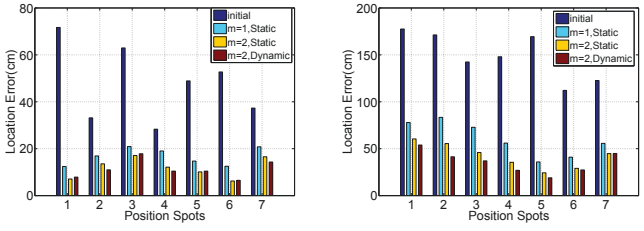


(a) BLE RSS (b) Acoustic RSS

Fig. 18. The polar distribution of normalized (a) BLE and (b) acoustic RSS.

conducted when one anchor was blocked, i.e., the non-line-of-sight (NLOS) case. To quantify the location performance and smooth out the random effect, we conducted measurements at 20 different location points for each configuration. And every location measurement contains more than 400 test results in each location point. In total, we have 8000 results for each bar in Fig. 19. The 'initial' approach is based on state-of-the-art BLE-based localization (Apple's iBeacon), where there is no acoustic beacon received. The 'm=1, static' leverages the acoustic beacon from one anchor node and the 'm=2, Dynamic' leverages the one acoustic beacon plus the motion results.

As shown in Fig. 19, when the anchor number is $m = 1, 2$, our proposed approach could significantly improved the location accuracy compared with normal BLE-based localization. Applying the dynamic part of (15) by leveraging the relative TOA distance measurement and moving direction, the performance was even improved as shown in the case of "m=2, Dynamic". From Fig. 19, we know that accuracy improvement ranged from 2 to 11 times over the initial results.



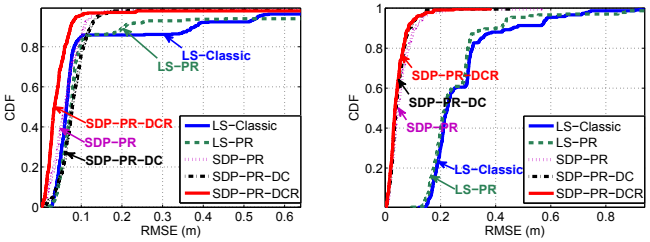
(a) LOS Results (b) NLOS Results

Fig. 19. The location error when the anchor number is $m = 0 \sim 2$ in (a) outdoor and (b) indoor environments.

G. Location Optimization via Semidefinite Programming

To evaluate and compare the performance of different localization algorithms when there are more anchor nodes available, we randomly selected 3, 4 and 5 anchors from all 8 anchors in Fig. 13, and average them together. To emulate the real application scenarios, the experimental environment was polluted with random voice sounds and other acoustic interferences.

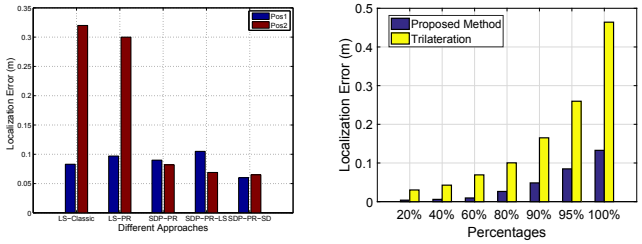
Performance Comparison. The algorithms compared are "LS-Classic" [10], "LS-PR" [14], "SDP-PR" [27], "SDP-PR-DC" and "SDP-PR-DCR". Fig. 20a and Fig. 20b show the CDF of the position error when the mobile phone is in LOS and NLOS environments, respectively. The SDP-based approaches perform better than the LS-based approaches in these two cases. By performing delay-constraint (DC) and the robust (R) approach (using Huber Estimator) during the SDP optimization, "SDP-PR-DCR" outperforms other approaches in most situations with different performance gains.



(a) LOS (b) NLOS

Fig. 20. Cumulative distribution of different algorithms when the mobile phone is in (a) LOS and (b) NLOS environment by averaging 40 cases with 3, 4 and 5 anchors.

Localization Error. For a more detailed comparison, we created Fig. 21a by listing the localization accuracy that was achieved by 80% measurements. From Fig. 21a, we know that the localization accuracy of "SDP-PR-SD" is near 6cm, i.e., 80% of position results within 6cm of error. Fig. 21b shows averaged localization results in different indoor environments (as shown in Fig. 13) with respect to different statistical probabilities. Comparing with the trilateration approach, our optimization solution achieved more than 3 times the performance improvement. Achieving such a high-accuracy position is low-cost and only relies on the normal mobile phone on the user side.

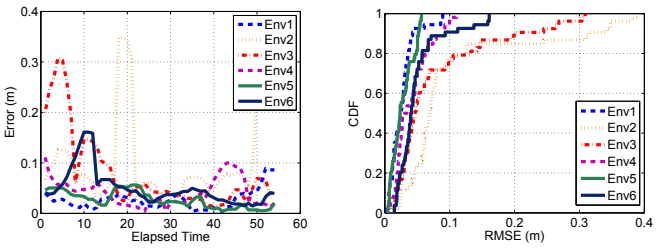


(a) Fixed Spots (b) Statistical Results

Fig. 21. The localization error of different algorithms with a) 80% probability when the mobile phone is placed near Pos1 $[5.13, 1.08]m$ and Pos2 $[5.5, 1.4]m$ and b) averaged results vs. different probabilities in CDF.

Dynamic Performance over Time. To demonstrate the dynamic localization performance over time, we calculated the localization accuracy of "SDP-PR-DCR" when its user stands still at different location spots. To support quantitative analysis, we use the time series (Fig. 22a) and CDF (Fig. 22b) of the localization error to illustrate the performance. If using 80% probability, the localization errors for all the spots are in the range between $4cm$ to $10cm$. These results are very accurate for indoor mobile phone based localization.

Location Scalability. Fig. 23 shows the achieved average location accuracy under different anchor numbers. Compared



(a) Time Series (b) CDF

Fig. 22. (a) The time series and (b) CDF of location error at different spots with 6 anchors.

with the normal trilateration case, we can achieve sub-meter level location accuracy when there are fewer than 3 anchors. This low requirement in terms of accessed anchor numbers will make the system scalable for various real world deployment conditions.

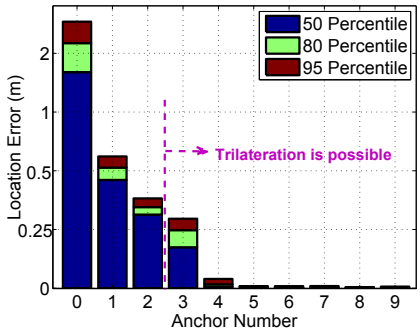
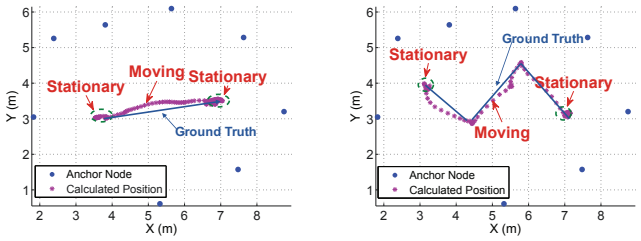


Fig. 23. The scalability of the location accuracy under different anchor coverage.

Moving Traces. Fig. 24 shows the location tracking traces of a smartphone when a user is stationary and moving. These smooth moving traces of line and curve illustrate the effectiveness of tracking a user with sufficient accuracy and update rate.



(a) Line (b) Curve

Fig. 24. Two tracking traces (a) line and (b) curve of a smartphone with 8 anchors.

VII. Related Work

Localization via anchor nodes: Conventional highly accurate infrastructure-based localization systems rely on dense anchor nodes for trilateration computations, and require special devices on the user side for ranging purposes, e.g., ultrasound [15], [10], [12], [14], [3], [28], [7]. Recent approaches relying on the high-band of the microphone sensor introduces a convenient approach for trilateration without additional hardware attachment on a user's smartphone [15], [10], [12], [14]. Liu et al. [14] utilized low-complexity anchor nodes for broadcasting unnoticeable acoustic beacon with high accuracy. However, at least three anchor nodes are needed for one location calculation with 2-D coordinates, and more nodes are needed for covering large areas, all of which inhibits wide deployment.

Localization via hybrid approaches: Leveraging multiple sensors in a mobile phone and optimizing the location accuracy via moving traces without anchor nodes are proposed in [29], [30], [31]. However, the moving traces obtained by accelerometer and compass are inaccurate and highly dependent on the prior information of the foot-step length. Authors [11], [32] proposed a localization optimization approach via peer-to-peer ranging. However, the two-way ranging process among all user-pairs is too time-consuming and inconvenient. Rajalakshmi et al. [10] proposed systems named EchoBeep and DeafBeep that fuse RF and acoustic-based techniques into a single framework. However, all these scenarios are based on fixed desktops without any mobility considerations and are not directly applicable to smartphones. Moreover, the requirement on the two-way ranging for EchoBeep and the triangular ranging for DeafBeep would limit the user numbers (as they only support one user) and introduce complex ranging protocols and long delay, which is impractical in a smartphone-based mobile system.

Proximity detection without localization: Relying on the proximity detection, one anchor node can provide location references to the users whose accuracy depends on the density of the anchor deployment. The RFID network and the recently introduced BLE network, e.g., Apple's iBeacon [19] and Qualcomm's Gimbal proximity beacons [20], and Estimote [33], are examples of using proximity detection approaches. However, proximity-based approaches are simple but inaccurate, so relying purely on the anchor density to improve accuracy is neither an efficient nor economic method.

VIII. Conclusion

We proposed location optimization approaches in mobile phones via opportunistic anchor sensing. Using the obtained coarse-grained absolute and fine-grained relative ranging information from accessible anchors, location accuracy achieved significant improvement – even with only one or two anchors. When sufficient anchors are available for trilateration, we proposed delay-constraint robust semidefinite programming to ensure robustness in the presence of ranging outliers. The achieved results show 2 to 11 times greater performance with limited anchors and sub-second delay for supporting unlimited users, and they achieve 80% accuracy of 8cm with sufficient anchors. The flexibility and accuracy of the proposed approaches provide strong incentives for service operators to deploy this low-complexity system for various location resolution demands.

REFERENCES

[1] I. Constandache, S. Gaonkar, M. Sayler, R. Choudhury, and L. Cox, "Enloc: Energy-efficient localization for mobile phones," in *INFOCOM 2009, IEEE*. IEEE, 2009, pp. 2716–2720.

[2] M. Kjærgaard, J. Langdal, T. Godsk, and T. Toftkjær, "Entracked: energy-efficient robust position tracking for mobile devices," in *Proceedings of the 7th international conference on Mobile systems, applications, and services*. ACM, 2009, pp. 221–234.

[3] S. Kumar, S. Gil, D. Katabi, and D. Rus, "Accurate indoor localization with zero start-up cost," in *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 2014, pp. 483–494.

[4] H. Wang, S. Sen, A. Elgohary, M. Farid, M. Youssef, and R. Choudhury, "No need to war-drive: unsupervised indoor localization," in *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 2012, pp. 197–210.

[5] Y. Chen, D. Lymberopoulos, J. Liu, and B. Priyantha, "Fm-based indoor localization," in *Proceedings of the 10th international conference on Mobile systems, applications, and services*. ACM, 2012, pp. 169–182.

[6] I. Constandache, R. Choudhury, and I. Rhee, "Towards mobile phone localization without war-driving," in *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010, pp. 1–9.

[7] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures," in *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 2014, pp. 617–628.

[8] K. Liu, X. Liu, and X. Li, "Acoustic ranging and communication via microphone channel," in *Global Communications Conference (GLOBECOM), 2012 IEEE*. IEEE, 2012, pp. 291–296.

[9] K. Liu, X. Liu, L. Xie, and X. Li, "Towards accurate acoustic localization on a smartphone," in *INFOCOM, 2013 Proceedings IEEE*. IEEE, 2013, pp. 495–499.

[10] R. Nandakumar, K. Chintalapudi, and V. Padmanabhan, "Centaur: locating devices in an office environment," in *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM, 2012, pp. 281–292.

[11] H. Liu, Y. Gan, J. Yang, S. Sidhom, Y. Wang, Y. Chen, and F. Ye, "Push the limit of wifi based localization for smartphones," in *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM, 2012, pp. 305–316.

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

[13] G. E. Santagati and T. Melodia, "U-wear: Software-defined ultrasonic networking for wearable devices," in *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 2015, pp. 241–256.

[14] K. Liu, X. Liu, and X. Li, "Guoguo: Enabling fine-grained indoor localization via smartphone," in *Proceedings of the 11th international conference on Mobile systems, applications, and services*. ACM, 2013.

[15] C. Peng, G. Shen, Y. Zhang, Y. Li, and K. Tan, "Beepbeep: a high accuracy acoustic ranging system using cots mobile devices," in *Proceedings of the 5th international conference on Embedded networked sensor systems*. ACM, 2007, pp. 1–14.

[16] N. Priyantha, A. Chakraborty, and H. Balakrishnan, "The cricket location-support system," in *Proceedings of the 6th annual international conference on Mobile computing and networking*. ACM, 2000, pp. 32–43.

[17] K. Liu, X. Liu, and X. Li, "Guoguo: Enabling fine-grained smartphone localization via acoustic anchors," *Mobile Computing, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.

[18] K. Liu, Q. Huang, J. Wang, X. Li, and D. O. Wu, "Improving gps service via social collaboration," in *Mobile Ad-Hoc and Sensor Systems (MASS), 2013 IEEE 10th International Conference on*. IEEE, 2013, pp. 393–401.

[19] Apple ibeacon, en.wikipedia.org/wiki/ibeacon. [Online]. Available: en.wikipedia.org/wiki/IBeacon

[20] Qualcomm gimbal proximity beacon, http://www.qualcomm.com/solutions/gimbal. [Online]. Available: http://www.qualcomm.com/solutions/gimbal

[21] Redis.io. http://redis.io/. [Online]. Available: http://redis.io/

[22] Apple cmattitude. [Online]. Available: https://developer.apple.com

[23] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ Pr, 2004.

[24] C. Meng, Z. Ding, and S. Dasgupta, "A semidefinite programming approach to source localization in wireless sensor networks," *Signal Processing Letters, IEEE*, vol. 15, pp. 253–256, 2008.

[25] G.-L. Sun and W. Guo, "Bootstrapping m-estimators for reducing errors due to non-line-of-sight (nlos) propagation," *Communications Letters, IEEE*, vol. 8, no. 8, pp. 509–510, 2004.

[26] K. Liu and X. Li, "Enabling context-aware indoor augmented reality via smartphone sensing and vision tracking," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 12, no. 1s, p. 15, 2015.

[27] P. Biswas, T. Liang, K. Toh, Y. Ye, and T. Wang, "Semidefinite programming approaches for sensor network localization with noisy distance measurements," *Automation Science and Engineering, IEEE Transactions on*, vol. 3, no. 4, pp. 360–371, 2006.

[28] Y. Zheng, G. Shen, L. Li, C. Zhao, M. Li, and F. Zhao, "Travi-navi: Self-deployable indoor navigation system," in *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 2014, pp. 471–482.

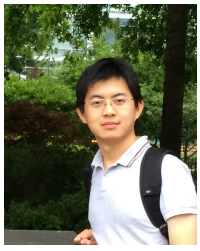
[29] I. Constandache, X. Bao, M. Azizyan, and R. Choudhury, "Did you see bob?: human localization using mobile phones," in *Proceedings of the sixteenth annual international conference on Mobile computing and networking*. ACM, 2010, pp. 149–160.

[30] Z. Yang, C. Wu, and Y. 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*. ACM, 2012, pp. 269–280.

[31] S. Sorour, Y. Lostanlen, S. Valaee, and K. Majeed, "Joint indoor localization and radio map construction with limited deployment load," *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 1031–1043, 2015.

[32] K. Liu, Q. Huang, J. Wang, X. Li, and D. Wu, "Improving gps service via social collaboration," in *Mobile Adhoc and Sensor Systems (MASS), 2013 IEEE 10th International Conference on*. IEEE, 2013.

[33] Estimote. http://estimote.com/. [Online]. Available: http://estimote.com/



Kaikai Liu is an assistant professor in the Department of Computer Engineering. His research interests include Mobile and Cyber-Physical Systems (CPS), Smart Sensing, Internet-of-Things (IoT), Software-Defined Computing and Networking. He received a Ph.D degree in Computer Engineering from UF under the direction of Dr. Xiaolin (Andy) Li. He is a recipient of the Outstanding Achievement Award at UF (four times), the Apple WWDC Student Scholarship (2013 and 2014), the Innovator Award from the Office of Technology Licensing at UF (2014), the Top Team Award at NSF I-Corps Winter Cohort (Bay area, 2015), the 2015 Gator Engineering Attribute Award for Creativity at UF, IEEE SWC 2017 Best Paper Award, IEEE SECON 2016 Best Paper Award, ACM SenSys 2016 Best Demo - Runner up, 2016 CoE Kordestani Endowed Research Professor, 2017 and 2018 CoE Research Professor Award.

Xiaolin Li is an associate professor and area chair of Computer Engineering Division in Department of Electrical and Computer Engineering at University of Florida. He received a PhD in computer engineering from Rutgers University. His research interests include cloud computing, mobile computing, big data analytics, HPC, CPS/IoT, and security & privacy. He has published over 80 peer-reviewed papers in journals and conference proceedings, 4 books, and 4 patents. His team has created large-scale software systems for clouds, big data, HPC, SDN, CPS, and mobile social networks. He is a recipient of National Science Foundation CAREER Award in 2010, Internet2 Innovative Application Award in 2013, the Top Team Award at NSF I-Corps Winter Cohort (Bay area) in 2015, and two best paper awards (ACM CAC 2013 and IEEE UbiSafe 2007). He is directing the Scalable Software Systems Laboratory (http://www.s3lab.ece.ufl.edu).