WiTraj: robust indoor motion tracking with WiFi signals

Dan Wu, Youwei Zeng, Ruiyang Gao, Shengjie Li, Yang Li, Rahul C. Shah, Hong Lu, Daqing Zhang

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing

WiTraj: Robust Indoor Motion Tracking with WiFi Signals

Dan Wu, Youwei Zeng, Ruiyang Gao, Shenjie Li, Yang Li,
Rahul C. Shah, Hong Lu, and Daqing Zhang, *Fellow, IEEE*

***Abstract*—WiFi-based device-free motion tracking systems track persons without requiring them to carry any device. Existing work has explored signal parameters such as time-of-flight (ToF), angle-of-arrival (AoA), and Doppler-frequency-shift (DFS) extracted from WiFi channel state information (CSI) to locate and track people in a room. However, they are not robust due to unreliable estimation of signal parameters. ToF and AoA estimations are not accurate for current standards-compliant WiFi devices that typically have only two antennas and limited channel bandwidth. On the other hand, DFS can be extracted relatively easily on current devices but is susceptible to the high noise level and random phase offset in CSI measurement, which results in a speed-sign-ambiguity problem and renders ambiguous walking speeds.**

This paper proposes WiTraj, a device-free indoor motion tracking system using commodity WiFi devices. WiTraj improves tracking robustness from three aspects: 1) It significantly improves DFS estimation quality by using the ratio of the CSI from two antennas of each receiver, 2) To better track human walking, it leverages multiple receivers placed at different viewing angles to capture human walking and then intelligently combines the best views to achieve a robust trajectory reconstruction, and, 3) It differentiates walking from in-place activities, which are typically interleaved in daily life, so that non-walking activities do not cause tracking errors. Experiments show that WiTraj can significantly improve tracking accuracy in typical environments compared to existing DFS-based systems. Evaluations across 9 participants and 3 different environments show that the median tracking error $< 2.5\%$ for typical room-sized trajectories.

***Index Terms*—Channel Quotient, WiFi Sensing, Channel State Information (CSI), Indoor Motion Tracking**

I. Introduction

Indoor location trajectory is very important contextual information for many applications. It is a key input for many scenarios, such as warehouse logistics, workflow automation, medical rehabilitation, home activity recognition, and computer-assisted living. While global satellite positioning systems such as GPS, Galileo, GLONASS, and Beidou positioning systems can achieve meter-level accuracy at low cost in outdoor environments, there is no successful widely deployed technology for indoor environments.

There are a variety of systems that are designed for different motion tracking applications in indoor scenarios [1], [2]. They are diverse in terms of device and infrastructure requirements, performance, and robustness. Vision-based approaches such as camera [3], infrared [4], etc. can achieve very high accuracy, however, they require expensive equipment, sophisticated calibration, suffer from field-of-view limitations, and have significant privacy concerns. Acoustic-based schemes [5] cover a limited range and do not scale to a large number of users. IMU-based methods [6] are easy to implement and deploy, but require the user to carry a device and are limited by error accumulation over time. For RF-based indoor localization systems, often times the user needs to carry an RF device. However, a number of device-free RF approaches have been developed including radar [7], [8], RFID [9], UWB [10], and WiFi localization systems [11], [12], [13], [14]. Among them, WiFi-based device-free motion-sensing systems are particularly attractive due to their low cost, ubiquitous deployment, and usability.

WiFi-based device-free motion tracking is made possible by either localizing the human target or tracking its relative displacement. Indoor positioning techniques leverage information such as CSI fingerprint collected at different locations (FIMD [13], MonoPHY [15], Pilot [16]), Angle-of-Arrival (AoA) estimation of the signal reflected off human body (MaTrack [14]), Time-of-Flight (ToF) measurement (Multi-Track [17]) to locate a person in a room. While motion tracking techniques obtain a person's displacement by calculating Path-Length-Change-Rate (PLCR) [18] and Doppler-Frequency-Shift (DFS) [19], [20] from WiFi CSI. Sometimes, information of ToF, AoA, and DFS are combined to determine the human location and track them [19], [20], [21].

Although showing promising progress, existing WiFi-based device-free motion tracking systems are still far from robust. There are three main reasons. Firstly, WiFi-based measurements are often noisy in real-world environments. Accurate AoA and ToF estimations require a large antenna array and high bandwidth, which is not available in commodity WiFi. PLCR and DFS estimations from WiFi CSI are also unreliable. Due to the uncertainty of power scaling and unsynchronized clocks between transceiver devices [22], [23], CSI obtained from commercial WiFi devices often contains amplitude noise and random phase offset, hence the inferred motion information can be noisy and unreliable.

Moreover, the accuracy of DFS is also human position and orientation dependent. The quality of DFS estimations fluctuates as the subject moves. Many existing works rely solely on such unreliable estimations as if they are accurate all the time. For instance, WiDar [18] and IndoTrack [19] use DFS information from two perpendicularly placed WiFi links to construct motion trajectory. WiDar2 [20] and mD-Track [21] use only one WiFi-link but combine AoA, ToF and DFS information. For all these approaches, there is no redundancy in the system to provide robust tracking over the entire motion trajectory.

Finally, even if continuous walking can be reconstructed

well, in realistic situations, walking is interleaved with non-walking activities such as swinging hands, vacuuming, cooking, other upper body motion, etc. Without recognizing in-place activities in real-time, DFS-based speed estimation can add substantial additional drift to the reconstructed trajectory.

WiTraj takes all these factors into consideration for tracking trajectories in a realistic indoor environment and addresses these challenges by design. In this paper, WiTraj proposes a novel DFS-based motion tracking system, which extracts DFS and reconstructs the walking trajectory by leveraging the CSI quotient of two antennas on commodity WiFi devices. This technique not only improves the SNR but also obtains a more accurate Doppler speed. An important aspect to note is that the accuracy of the Doppler speed estimation is dependent on the orientation of the person (the estimation is better when a person's chest or back is facing the WiFi devices compared to a sideways orientation) and also related to the location of the WiFi devices. Based on our study, the estimation is more accurate when the angle between the person's walking direction and the line-of-sight of the transceiver devices is greater than 30°. Therefore, we propose using three WiFi links to track human motion in a room, which ensures that at least two links can provide reliable Doppler speed estimations. Through the selection and fusion of the two best views, trajectory reconstruction is robust to user posture, orientation, and walking direction. Moreover, our system is able to separate in-place activities from walking motion, ensuring the reconstruction of the walking trajectory is reliable and accurate.

To summarize, this work makes the following contributions:

- We propose a DFS acquisition method leveraging a CSI-quotient model. It uses the CSI of two antennas on a WiFi receiver to improve the SNR of Doppler speed estimation and resolve the speed sign ambiguity problem (confusion about whether the motion is towards or away from the WiFi link).
- Leveraging multiple WiFi device links (three or more), we ensure there are at least two receiving devices (views) that get a good estimation of Doppler speed at any time, regardless of user location and orientation, so that the accuracy of a person's moving speed and direction can be guaranteed.
- We can distinguish walking from other in-place activities using cumulative displacement information, so that these activities do not add errors to the walking trajectory reconstruction and better support practical motion tracking in real-world applications.
- We implement WiTraj on commodity WiFi devices and conduct large-scale experiments for evaluation demonstrating high accuracy of trajectory reconstruction. Evaluation of 9 participants in three different environments shows that the median tracking error is < 0.3 meters or $< 2.5\%$ for a typical room-sized trajectory.

The remainder of the paper is organized as follows. We propose a CSI-quotient-based DFS extraction method and compare it with state-of-the-art in Sec. II. After that, we leverage multi-view to ensure that at least two reliable DFS estimations are always available in Sec. III, then we show in-place activities can be distinguished from walking in Sec. IV. We provide details of the implementation in Sec. V and evaluate the performance in various settings in Sec. VI. We discuss the limitations of the approach and future opportunities in Sec. VII and survey the related work in Sec. VIII. Finally, we conclude the paper in Sec. IX.

II. Extracting DFS using CSI Quotient

Doppler frequency shift (DFS) extraction from WiFi CSI would be easier if CSI has no phase offset. However, real CSI readings from commodity WiFi devices contain both phase and amplitude noise [24], [25]. Consequently, the existing DFS estimation method often suffers from the ambiguous sign of speed direction [26], [19], [20]. In this section, we propose a novel DFS extraction method based on the CSI-quotient model and show its superiority in terms of better SNR and unambiguous speed estimation compared to existing methods.

A. Limitation of existing DFS extraction methods

In WiFi-based device-free sensing scenarios, the channel state can be expressed as the superposition of static signal components (combination of LoS and static multipath reflections) and dynamic signal components that are reflected off the moving human body [27], [28]. Considering the CSI noise in amplitude [27] and phase [22], [23], [29], in the case of one moving object, the received CSI can be expressed as:

$$CSI(f,t) = A_{\text{noise}}(f,t)e^{-j\theta_{\text{offset}}(f,t)}(H_s(f)+H_d(f,t)) \quad (1)$$

where A_{noise} is the power amplifying uncertainty in amplitude, θ_{offset} is the random phase offset, $H_s(f)$ is the static phasor component and $H_d(f,t)$ is the dynamic phasor component.

Existing research takes the conjugate multiplication of two CSI signals from two antennas of a WiFi card as the input for Doppler velocity extraction [26], [19]. The conjugate multiplication of two CSI signals takes the form of Eq. 2. It contains four terms. Among them, the term ① is time-invariant. The cross-term ④ is orders weaker than the term ② and term ③, so we can safely omit it [26]. The other two terms ② and ③ are time-variant. Term ③ contains the Doppler shift of interest, while term ② contains an arithmetically opposite number which may produce ambiguous Doppler speed estimation. By design, it is impossible to separate the term ③ out directly.

One drawback of conjugate-multiplication-based DFS extraction methods is the amplified noises. From Eq. 2 we can see the conjugate multiplication operation eliminates the CSI phase offset, but further amplifies the CSI amplitude noise.

Another drawback of conjugate-multiplication-based DFS extraction methods is that they have to use heuristics to deal with an ambiguous speed problem. WiDance [26] selects antennas and assigns the order of antennas carefully in conjugated multiplication to ensure the right estimation. Doppler-MUSIC [19] amplifies the CSI amplitude of one antenna and reduces the CSI amplitude of the other antenna to alleviate the ambiguity. Although the above methods work to some extent, the ambiguity is still a challenge in practice. The ambiguous estimates come from the requirement of carefully comparing

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing

$$
\begin{aligned}
H_{cm}(f,t) &= CSI_1(f,t)\overline{CSI_2(f,t)} \\
&= \left(A_{\text{noise}}(f,t)e^{-j\theta_{\text{offset}}(f,t)}(H_{s1}(f)+H_{d1}(f,t))\right)\overline{\left(A_{\text{noise}}(f,t)e^{-j\theta_{\text{offset}}(f,t)}(H_{s2}(f)+H_{d2}(f,t))\right)} \\
&= \left(A_{\text{noise}}(f,t)e^{-j\theta_{\text{offset}}(f,t)}(H_{s1}(f)+H_{d1}(f,t))\right)\left(A_{\text{noise}}(f,t)e^{j\theta_{\text{offset}}(f,t)}\overline{(H_{s2}(f)+H_{d2}(f,t))}\right) \\
&= A_{\text{noise}}(f,t)^2\left(H_{s1}(f)+H_{d1}(f,t)\right)\left(\overline{H_{s2}(f)}+\overline{H_{d2}(f,t)}\right) \\
&= A_{\text{noise}}(f,t)^2(\underbrace{H_{s1}(f)\overline{H_{s2}(f)}}_{①}+\underbrace{H_{s1}(f)\overline{H_{d2}(f,t)}}_{②}+\underbrace{\overline{H_{s2}(f)}H_{d1}(f,t)}_{③}+\underbrace{H_{d1}(f,t)\overline{H_{d2}(f,t)}}_{④}) \\
&\approx A_{\text{noise}}(f,t)^2\left(H_{s1}(f)\overline{H_{s2}(f)}+H_{s1}(f)\overline{H_{d2}(f,t)}+\overline{H_{s2}(f)}H_{d1}(f,t)\right)
\end{aligned}
\tag{2}
$$

the two CSIs and arranging the order in the conjugate multiplication. Unfortunately, the correct order may not be consistent over time especially while the subject is moving.

B. Using CSI quotient to estimate DFS

We propose to estimate DFS from CSI channel quotient with a transform model known as the CSI-quotient model or CSI-ratio model [30], [31]. The channel quotient has shown its success in sensing micro-scale motions such as respiration monitoring [31], [32] and finger motion tracking [30]. In this section, we show it also has great potential in modeling macro-scale indoor motion and resolving the limitation of previous conjugate-multiplication-based Doppler acquisition methods.

The CSI-quotient model takes the ratio of CSI readings of two antennas on the same receiver as a new base signal. It has the form of:

$$H_q(f,t) = \frac{CSI_1(f,t)}{CSI_2(f,t)} = \frac{H_{s1}(f)+H_{d1}(f,t)}{H_{s2}(f)+H_{d2}(f,t)} \tag{3}$$

where $H_{s(1,2)}(f)$ are the static phasor components and $H_{d(1,2)}(f,t)$ are the dynamic phasor components for the two antennas of the same receiver, respectively.

In contrast to prior work that requires dedicated steps to reduce the CSI noise [27] and sanitize the phase [23], [29], [24], [25], CSI quotient removes CSI noise efficiently by just dividing two raw CSI readings from different receiver antennas. As the antennas on a receiver share the same RF chain and clock, the amplitude noise $A_{\text{noise}}(f,t)$ and random phase offsets $\theta_{\text{offset}}(f,t)$ for each of the antennas are almost identical. Therefore the CSI quotient as shown in Eq. 3 cancels both amplitude noise and phase offset, providing a much higher SNR than conjugate-multiplication-based methods.

The rationale that the CSI quotient model can be used to extract DFS stems from its periodic characteristic and the correlation to the reflection path length change. Under the assumption of only one dominant dynamic reflection path, the CSI quotient of two antennas is a Möbius transform [30]:

$$H_q(f,t) = \frac{H_{s1}(f)+A_1(f)e^{-j2\pi\frac{d_{\text{diff}}}{\lambda}}e^{-j2\pi\frac{d_2(t)}{\lambda}}}{H_{s2}(f)+A_2(f)e^{-j2\pi\frac{d_2(t)}{\lambda}}} = \frac{az+b}{cz+d} \tag{4}$$

where $A_1(f,t)$ is the attenuation of dynamic component $H_{d1}(f,t)$ for antenna 1 of the receiver, $A_2(f,t)$, $e^{-j2\pi\frac{d_2(t)}{\lambda}}$ and $d_2(t)$ are the attenuation, the phase shift and the path length of dynamic component $H_{d2}(f,t)$ for antenna 2, and d_{diff} is the difference in length of the dynamic signal propagation path between antenna 1 and antenna 2. If we define $z = e^{-j2\pi\frac{d_2(t)}{\lambda}}$, then it is a unit complex variable with its phase representing the change in length of the reflected path, and the coefficients a, b, c, d are constant complex numbers. Obviously, as the reflection path length changes one wavelength, z rotates for a circle. According to the Möbius transform, the signal of the channel quotient rotates for a perfect circle on the complex plane.

The Doppler frequency shift of a moving object is defined as:

$$f_D = -\frac{1}{\lambda}\frac{d\mathrm{d(t)}}{d\mathrm{t}} \tag{5}$$

Putting together Eq. 4 and Eq. 5, we can see that the Doppler speed measured at the dynamic reflection path actually equals the rotations per second for the CSI quotient on the complex plane, with the rotation direction indicating the sign of the Doppler speed.

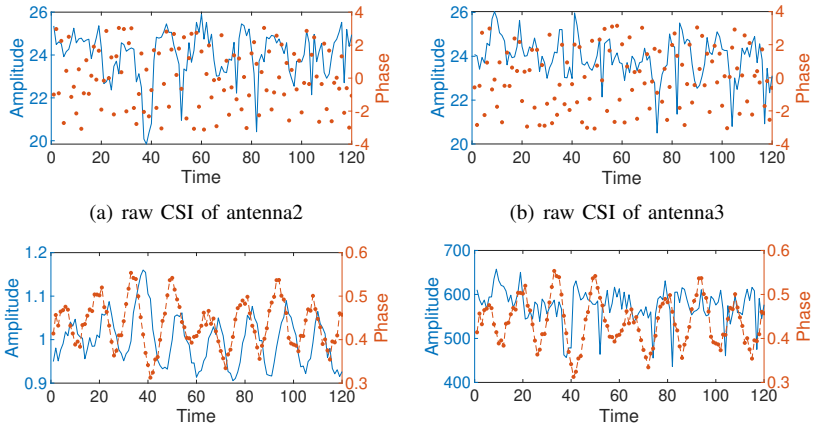
C. CSI Quotient vs. Existing DFS Extraction Methods

One benefit of the CSI-quotient model compared to conjugate multiplication is better noise cancellation capability. CSI-quotient cancels both the phase offset and the amplitude noise effectively, thus providing a better signal-to-noise ratio (SNR) than the conjugate-multiplication-based methods.

Fig. 1 shows a typical raw CSI signal of human walking and the CSI quotient counterpart. Fig. 1(a) and Fig. 1(b) show the raw CSI of two antennas on a receiver; as can be seen, the phase of CSI samples are randomly distributed in $[-\pi, \pi]$ and we observe large power amplification noise in the amplitude. Fig. 1(c) shows that by dividing the two CSI signals, most of the noise in both amplitude and phase gets removed. On the other hand, Fig. 1(d) shows the conjugate multiplication of the two CSI signals. It removes phase offset just as effectively but is not capable of reducing noise in the amplitude.

Another advantage that the CSI-quotient model offers is unambiguous speed estimation. The Möbius transformation keeps the rotation direction the same as the CSI, as long as the denominator satisfies a simple rule that the static signal component is stronger than the dynamic component. This constraint generally holds in practice, as the signals propagating along the LoS path and environmental reflection paths are

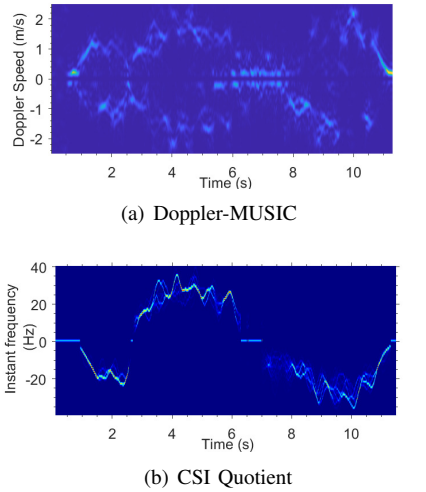
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing



(a) raw CSI of antenna2 (b) raw CSI of antenna3

(c) CSI quotient (d) CSI conjugate multiplication

Fig. 1. The comparison of the raw CSI of two antennas, the CSI quotient, and the CSI conjugate multiplication signals of the same human walking. The blue line is amplitude, and the red line is phase. We can see the amplitude from CSI quotient model is much cleaner. All the above signals are not filtered.



(a) Doppler-MUSIC

(b) CSI Quotient

Fig. 2. The instant speed spectrum of the same walking data. The spectrum of Doppler-MUSIC algorithm contains noticeable speed ambiguity, while the CSI quotient removes this ambiguity and renders speed of better SNR.

much stronger than that reflected off human bodies [27], [19]. Note that there is no requirement for the numerator. This means that given the CSIs of different antennas on a WiFi card, all we need to do is choose an appropriate CSI that meets the above condition as the denominator and any other CSI can be chosen as the numerator to ensure unambiguous DFS estimation.

Fig. 2 shows the extracted Doppler frequency shift information of a person walking along a circular route. Compared with the result of the conjugate-multiplication-based Doppler-MUSIC method [19], the DFS estimated by the CSI quotient model is clear and has no ambiguity.

III. Ensuring Reliable DFS via Multiple-view

The CSI quotient model extracts DFS under the assumption of only one dynamic reflection path (i.e. a single moving object). For macro-scale motions such as walking, this assumption does not always hold. In practice, this means that a single WiFi link can not guarantee reliable DFS estimation at all times. However, we show in this section that by combining multiple views from different WiFi links, we can reliably estimate the DFS and track motion robustly.

A. Drawbacks of Single-view DFS Estimation

Existing motion tracking systems [19], [20], as well as the basic CSI quotient model, treat the human body as a single point that reflects a single dynamic path. However, due to the motion of multiple body parts, this assumption is not accurate in reality. This is clearly demonstrated by an empirical study.

1) Experimental Settings: We use a single pair of WiFi transmitter and receiver and fix the locations of the devices while letting a person walk along several straight lines as illustrated in Fig. 3. The WiFi transmitter and receiver devices are mounted on tripods and separated by 4 meters. Each WiFi device consists of an Intel 5300 wireless card to collect the CSI data. We mark 16 intersections in the sensing area, arranged in a 4×4 grid that marks the points separated by 1 meter. The path *Trace 1* connects points (1,1) and (1,4), while the path *Trace 2* connects (1,1) and (4,2). As can be seen, the paths have different angles concerning the LOS link between the transceivers. By visualizing the Fresnel zones around WiFi transceivers, we can see that the participant walks almost perpendicular to the Fresnel zone boundaries in *trace 1*, while the participant walks at a relatively small angle with respect to the Fresnel zone boundaries in *trace 2*.

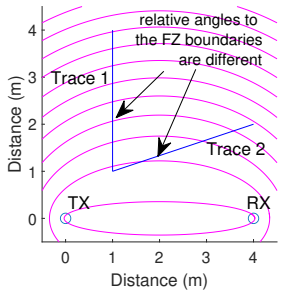
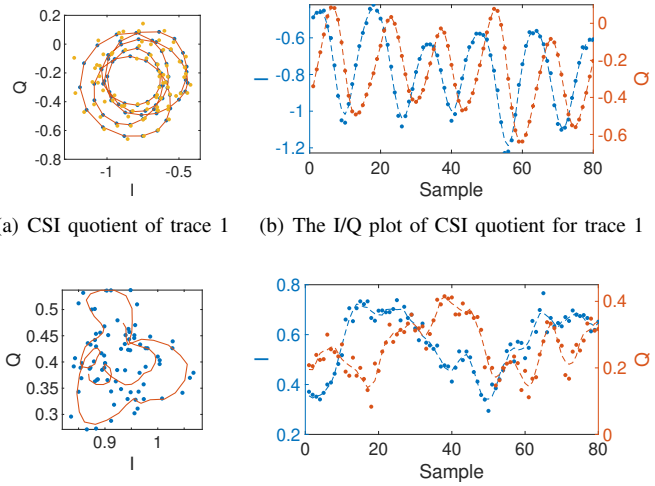


Fig. 3. A participant walks along two traces in the Fresnel Zone (FZ) of a pair of transceivers. The traces have different relative angles to the FZ boundaries. We draw an ellipse every 15 FZs for a clearer view.

2) Observations: Fig. 4 shows the *signal* of the computed CSI quotient for the two walking traces for a single sub-carrier. It is obvious that the signal has a regular pattern for *trace 1*, while it is fairly noisy for *trace 2*. Clearly, it is much simpler to determine the number of Fresnel zone boundaries crossed in an ideal case like *trace 1*, but it is much more complex in the case of *trace 2*. Considering the walking posture, it is clear that the participant walks with their chest/back facing the WiFi devices in *trace 1*. Hence, in this case, the torso reflection is stronger than the reflection from other moving body parts such as arms and legs, thus it is closer to the assumption of single dynamic

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing



(a) CSI quotient of trace 1 (b) The I/Q plot of CSI quotient for trace 1

(c) CSI quotient of trace 2 (d) The I/Q plot of CSI quotient for trace 2

Fig. 4. the comparison of the CSI quotient signal for the two walking traces. The dots are the raw CSI quotient signals, and the curves are the smoothed ones.

reflection. Conversely in *trace 2*, the participant's arms and legs alternately move toward the WiFi transceivers, and the reflective surface of the torso may be blocked by the arms and shoulders alternately so the multipath superposition is much more complex. In this case, the estimation of the change of distance using the assumption of a single reflection point can lead to large errors. From this, it is easy to deduce that the smaller the trajectory angle with respect to the elliptical boundary of the Fresnel zone, the more irregular the pattern in the CSI signal, which leads to a correspondingly larger estimation error.

In order to further investigate the distribution of errors, we collected data from multiple walking trajectories along with different angles and summarized the error distribution of DFS estimation. We use lines of pairwise connection of 16 points in Fig. 3, such as $(1,1) \rightarrow (4,1)$, $(1,1) \rightarrow (4,2)$, $(1,1) \rightarrow (4,3)$, $(4,2) \rightarrow (1,3)$, $(4,2) \rightarrow (1,4)$. Then we asked five participants to walk along each path ten times, and collected a total of 400 data traces. The angle between the walking trajectory and the Fresnel zone is calculated as the average of the angles that the motion direction makes with the multiple Fresnel elliptical boundaries during the entire trajectory. The ground-truth of the distance change for a reflection path is calculated based on the start and end locations (d_{ref}). By definition, the change in the length of the reflection path should be reflected in the number of full rotation circles in the CSI quotient signal on the complex plane. So from the number of rotation circles in the signal, we compute the estimated walking displacement (d_{est}). The error is calculated as $(d_{\text{est}} - d_{\text{ref}})/d_{\text{ref}}$. Fig. 5 shows the scatter and box plots of the distribution of errors of the displacement estimates at different angles. For walking angles between 30 to 90 degrees, the error is mostly within 5% and the total standard deviation is 0.079. On the other hand, for walking angles less than 30 degrees, the error range is much wider, sometimes exceeding 150%. This confirms that the computed displacement is much more accurate when

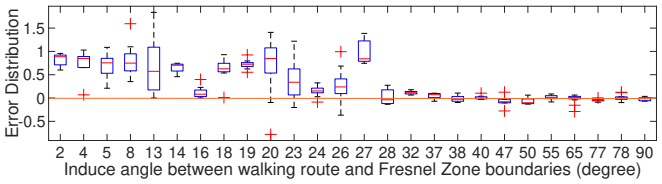


Fig. 5. The error distribution of motion displacement estimation for different trace angles relative to the Fresnel Zone boundaries. The error is closer to zero, i.e. the estimation is more accurate as the trace angle is greater.

the walking trajectory is perpendicular to the Fresnel zone boundaries.

From the above observations, we see that the accuracy of the DFS estimation varies depending on the location of the person and the angle at which they are walking relative to a WiFi link, and can conclude that a single WiFi link cannot guarantee the motion estimation accuracy.

B. Improving DFS Estimation with Multi-View

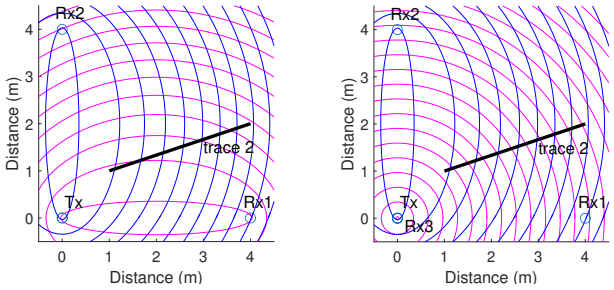
Knowing the limitation of a single WiFi link, we explore using multiple WiFi receiving devices placed at different locations to provide multiple viewing angles. At every point of the walking path, we expect that the additional DFS information extracted from these views can capture a better signal for reliable motion estimation as shown below.

1) A Two-view Setting is Not Robust Enough: A single pair of WiFi transceiver devices only give one-dimensional DFS information. This can naturally be extended to a two-view setting which will provide the minimum information along with two directions such that the walking speed on a 2D plane can be determined. This is shown in Fig. 6(a) which illustrates the default setting for many human tracking systems [28], [18], [19]. These systems place one transmitter and two receivers (marked Rx1, Rx2) in three vertexes of the square sensing area.

As there are only two receiving devices, the reconstruction of the walking trajectory has no redundant information from other views. It requires both the DFS estimations to be accurate for a good trajectory reconstruction. However, as observed in Sec. III-A, the accuracy of DFS estimation of a person depends largely on the angle between the walking route and the elliptical boundary of Fresnel zones. Any inaccurate DFS estimations cause unwanted drifts on the reconstructed trajectory. As an example, consider a subject walking along a straight line *trace 2*. This route has a large angle to the Fresnel zones of Tx-Rx2, but it has a small angle to the Fresnel zones of Tx-Rx1. This means that the estimated target displacement on the Rx2 device will be more accurate, while Rx1 will not be that accurate.

2) Robust Trajectory Estimation with Three-views: To ensure robust motion tracking, at least two reliable DFS estimations are required regardless of human position and walking direction. Hence our insight is that we need to obtain redundant motion estimates by observing from additional viewing-angles, such that we have a good chance to have at least two accurate estimates at any time to reconstruct a robust motion trajectory.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing



(a) Traditional setting (b) Multi-view setting

Fig. 6. The multi-view setting for 2D motion tracking. Different views estimate displacement with different errors for the same walking trace. By selecting proper views where the walking trace crosses the Fresnel zones with larger angles, the estimation errors can be reduced significantly.

This can be achieved by adding a third receiving device (Rx3) as in Fig. 6(b). The new transceiver pair Tx-Rx3 forms a new set of Fresnel zones. For the case of *trace 2*, the trajectory now has a large included angle between the walking route and the Fresnel zones of Rx3. As a result, displacement estimation on Rx2 and Rx3 is more reliable and the reconstructed motion trajectory using information from these two devices will be more reliable than other combinations.

IV. Differentiating In-place Activity

As shown in the previous section, we can use multiview to accurately reconstruct the walking trajectory of a person. However, in daily life, a person does not keep walking continuously but stops to perform activities in-between. These activities (e.g., using a computer, brushing teeth, or cooking) can accumulate errors in the trajectory estimation over time. To prevent this, WiTraj reconstructs the trajectory only when the target subject is walking by distinguishing it from the motion of other in-place activities.

A. Differentiating Walking Motion

The intuition of differentiating walking from non-walking motion lies in the fact that the target's position during in-place activities only changes irregularly in a small range. On the other hand, when the target is walking, his/her motion will be smooth and consistent in a certain direction which will result in a large and regular position change. Based on this understanding, WiTraj calculates the 2-D motion displacement and obtains the position sequence of the sensing target. Next, WiTraj computes a circle to cover all the estimated positions in a short period, with the diameter indicating the range of the position changes. Fig. 7 shows the diameter of the covering circle for a series of daily activities. As we can see, the circle's diameter corresponding to walking is significantly larger than other in-place activities. Hence it is possible to differentiate walking from in-place activities by comparing with a threshold diameter. To reconstruct the trajectory, we identify all the walking segments in an activity session and splice them together to record the target's footprint.

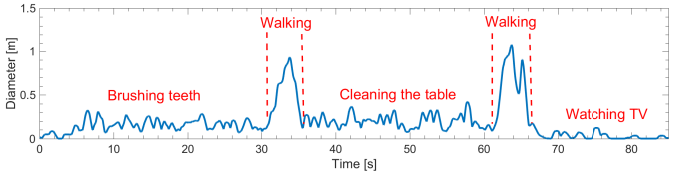


Fig. 7. The change in the diameter of the position covering circle for a series of daily activities.

B. Testing results

To validate this idea, we collect data of typical activities in an apartment and compare them with walking. We use one transmitter and three receivers as described in Sec. VI-A. Specifically, we put Tx and Rx3 at the coordinate of (0,0), Rx1 at (2,0), and Rx2 at (0,3). We chose two locations to test the in-place activities, namely (1,2) and (2,3). The participant conducts four in-place activities, including turning around, doing jumping jacks, waving hands, sitting down, and standing up. Each activity is performed ten times by the participant at each testing location. These activities are compared with walking in terms of the diameter of the trajectory covering circle as described above. We test walking at three speeds, i.e, strolling, normal walking, and brisk walking, and estimate the 2D displacement and maximum displacement in a time window of 1 second for each activity.

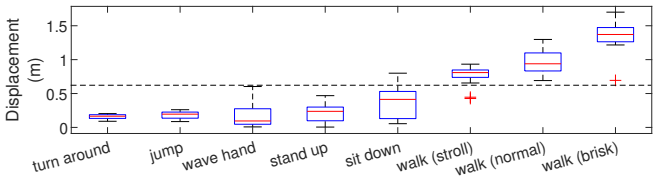


Fig. 8. The non-walking activities accumulate less displacement over a time-window of 1 second compare with walking, thus the two can be differentiated.

Fig. 8 shows the maximum displacement of all the in-place activities and walking in a 1-second-window. We can see walking has a much larger displacement than most in-place activities, even for a slow walking speed. Activities such as sitting down and standing up have larger displacements than other in-place activities, but are still clearly smaller than walking. Hence, to distinguish walking from non-walking, we choose 0.7 meters as the empirical threshold in our test and build the confusion matrix, as illustrated in Fig. 14. Most non-walking activities and walking can be differentiated well, except for a few cases of confusion between sitting down and slow walking.

V. WiTraj System Design

In previous sections, we showed how the CSI-quotient model and the multi-view setting improve the DFS estimation quality, as well as the basic idea on differentiating walking from in-place activities. Putting all these together, we now describe the full motion tracking system.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing

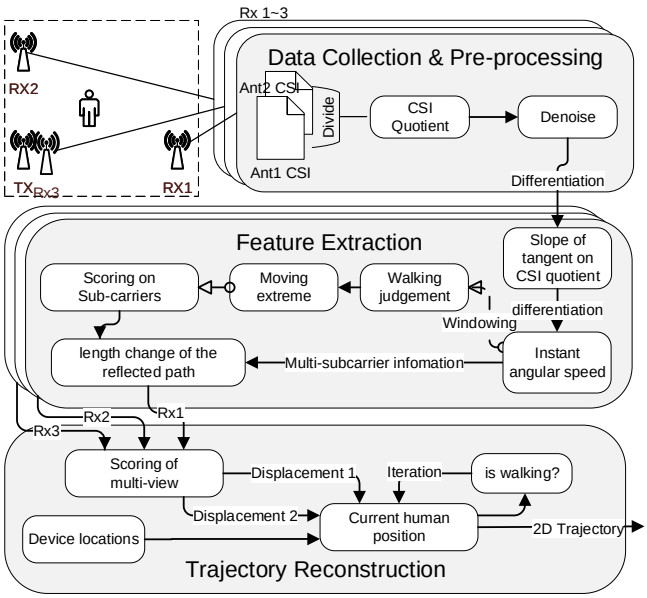
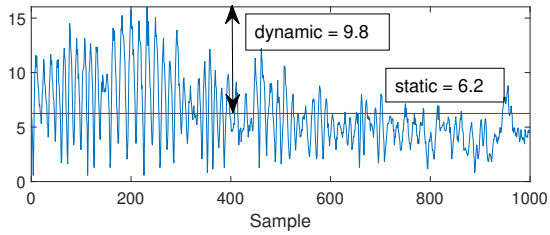
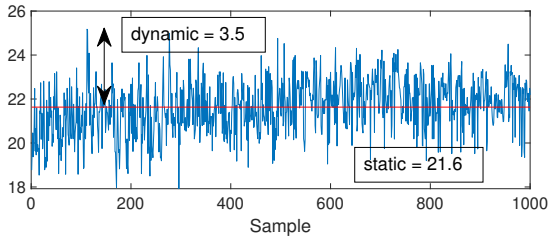


Fig. 9. WiTraj system overview



(a) CSI amplitude of antenna 1



(b) CSI amplitude of antenna 2

Fig. 10. The denominator of CSI quotient is decided by the ratio of the power of static and dynamic components. It can be calculated as the average and the variation of the CSI amplitude. In this figure, we can safely use the CSI of antenna 2 as the denominator.

A. System Overview

The WiTraj system[1] is built with commodity WiFi hardware. It includes four WiFi devices, i.e. a WiFi access point (AP) and three WiFi clients (as receivers) equipped with Intel 5300 wireless cards. Each receiver is equipped with at least two external omnidirectional antennas. The placement of the devices is shown in Fig. 6(b).

The WiTraj software implementation consists of three modules: data collection & pre-processing, reflection-path length change estimation, and trajectory reconstruction. The three modules are sequentially connected. The data collection module gathers CSI information from the WiFi receivers and computes the CSI quotient data streams for each receiver. The reflection-path length change estimation module extracts length changes of the reflection path in Fresnel Zones from the CSI quotient streams. And finally, the trajectory reconstruction module estimates the walking trajectory based on the calculated reflection path length change information. The complete system information flow is illustrated in Fig. 9. We built a real-time prototype system in MATLAB that collects the CSI streams over TCP/IP sockets and reconstructs motion trajectories.

B. CSI Data Collection

We collect CSI data from all the receivers with each of them having at least two antennas. In the case of the Intel 5300 wireless cards, CSI data contains information from 30 sub-carriers. Since the 5300 card supports three antennas, each CSI data frame is a 3×30 complex matrix. We send this CSI data to a dedicated computer via TCP/IP socket for further processing.

[1]The source code is available at https://github.com/Soccerene/WiTraj

C. CSI Quotient Calculation

We take an element-wise division of the CSI data of two antennas on the same receiver to get the channel quotient. To ensure unambiguous speed mapping, the denominator needs to be the CSI value for which the static signal component is stronger than the dynamic signal component, as explained in [30]. To determine this, the mean power of the static signal component is calculated as the average raw CSI amplitude of each antenna, and the maximum possible power of the dynamic signal component is calculated as the difference between the maximum and average raw CSI amplitude. Fig. 10 illustrates the CSI amplitude of two antennas. The CSI of antenna 2 has a static to dynamic power ratio greater than 1 in all the data segments, so we can safely use it as the denominator. As a comparison, the CSI of antenna 1 has static to dynamic power ratio of less than 1 in the first half of the data. If we use it as the denominator of the CSI quotient, then the speed estimations in the first half of the data will have negative speeds. Please note that we don't need to consider the severe noise in antenna 2, as the noise will be canceled completely after the division.

D. Data Denoising

Data preprocessing involves data denoising and segmentation. Although the division of CSI over two antennas greatly suppresses CSI noise, the CSI quotient data still needs to be smoothed to clearly reveal the circular shape for phase change extraction. We use a Savitzky-Golay filter to smooth out the CSI quotient data. Savitzky-Golay smoothing fits successive subsets of data points with a low degree polynomial using the method of linear least square. It introduces less distortion than band-pass filters in [28] and is also efficient in smoothing out

complex data. We show an example of denoised CSI quotient data in Fig. 4.

E. Computing reflected path length change on a Single-view

1) Calculate reflected path length change from CSI Quotient: WiTraj uses the CSI quotient for motion displacement sensing. First, we need to compute the change in the length of the reflected path. It can be derived from DFS information f_D, which has been discussed in Sec. II. Mathematically, we have the relationship:

$$\triangle d = f_D \triangle t = f_D / fn = \frac{\lambda \triangle \rho}{2\pi} \tag{6}$$

where $\triangle d$ is the distance change of the reflected path, fn represents the sampling rate of CSI, and $\triangle \rho$ represents the phase change of the rotating circles in the CSI quotient signal between the sampling intervals. For the sake of clarity, please note that the actual distance walked by a person will be different from the change in the reflected path, and for now we are referring to the latter.

To calculate the path length change, we need to measure the cumulative phase change $\triangle \rho$ of the circle rotation. An easy way to measure phase changes is to find the center of a circle and extract the rotation angle of its radius. This method uses a fixed-length sliding window to average the complex CSI-quotient data, then subtracts it from the CSI-quotient to obtain the dynamic phasor vector and then extracts its rotating angle, as done in [33]. However, there are two problems with this method. The first drawback is that using a fixed-length sliding window does not fit the changing cycle lengths. Even when walking in a straight line, the instantaneous walking speed of a person is always changing, so the cycle length on the signal also varies. Therefore taking the average may not be a good way to pinpoint the center of the circles. The second drawback is that the real signal is always very different from the ideal model. Because of the change of the human reflective surface during walking [34], the rotating circle of the CSI-quotient on the complex plane will gradually change the position of its center and the radius size. These two factors make it difficult to obtain the phase information of the dynamic vector accurately using traditional methods.

In WiTraj, we use a similar method as in [30] to calculate the phase change $\triangle \rho$. It takes advantage of the simple fact that the tangent on the circle is always perpendicular to the radius at that point, so the derivative of the tangent slope is equivalent to the derivative of the phase of the dynamic phasor component over time. The tangent $\perp \rho$ on the circle can be calculated by subtracting two adjacent sampling points on the circle, while the tangent slope change $\triangle \rho$ can be calculated by subtracting the slope of two adjacent tangents. Compared to the Doppler-MUSIC algorithm [19], it is computationally efficient because the calculation involves only two subtractions. After the phase change $\triangle \rho$ is obtained, the distance change of the reflection path can be easily calculated based on Equ. 6.

The above analysis is based on the ideal CSI quotient model. However, there is a difference between real and ideal signals. As seen in the previous empirical observations in Sec. III-A, the signal quality of each subcarrier or antenna pair is different. We need to discard segments of data that may produce the wrong estimation and pick up the best signals from the redundant data streams of subcarriers segment-by-segment to estimate the path changes. As the 'perfect' circle provides a better SNR, it is more likely to produce correct path length change estimates.

2) Picking and Combining Motion Information from Sub-carriers: All the sub-carriers of the CSI-quotient provide similar motion information but with varying signal quality. High-quality signals result in good motion estimates, while poor-quality signals do not. Fig. 11(a) and Fig. 11(b) show the comparison of CSI quotient signal of two sub-carriers. During the first half of the time in the plot, we can see that sub-carrier 1 is of low quality, while sub-carrier 29 has a very clear pattern. For both the sub-carriers, the quality of the signal changes over time. It is often the case that the signal quality deteriorates in one sub-carrier, while the signal quality in another sub-carrier gradually improves. Hence by picking good quality information from different sub-carriers and combining them, we can achieve a high-quality signal for the entire time period. As the quality changes for different sub-carriers during walking, we adapt to it dynamically.

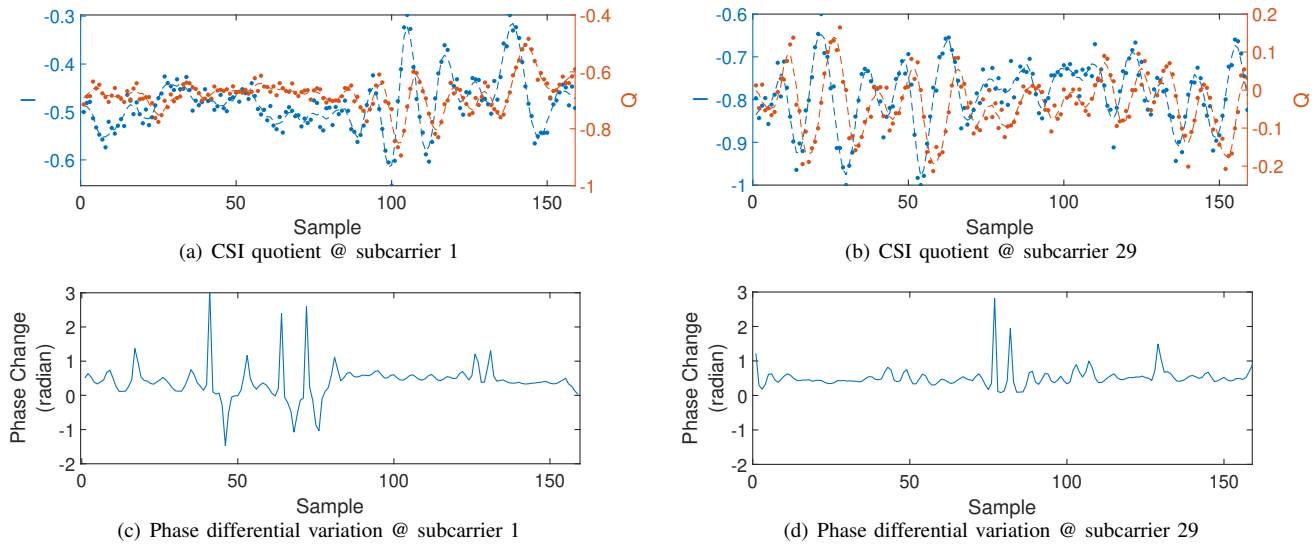
The quality of subcarriers can be measured as the variation range of the derivative of the tangent slope. To score the sub-carriers and select the signal of the best quality, we calculate the variation range with a sliding window whose size is longer than a full cycle. The variation in curvature can be approximated by the tangent slope differentials. Fig. 11(c) and Fig. 11(d) show the corresponding phase differential variations. As can be seen from the figure, both subcarriers contain data segments that have high phase differential variations which indicate poor signal quality. On the other hand, data segments of smooth variation indicate a good sensing quality. By assigning better scores for smoother phase differential variations, we can select the best signal, i.e. the best subcarrier piece-by-piece from all candidate sub-carriers.

We use the moving range to score signal quality. The moving range is defined as the range between the minimum and maximum values of the phase differential data within a sliding window. The score of the CSI-quotient of all the sub-carriers is calculated individually. Then we pick the displacement information of corresponding sub-carriers segment-by-segment with the highest score and combine them to produce a final estimation.

F. Synthesizing Information from Multiple Views

Recovering 2-dimensional displacement and reconstructing the trajectory requires accurate displacement information from each dimension. In the previous step, we calculated the reflection-path-length-change without considering the human trajectory estimation accuracy. According to the empirical study in section III-A, for a walking track, the estimated displacement on a single transceiver pair is accurate only when the angle between the walking route and Fresnel zones is greater than 30 degrees. However, with information of one transceiver device pair alone, it is not possible to tell which piece of the data is accurate in estimating displacement and which is not.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing



(a) CSI quotient @ subcarrier 1 (b) CSI quotient @ subcarrier 29

(c) Phase differential variation @ subcarrier 1 (d) Phase differential variation @ subcarrier 29

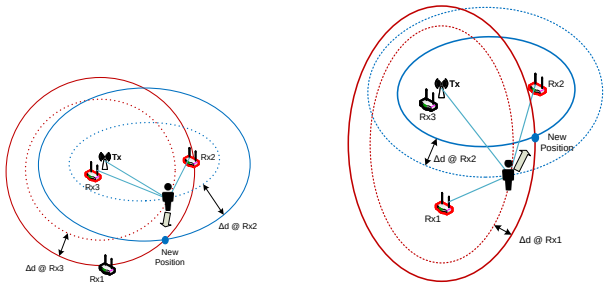
Fig. 11. The human motion induces different signal variations over sub-carriers. This figure shows the same piece of CSI quotient signal and its dynamic phase differentials of two sub-carriers. The signal quality is well described by the dynamic phase differentials, as the clearer signals have smoother phase changes. By combining all the sub-carriers, we can pick the best signal segment-by-segment to fully capture the motion displacement on a receiver.

The idea here is to make use of the complementary nature of the viewing angles of multiple WiFi transceiver links. That is, as multiple WiFi receivers capture the reflection signal of walking from different viewing angles, there is always a viewing angle that can provide more accurate reflected path length change estimates. By comparing the confidence of estimation from multiple alternative views over time, we can discard the estimations from the devices that have a large estimation error (which also have poor confidence), without having to judge whether the walking angle exceeds the 30-degree threshold.

We can achieve this by dividing the walking route into small segments and selecting path length change information from two best-scored views among multiple WiFi links to recover the 2-D displacement. Our observations are very similar to that in Sec. V-E2. As observed in Sec. III-A, when a person's walking route is roughly perpendicular to the Fresnel zone boundaries, the CSI quotient signal looks like perfect circles on the complex plane. The smaller the angle to Fresnel zone boundaries, the more distortion to the circles in the CSI quotient signal. We use a similar scoring criterion as in Sec. V-E. We use the minimum sliding range to rate the confidence of the estimation on different receiving devices. The minimum sliding range is the minimum value of the sliding ranges between multiple sub-carriers on that receiving device. With the minimum sliding-range data of each receiver, we can always select two estimates with the least estimation error from multiple views. Based on the geometric positions of the chosen WiFi devices, and the corresponding displacements on that devices, we can compute the reflected path length changes, and then reconstruct the 2-D motion trajectory.

G. 2-D Displacement Reconstruction

The information computed above describes how the length of the reflected path changes, not the moving subject itself. With the length change of the reflected path, the subject can be at any position on the ellipse of a Fresnel zone boundary. Hence one set of transceivers is not enough to reconstruct the whole trajectory as a 2D motion trajectory is composed of displacement in two orthogonal directions. We need at least two displacement information from different transceiver pairs. The two transceiver pairs can form crossed ellipses, where the subject location is marked as the intersection point of the two ellipses.

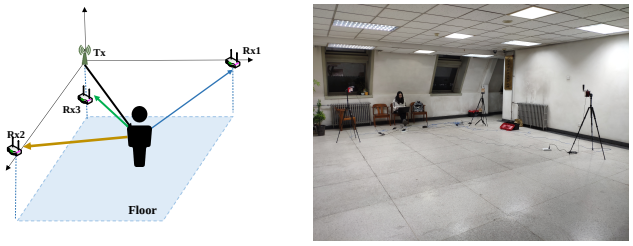


(a) using Rx2 and Rx3 (b) using Rx1 and Rx2

Fig. 12. The receiver devices are first selected dynamically based on their confidence scores, and then the trajectory are calculated iteratively by solving a group of ellipse equations based on the selected transceiver device pair (marked in red) and the length changes of propagation path on that devices

We reconstruct the motion trajectory iteratively. We first crop out all the non-walking data segments, as described in Sec. IV. Then, we mark the position of each WiFi device, calculate the position of the target subject based on the information of the previous position and the displacement related to the two pairs of transceivers, as shown in Fig. 12. The calculation is done by solving a group of ellipse equations, similar to previous approaches [19], [30]. Solving a group of ellipse equations may result in multiple location candidates.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing



(a) Device placement (b) The setup in a corridor

Fig. 13. Experimental setup with 1 transmitter and 3 receivers.

However, with the constraint of the sensing area, we will be left with only one possible solution. Please note that we use the method proposed in [19] to estimate the initial position. This will be discussed further in Sec. VII-C.

VI. Evaluation

In this section, we evaluate the performance of WiTraj with commodity WiFi devices. We first describe the experimental setup and illustrate some reconstructed trajectories, after which we evaluate the accuracy of WiTraj in various settings. Next, we compare WiTraj performance with state-of-the-art and finally demonstrate WiTraj in a natural setting.

A. Experimental Settings

In all the experiments, we used four GigaByte mini PCs as WiFi transceivers. Each receiver was equipped with an Intel 5300 wireless card and three omnidirectional antennas. The receivers were configured to work under monitor mode, so all the receivers can capture packets from the transmitter simultaneously. We mounted one transmitter (marked as Tx) and three receivers (marked as Rx1, Rx2, and Rx3, respectively) on tripods, as illustrated in Fig. 13. In the experiment, Tx, Rx1, and Rx2 were fixed on the three vertexes of a rectangular sensing area. In most cases, Rx3 was placed just below the Tx. The exact locations of the devices were measured by a laser range-finder. The height of the Tx, Rx1, and Rx2 devices was 1.3 meters, and the height of Rx3 was 0.9 meters. We used the open-source Linux CSI tool [35] released by the University of Washington to collect CSI data from the receivers. The frequency was set to 5.24 GHz, bandwidth to 20 MHz and the transmitter sent 400 packets per second at a transmit power of 15 dBm.

We evaluated the tracking accuracy in both outdoor and indoor environments. Specifically, we tested in four different environments, including an outdoor parking lot, a meeting room (7 m $\times$ 6 m), a corridor (6 m $\times$ 8 m), and an apartment house. In contrast to the outdoor lot and almost empty corridor, the other two rooms had a lot of furniture such as tables, chairs, and sofas. We placed the WiFi devices in the corners half a meter away from the wall. In most cases, the LoS between the transceiver devices was between 4 and 6 meters. Fig. 13 shows the setting of a corridor environment. We recruited 9 participants to participate in the experiments, including seven males and two females, ranging in age from 20 to 40, with their BMI between 18 and 24.

B. Tracking Accuracy

To evaluate the tracking accuracy of WiTraj, we put some labels on the ground to mark the key points of the walking routes and used a camera to record video as the ground-truth. The WiFi devices were placed around a square area of 6-by-6 meters while the sensing area of walking was set to 4 $\times$ 4 meters. We invited five participants to walk along with four sets of tracks including a circle, square, Z-shaped, and diamond-shaped trajectories. The participants were asked to walk naturally, such as swinging their hands during walking or stopping briefly at the turn points of the trajectories. Closed trajectories such as the circle, square and diamond were performed in both clockwise and counterclockwise directions, thus we have twice the amount of data for these trajectories. The initial position of all tracks was fixed and known at the time of calculation. The data was collected over a time span of two weeks. In total, we collected 60 sets of data for each volunteer. A similar method as IndoTrack [19] was used to calculate the error.

Fig. 15 shows the example of the reconstruction of three tracks, i.e. a circle, a square, and a diamond. The reconstructed trajectory is painted in color, showing a sequence of time from blue to red. For each track, the error is calculated as the distance between the calculated coordinates of each human position and the reference track. To be specific, for square, diamond, and Z-shaped tracks which consist of straight-line segments, we first find the turning point between the segments on the calculated trajectory, and then calculate the distance between the coordinates of each point in each trajectory segment and the reference segment to calculate the Euclidean distance. Based on this WiTraj had an overall median tracking error of 0.26 meters and a 90^{th} percentile error of 0.82 meters for all trajectory data. Given the trajectory lengths for the various shapes, the median tracking error translates to $< 2.5\%$.

1) Impact of different environments: To study the impact of different environments, we compared the performance of WiTraj in two indoor settings, i.e., a corridor and a conference room, as well as an outdoor environment. While the shape of the two rooms and the interior furniture layout is different, we used the same settings of the device placements, the LoS spacing between transceivers, and the size of the sensing area between the different environments to better compare them. We collected the data of four walking trajectories. Fig. 16(a) shows the cumulative distribution function (CDF) of target tracking error in these environments. As shown in the figure, outdoor tracking of the target is more accurate due to the simpler multi-path reflection of the environment, with a median error of 0.13 meters, while the indoor environment is worse, with the median tracking error in the corridor and conference room being 0.36 meters and 0.44 meters, respectively.

2) Impact of trajectory shapes: We selected four types of shapes to evaluate the WiTraj system's tracking performance, including zigzag, rectangular, circular, and diamond tracks. The data was collected in all three environments by 9 participants. For zigzag-shaped trajectory, we collected data in two ways - where the initial part of the path is parallel and perpendicular to the x-axis, hence forming the shape of 'Z'

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing

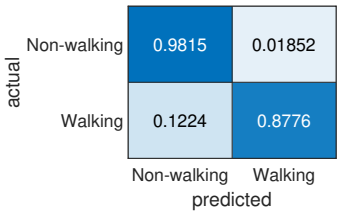
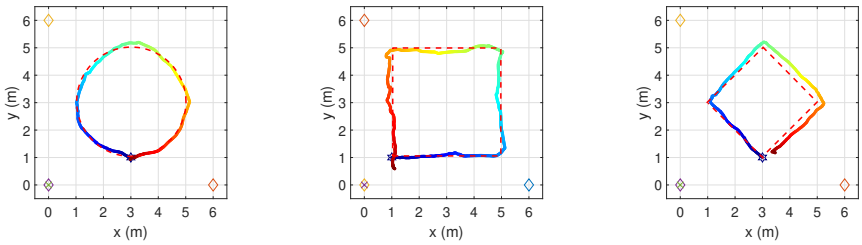
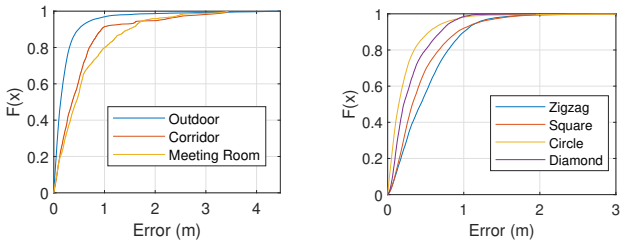


Fig. 14. The confusion matrix of in-place activities and walking

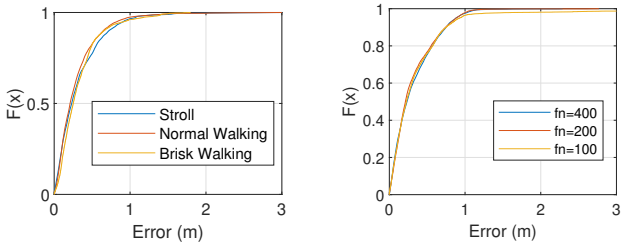


(a) A circle (b) A square (c) A diamond

Fig. 15. Illustration of WiTraj trajectory reconstruction. The reconstructed track is painted from blue to red, with the dashed lines marking the ground-truth. × and ◇ mark the positions of WiFi transmitter and receivers, respectively.



(a) different environments (b) different trajectories



(c) different speeds (d) different sampling rates

Fig. 16. The CDF of tracking error under various settings

and 'N', respectively. For rectangular, circular, and diamond-shaped trajectories, we collected two sets of walking data in both clockwise and counterclockwise directions. We combined these two data sets in each shape category and sum up the results of the three environments. As shown in Fig. 16(b), the median tracking errors for the four tracks are 0.43 m, 0.40 m, 0.15 m and 0.21 m, respectively. Of these four trajectories, the accuracy of circular and diamond shapes is better. By analyzing the extracted Doppler speed, we found that the diamond and circle tracks cross the Fresnel zone boundaries of at least two of the three WiFi device pairs at a large angle most of the time. Therefore, the estimated speed information is more accurate, hence a better-reconstructed trajectory. On the other hand, the zigzag-shaped and rectangular trajectories have some piece of track segments that have small angles between the walking route and Fresnel zones of two of the three receiving devices, making the error larger than the circular and diamond shapes.

3) Impact of walking speed: We evaluated different walking speeds of some tracks in an outdoor environment. Specifically, we set the distance between Tx and Rx to 4 m and 6 m respectively in two tests and evaluated the tracking accuracy of square and circular trajectories in the sensing area of 3×3 m and 4×4 m, respectively. Two participants traveled along these tracks in clockwise and counterclockwise directions, and 10 data were recorded in each case. We asked the participants to walk at three speeds. The slow speed is less than 1 m/s, which is equivalent to strolling speed; the normal speed is 1–1.5 m/s, which is about the normal outdoor walking speed; and the fast movement is 1.5-3 m/s, close to brisk walking speed. Fig. 16(c) shows the result for different walking speeds. WiTraj is not affected by the walking speeds and we did not observe differences in the reconstruction quality among all three speeds. As a result, WiTraj can track the motion of a human target at most walking speeds.

4) Impact of sampling rates: We use down-sampling to illustrate the impact of sampling rates on tracking accuracy. We use the same data set of circular, square, and diamond tracks recorded by two participants at a sampling rate of 400 Hz. Then we down-sample the data by half several times in order to evaluate the reconstruction error at each sampling rate. The participants walk at a normal speed (approximately 1m/s in the experiment). All the experimental data in this test were taken in an outdoor environment, with the distance between Tx and Rx set to 4 meters and the sensing area set to 3×3 meters. For each trajectory, we collect 20 data samples, 10 for clockwise walking and 10 for counterclockwise walking. The results are shown in Fig. 16(d). We see no obvious difference when the sampling rate is greater than 100 Hz. The reconstruction is successful when the sampling rate is equal to or greater than 100 Hz, but failed for the sampling rate of 50 Hz or lower. The main reason is that the signal waveform at this sampling rate is no longer sufficient to describe the signal fluctuation period brought about by walking. According to the Fresnel Zone model [28], [36], the CSI signal frequency generated by normal speed walking is between 10 Hz and 40 Hz. Hence according to the Nyquist sampling theorem, a minimum sampling rate of at least 80 Hz is required to fully capture this information. In practice, we set the sampling rate higher than 150 Hz for motion tracking, in order to cover a wider walking speed range.

5) Impact of device placements: To study how multi-view helps in reconstructing the trajectory, we changed the number of receiver devices as well as the device placement and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing

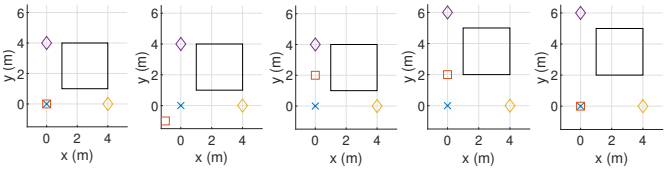


Fig. 17. The five device placement settings. The square in the figure marks the sensing area of 3x3 meters. ×, ◇, and □ mark the positions of WiFi transmitter, two receivers, and the third receiver, respectively.

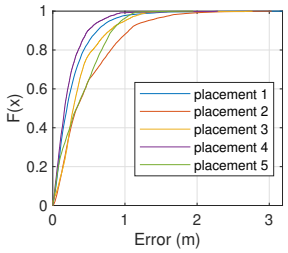


Fig. 18. Tracking error of different device placements

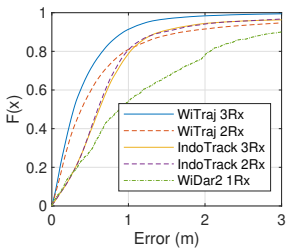


Fig. 19. Tracking error comparison of three algorithms

evaluated the tracking performance in outdoor environments. The sensing area is set to a 3×3 m square in all settings. Two participants walk Z-shaped, square, round, and diamond-shaped tracks in the sensing area. We tested 5 settings of device placements as shown in Fig. 17. Each participant walked each track 10 times to collect data. We evaluated the impact of device placement in five settings. In Fig. 17 (a-c), the distance between Tx-Rx is 4 meters, however, the three cases have different locations of the third receiving device. In Fig. 17 (d-e), the LoS distance of one of the Tx-Rx pairs increases to 6 meters. The results are shown in Fig. 18. There are some differences in the tracking performance, but the difference is minor with the median errors in all these settings less than 0.33 meters.

6) Impact of the number of views: Since two receiving devices can provide information to reconstruct a trajectory, and additional receiving devices can improve the reliability of the reconstruction, we evaluate how additional views improve tracking performance in outdoor and indoor corridor environments in this section.

For the outdoor environment, we use the same data set in Sec. VI-B5. We only used the data of Rx1 and Rx2 in the five settings of Fig. 17 to simulate the settings of two views. In addition, we also added a fourth WiFi receiving device in Fig. 17 (c-d), located at the coordinate of (0,0), to simulate the settings of four views. In the case of more than two views, such as the cases of three and four views, we use the same algorithm to select information of the two best views to reconstruct the tracks. The dataset contains four tracks: Z-shaped, square, circular, and diamond. The results are compared in Fig. 20(a). It can be seen that in the outdoor environment, two receiving devices can reconstruct walking tracks, while three receiving devices produce fewer errors. However, in the case of 4 views, the information from more devices has a limited effect on improving track accuracy.

For the indoor environment, we use one transmitter and

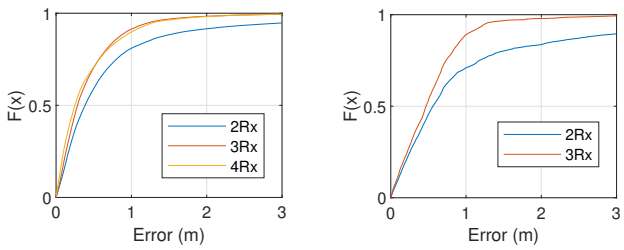


Fig. 20. Tracking error of different device numbers.

three receivers, as the setting in Fig. 17. Four participants walk along square and circular trajectories. We collected 20 data samples for each trajectory. Similar to the outdoor test, we removed the third receiving device represented by the square mark, to simulate the case of two receiving devices, so that the reconstruction by 2 devices and 3 devices can be compared on the same data set. The result is shown in Fig. 20(b). As can be seen, although the improvement of median error is moderate, the multi-view approach can significantly improve the performance of 90^{th} percentile error.

7) Impact of Environment change and interference: We evaluate the performance impact in the case of environmental change and human interference. For this, we conducted experiments in a room of 7×7 m. The floorplan of the room, device placement, and sensing area are illustrated in Fig. 21. We use two trajectory shapes, i.e., diamond and circle, to evaluate the performance impacts. We first let two participants walk along each trajectory ten times as the baseline.

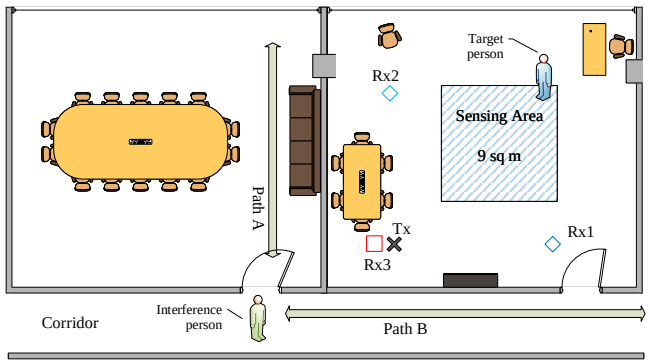
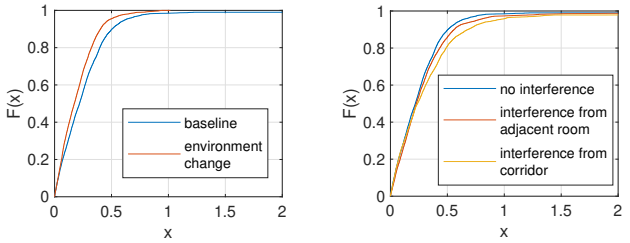


Fig. 21. The experiment design on evaluating the environment change and human interference.

To evaluate the impact of environmental change, we move tables and chairs to new positions, open the door of the room, and collect data for a second round. The accuracy is compared in Fig. 22(a). We see that environmental changes have little impact on tracking performance. This result is in line with expectations since the environment change only alters the static paths, while the CSI quotient model focuses on the dynamic reflection path. Hence the number of static paths and their distribution do not affect the performance.

To evaluate the impact of human interference, we designed two tests. In the first setting, we let a person walk back-and-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing



(a) Environment change (b) Human interference

Fig. 22. The CDF of tracking error under environment change and human interference.

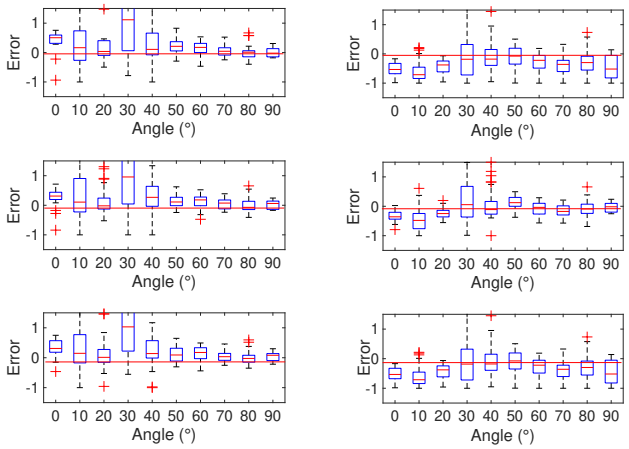
forth through *path A* in an adjacent room. In the second setting, we let a person walk back-and-forth through *path B* at the corridor. In each test, we collected the same amount of data as in the baseline test. The detailed settings are illustrated in Fig. 21, and the accuracy is compared in Fig. 22(b). We can see the accuracy is not impacted much in both settings. Since the walls block/weaken most of the signal reflecting off the person outside the room, a person moving in an adjacent room or outside the door does not have much impact on the tracking performance, provided people outside the room are not too close to the person being tracked, as shown in these two tests. It should be noted that the proposed system is designed to detect a single moving person as it is right now. We plan to tackle the multiple-subject tracking problem in our future work.

C. Comparison with state-of-the-art

In this section, we compare WiTraj with existing work. Since a lot of indoor motion tracking work [18], [19] is based on DFS information, we compare WiTraj with two DFS-based tracking systems, the IndoTrack system [19] and WiDar2 system [18]. We first compare the error of one-dimensional motion sensing (using one transceiver device pair) with the Doppler-MUSIC algorithm [19], the core component of IndoTrack and WiDar2, using the same data set of straight-line walking recorded in section III-A; then we compare the tracking accuracy of the three systems.

1) Comparison of system stability and one-dimensional estimation: In Sec. II, we compared the theoretical advantages and disadvantages of CSI quotient and CSI conjugate multiplication as the base signal. Since the Doppler-MUSIC algorithm takes the CSI conjugate multiplication as input, it suffers from the problem of speed ambiguity. As a consequence, the performance of Doppler-MUSIC-based tracking systems is unstable. Hence the Doppler-MUSIC algorithm can achieve good accuracy after carefully adjusting the placement of WiFi devices, but it is not robust and even a little change in the device positions can result in poor performance.

In this section, we compare the stability of the Doppler-MUSIC algorithm with WiTraj in one-dimensional displacement estimation. We adjusted the locations of WiFi devices slightly every time and observe the performance variance for each position. The device placement is illustrated in Fig. 3. We tested straight-line walking paths of different starting and



(a) WiTraj (b) Doppler-MUSIC

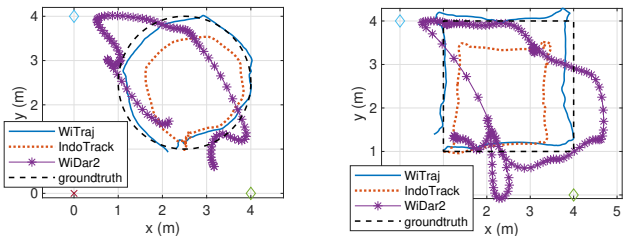
Fig. 23. The comparison of error distributions of WiTraj and Doppler-MUSIC in 1D displacement estimation. The three data sets use different device positions with WiTraj achieving more consistent estimation.

ending points. A total of three persons participated in the experiment. For each straight-line path, each person walked 10 times. We adjusted the location of the WiFi device by 1–3 cm each time, and repeated the above process three times. Hence we got three sets of data. The ground-truth of the reflected path length change d_{ref} is calculated based on the location of the start point and endpoint. The estimated path-length-change d_{est} was calculated using the WiTraj and Doppler-MUSIC algorithms. Error is calculated as $(d_{\text{est}} - d_{\text{ref}})/d_{\text{ref}}$, and the results are grouped by the angle of the walking route per 10 degrees relative to Fresnel zones, as shown in Fig. 23. As can be seen from Fig. 23, both the estimates of Doppler-MUSIC and WiTraj produce some deviations in these three sets of data. The WiTraj system produces more consistent estimations, with large errors in small angles and quite accurate estimations in angles larger than 40 degrees. As a comparison, the performance of Doppler-MUSIC changes at each setting of device placement, e.g. in position 2, the Doppler-MUSIC algorithm obtains accurate motion distance estimation. The error at a smaller walking angle is better than WiTraj. However, the overall estimation of position 1 and position 3 is obviously off, where the displacement is underestimated. After checking the data, we found that the error comes from the opposite sign of the estimated Doppler speed at some data segments so that it accumulates less displacement than it should.

Through the analysis of the above experimental results, the WiTraj system is more robust to changes in the environment where the multi-path condition can vary due to small changes in the environment and/or device positions. Doppler-MUSIC algorithm can provide good results if the device positions are carefully chosen. But once the environment or device placement changes, the error can increase significantly.

2) Two-dimensional tracking accuracy: In this section, we compare the tracking error of WiTraj in the 2D plane compared

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2021.3133114, IEEE Transactions on Mobile Computing



(a) circle (b) square

Fig. 24. The demonstration of the reconstructed trajectories for the three systems. The groundtruth is marked as the dashed lines.

to IndoTrack and WiDar2. The evaluation of IndoTrack and WiTraj uses the same data set of four trajectory shapes walking collected in Sec. VI-B. For these two systems, we calculate the displacement along each dimension and reconstruct the 2D trajectory using the algorithm described in Sec. V. We then evaluate the tracking accuracy for two cases: using information from only two receiving devices vs. three receiving devices.

On the other hand, we use a dedicated data set collected from a single device-pair for WiDar2 since that technique uses AoA and range (using ToF and DFS) information from a single WiFi link only. The sensing area and the shape of walking traces are the same as the settings of the other two systems. However, the performance of ToF and AoA estimation of WiDar2 is not very stable. We set up WiDar2 in a corridor by carefully arranging the placement of WiFi devices, then we collect the data set. The transceiver LoS is measured as 4.2 meters. WiDar2 requires an additional CSI phase offset calibration step for the correct AoA estimation, due to the initialization of the phase-locked loop of different antenna RF chains. We followed the same calibration process as described in [37]. Finally, we collected 10 traces for each of the four trajectory shapes. As both IndoTrack and WiDar2 use the Doppler-MUSIC algorithm for DFS estimation, we picked two suitable antennas from the 3 available antennas of 5300 in the conjugate multiplication step to alleviate the ambiguous speed estimation problem, as described in [19]. We demonstrate the reconstructed walking trajectories of a circle and a square for the three systems in Fig. 24. IndoTrack reconstructed a smaller shape than the ground truth. The error comes from the ambiguous sign of DFS, so it eventually accumulated lower displacement than expected. WiDar2 had the worst trajectory reconstruction. As the trajectory depends on AoA, ToF, and Doppler speed information, it is hard to ensure the correction of all the parameters simultaneously with only one WiFi link and the real-life constraints of limited antennas and bandwidth.

The tracking accuracy can be compared to Fig. 19. As can be seen from the figure, WiTraj performs better than IndoTrack on the same data set. With 3 viewing-angles, WiTraj achieves the best median error as well as 90^{th} percentile error. However, WiTraj has reduced accuracy when using the information of two viewing angles. Finally, WiDar2 uses only one viewing angle and has the worst performance among the three.

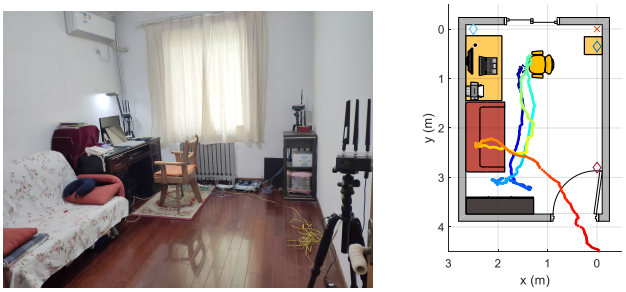
3) Computational Cost: The WiTraj system is more computationally efficient than the Doppler-MUSIC algorithm and is capable of running in real-time. We use 22-second CSI data received on an Rx device as input, and compare the processing time of the two algorithms at different sampling rates by applying the down-sampling method. We run Matlab code of both algorithms on a Microsoft Surface Book 2 that has an Intel Core i7 8650U processor. The running-time and speed-up are listed in Table I. As can be seen, the calculation of WiTraj takes much less time, with roughly 40x speed-up compared with the Doppler-MUSIC algorithm.

TABLE I
COMPUTATION TIME AND SPEED-UP OF THE TWO ALGORITHMS

FN	Samples	Doppler-MUSIC (s)	WiTraj (s)	Speed-Up
400	8800	21.42	0.49	43.7x
200	4400	10.67	0.29	36.6x
100	2200	5.07	0.15	34.0x

D. Example of real-world tracking

Putting these together, we can see that WiTraj can track the walking trajectory in real-world scenarios fairly well. Fig. 25(a) shows an example of tracking in a study with a size of 2.8 x 3.9 meters. The room is furnished with tables, a chair, a sofa, and a bookcase (where the bookcase is not visible in Fig. 25(a)). A participant walks from the chair in front of the desk to the bookcase, then returns to the desk. After that, the participant goes to the sofa to rest and finally goes out of the room. The floor-plan of the room and the reconstructed walking track are shown in Fig. 25(b). The trajectory is painted from blue to red. As can be seen, WiTraj reconstructs the entire walking trace quite well.



(a) indoor furniture includes tables, a chair, a sofa and a bookshelf (b) The reconstructed trajectory

Fig. 25. The tracking illustration using one transmitter and three receivers in a study.

VII. LIMITATIONS AND DISCUSSION

In this section, we discuss the limitations of the proposed method and future work.

A. Single-person tracking

Using a single WiFi transmitter and three receivers, WiTraj enables tracking of an individual in a room with default WiFi configuration and bandwidth constraints. WiTraj leverages the power of the CSI quotient model to sense the distance moved by a single person. It also uses multiple receivers to capture

walking from multiple viewing angles and selects proper views to accurately track the distance moved regardless of the motion direction and also reduces the negative impact of complex reflections of the human body during walking.

A limitation of the work is that it cannot track multiple persons at the same time. As the walking of multiple people produces complicated superposition on the received signal, better hardware (such as more antennas) and larger bandwidth are needed to separate the signals of each individual for motion tracking. Larger bandwidth by combining multiple WiFi channels has been shown to help to distinguish the signal arrival times of different paths (e.g. MultiTrack [17]), but that is hard to achieve with standards-based WiFi as noted earlier.

B. Accumulated Tracking Error

Limited by the default bandwidth of commercial WiFi devices, WiTraj is not based on absolute ranging information for motion tracking, but rather on the change in the distance the signal transmits over the reflected path to sense the displacement of the target. WiTraj uses an iterative approach to calculating walking trajectories. Such a system is susceptible to accumulating tracking errors over time.

Nevertheless, WiTraj attempts to alleviate the tracking errors by identifying the data segments of in-place activities, so that the corresponding DFS will not be used in the trajectory calculation. Additionally, the multiview approach reduces the error of trajectory reconstruction, so that it has less accumulated tracking errors than other traditional two-view methods.

C. Assumption of known initial position

WiTraj calculates walking trajectories iteratively based on displacement and previous human position. Thus, it is required to know a person's initial position.

This requirement can be satisfied in many cases by using existing methods. For example, by combining AoA and DFS, a person's initial position can be inferred, as shown in [19]. Another example is to monitor whether a person is walking through a door using WiBorder [38]. When detecting the person entering the room through the door, the location of the door can be used as the initial position. In practice, as there are usually several rooms in an apartment, a person's position while crossing any room door can be treated as anchor points using the WiBorder technique.

VIII. Related Work

There are a number of techniques in the literature that use WiFi for localization and motion tracking. In this section, due to space constraints, we review *device-free* approaches only.

A. WiFi Device-free Localization

RSSI was one of the initial pieces of information exploited in device-free passive localization (DFPL) systems due to easy availability on COTS devices. RSSI is coarse-grained and unstable, hence most of the RSSI-based DFPL methods leverage multiple TX-RX pairs to achieve high-accuracy localization [12]. Subsequently, CSI which is a finer-grained feature compared to RSSI was exploited. CSI measures the wireless channel between the transmitter and receiver in terms of phase and amplitude. Initial techniques include FIMD [13], MonoPHY [15] and Pilot [16] that explored CSI-based fingerprinting methods to localize objects. Since they are based on fingerprinting, they have the drawback of labor-intensive training and not being resilient to environmental changes.

Other systems use model-based approaches to localize targets without training. LiFS [39] leverages signal phase characteristics in and out of the Fresnel Zones to achieve tracking but requires the specific and dense deployment of WiFi devices. MaTrack [14] uses AoA estimation of multiple known WiFi receivers for triangulation, but it requires at least three antennas which are not available in today's IEEE 802.11ac WiFi clients (typically equipped with two antennas). Hence current device-free positioning systems that are compatible with current WiFi implementations are too coarse to support accurate motion tracking. MultiTrack [17] leverages ToF information and examines the path of the reflected signals at multiple links to simultaneously track multiple users, but requires high bandwidth by splicing all the WiFi channels. This frequency hopping scheme is not compatible with WiFi standards.

B. WiFi Device-free Motion Tracking

This class of systems senses a subject's motion rather than localizing it using COTS WiFi devices. WiDar [18] extracts unsigned PLCR information and simultaneously estimates a human's moving velocity (both speed and direction) and location with an optimization method using CSI amplitude signal. As speed direction is undetermined, it may produce wrong motion estimation. DFS-based solutions (e.g. WiDance [26] and IndoTrack [19]) take a different approach by using the DopplerMUSIC algorithm. The DFS of two device links are then combined to recover the moving speed of a person. On the other hand, WiDar2.0 [20], xD-Track [40] and mD-Track [21] use a single WiFi link, but jointly estimate AoA, DFS and ToF which enables sub-meter tracking.

Most of the above device-free motion tracking systems suffer from amplification of the CSI noise and ambiguous speed estimation problems [18], [26], [19], [20], [40], [21]. They stem from the conjugate multiplication of the CSI in the DFS calculation [26], [19], [20], [40], [21], or lack of sign in speed estimation [18]. In particular, when the signal strengths received on the two antennas are similar, it is difficult to determine the sign of the extracted Doppler speed, even with the trick mentioned in [19]. Thus, motion tracking becomes unstable in many practical scenarios.

In addition, these systems are often not robust to cases where any underlying parameters are inaccurate. For example, the Doppler speed of human walking as measured from a single link is not always accurate, especially in some walking directions/orientations, as the motion of all body parts contribute to a Doppler speed component.

In contrast, WiTraj uses the CSI quotient to estimate DFS. It cancels both the CSI noise in amplitude and phase to greatly increase the SNR while having no ambiguous speed problems.

It also ensures at least two accurate DFS estimations to ensure reliable motion tracking by relying on a multi-view approach.

IX. Conclusion

This paper presents WiTraj, an indoor motion tracking system with commodity WiFi devices. WiTraj achieves robust motion tracking by three steps: 1) It extracts reliable DFS by leveraging the CSI-quotient model, which provides higher SNR and unambiguous Doppler speed estimation. 2) By placing multiple receiving devices, WiTraj ensures that the system obtains reliable DFS from at least two viewing angles at any time regardless of the human position and walking direction. By assessing the quality of motion displacement estimates in different views, the information of two views with the least error is chosen to reconstruct the motion trajectory in the 2D plane, leading to a median tracking error of 0.3 meters. 3) WiTraj discriminates in-place activities from walking, so that typical daily activities do not add drift to the constructed trajectory. Based on the experiments and results, we believe WiTraj can serve as a practical basis for high-level human indoor mobility and behavioral analysis.

Acknowledgment

This research is supported by NSFC 62061146001, PKU-Intel collaboration project, PKU-Baidu Funded Project 2019BD005, and Huawei-PKU collaboration project.

References

[1] G. Welch and E. Foxlin, "Motion tracking: No silver bullet, but a respectable arsenal," *IEEE Computer graphics and Applications*, vol. 22, no. 6, pp. 24–38, 2002.

[2] H. Zhou and H. Hu, "Human motion tracking for rehabilitation—a survey," *Biomedical Signal Processing and Control*, vol. 3, no. 1, pp. 1–18, 2008.

[3] R. Mautz and S. Tilch, "Survey of optical indoor positioning systems," in *2011 international conference on indoor positioning and indoor navigation*. IEEE, 2011, pp. 1–7.

[4] E. M. Gorostiza, J. L. Lázaro Galilea, F. J. Meca Meca, D. Salido Monzú, F. Espinosa Zapata, and L. Pallarés Puerto, "Infrared sensor system for mobile-robot positioning in intelligent spaces," *Sensors*, vol. 11, no. 5, pp. 5416–5438, 2011.

[5] C. Sertatıl, M. A. Altınkaya, and K. Raoof, "A novel acoustic indoor localization system employing cdma," *Digital Signal Processing*, vol. 22, no. 3, pp. 506–517, 2012.

[6] Q. Yuan and I.-M. Chen, "Localization and velocity tracking of human via 3 imu sensors," *Sensors and Actuators A: Physical*, vol. 212, pp. 25–33, 2014.

[7] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3d tracking via body radio reflections," in *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*, 2014, pp. 317–329.

[8] F. Adib, Z. Kabelac, and D. Katabi, "Multi-person localization via rf body reflections," in *12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15)*, 2015, pp. 279–292.

[9] J. Wang, J. Xiong, H. Jiang, X. Chen, and D. Fang, "D-watch: Embracing "bad" multipaths for device-free localization with cots rfid devices," *IEEE/ACM Transactions on Networking*, vol. 25, no. 6, pp. 3559–3572, 2017.

[10] Z. Sahinoglu, S. Gezici, and I. Guvenc, "Ultra-wideband positioning systems," *Cambridge, New York*, 2008.

[11] P. Bahl and V. N. Padmanabhan, "Radar: An in-building rf-based user location and tracking system," in *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No. 00CH37064)*, vol. 2. Ieee, 2000, pp. 775–784.

[12] M. Seifeldin, A. Saeed, A. E. Kosba, A. El-Keyi, and M. Youssef, "Nuzzer: A large-scale device-free passive localization system for wireless environments," *IEEE Transactions on Mobile Computing*, vol. 12, no. 7, pp. 1321–1334, 2012.

[13] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni, "Fimd: Fine-grained device-free motion detection," in *2012 IEEE 18th International conference on parallel and distributed systems*. IEEE, 2012, pp. 229–235.

[14] X. Li, S. Li, D. Zhang, J. Xiong, Y. Wang, and H. Mei, "Dynamic-music: accurate device-free indoor localization," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 196–207.

[15] H. Abdel-Nasser, R. Samir, I. Sabek, and M. Youssef, "Monophy: Mono-stream-based device-free wlan localization via physical layer information," in *2013 IEEE wireless communications and networking conference (WCNC)*. IEEE, 2013, pp. 4546–4551.

[16] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni, "Pilot: Passive device-free indoor localization using channel state information," in *2013 IEEE 33rd International Conference on Distributed Computing Systems*. IEEE, 2013, pp. 236–245.

[17] S. Tan, L. Zhang, Z. Wang, and J. Yang, "Multitrack: Multi-user tracking and activity recognition using commodity wifi," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[18] K. Qian, C. Wu, Z. Yang, Y. Liu, and K. Jamieson, "Widar: Decimeter-level passive tracking via velocity monitoring with commodity wi-fi," in *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2017, pp. 1–10.

[19] X. Li, D. Zhang, Q. Lv, J. Xiong, S. Li, Y. Zhang, and H. Mei, "Indotrack: Device-free indoor human tracking with commodity wi-fi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 1–22, 2017.

[20] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2. 0: Passive human tracking with a single wi-fi link," in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, 2018, pp. 350–361.

[21] Y. Xie, J. Xiong, M. Li, and K. Jamieson, "md-track: Leveraging multi-dimensionality for passive indoor wi-fi tracking," in *The 25th Annual International Conference on Mobile Computing and Networking*, 2019, pp. 1–16.

[22] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "Spotfi: Decimeter level localization using wifi," in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, 2015, pp. 269–282.

[23] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity wi-fi," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 2015.

[24] Y. Zhuo, H. Zhu, H. Xue, and S. Chang, "Perceiving accurate csi phases with commodity wifi devices," in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*. IEEE, 2017, pp. 1–9.

[25] J. Zhu, Y. Im, S. Mishra, and S. Ha, "Calibrating time-variant, device-specific phase noise for cots wifi devices," in *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*, 2017, pp. 1–12.

[26] K. Qian, C. Wu, Z. Zhou, Y. Zheng, Z. Yang, and Y. Liu, "Inferring motion direction using commodity wi-fi for interactive exergames," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 1961–1972.

[27] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *Proceedings of the 21st annual international conference on mobile computing and networking*, 2015, pp. 65–76.

[28] D. Wu, D. Zhang, C. Xu, Y. Wang, and H. Wang, "Widir: walking direction estimation using wireless signals," in *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*, 2016, pp. 351–362.

[29] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single wifi access point," in *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, 2016, pp. 165–178.

[30] D. Wu, R. Gao, Y. Zeng, J. Liu, L. Wang, T. Gu, and D. Zhang, "Fingerdraw: sub-wavelength level finger motion tracking with wifi signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 1, 2020.

[31] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, "Farsense: Pushing the range limit of wifi-based respiration sensing with csi ratio of two antennas," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–26, 2019.

[32] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, "Multisense: Enabling multi-person respiration sensing with commodity wifi," *Pro-*

ceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 4, no. 3, pp. 1–29, 2020.

[33] N. Yu, W. Wang, A. X. Liu, and L. Kong, "Qgesture: Quantifying gesture distance and direction with wifi signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1–23, 2018.

[34] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 6, pp. 1–13, 2015.

[35] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, pp. 53–53, 2011.

[36] D. Zhang, H. Wang, and D. Wu, "Toward centimeter-scale human activity sensing with wi-fi signals," *Computer*, vol. 50, no. 1, pp. 48–57, 2017.

[37] J. Xiong and K. Jamieson, "Arraytrack: A fine-grained indoor location system," in *Presented as part of the 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*, 2013, pp. 71–84.

[38] S. Li, Z. Liu, Y. Zhang, Q. Lv, X. Niu, L. Wang, and D. Zhang, "Wiborder: Precise wi-fi based boundary sensing via through-wall discrimination," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 4, no. 3, pp. 89:1–89:30, 2020.

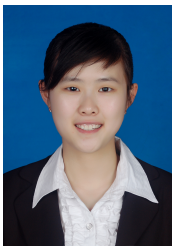
[39] J. Wang, H. Jiang, J. Xiong, K. Jamieson, X. Chen, D. Fang, and B. Xie, "Lifs: low human-effort, device-free localization with fine-grained subcarrier information," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, 2016, pp. 243–256.

[40] Y. Xie, J. Xiong, M. Li, and K. Jamieson, "xd-track: leveraging multi-dimensional information for passive wi-fi tracking," in *Proceedings of the 3rd Workshop on Hot Topics in Wireless*, 2016, pp. 39–43.

Dan Wu received a Ph.D degree in Computer Science from Peking University in 2020. He is currently an associate research fellow with the School of Computer Science at Peking University, China. His research interests include wireless sensing and mobile computing. Contact him at dan@pku.edu.cn.

Youwei Zeng received a B.E degree in Software Engineering from Zhejiang University in 2016. He is currently working toward a PhD degree in Computer Science in the School of Computer Science, Peking University. His research interests include ubiquitous computing and mobile computing.

Ruiyang Gao received a B.E degree in computer science and technology from Shandong University, China, in 2016. He is currently working toward a PhD degree in the School of Computer Science, Peking University. His research interests include mobile crowd-sensing and ubiquitous computing.



Shenjie Li received a Ph.D degree in Computer Science from Peking University in 2021. She is now with JD.com. Her research interests include wireless sensing, signal processing, and mobile health. Before coming to PKU, she received her Bachelor degree in 2015 from Jilin University.

Yang Li received a B.Eng degree in Software Engineering from Northwestern Polytechnic University in 2020. He is currently a Ph.D student in Computer Software with the School of Computer Science at Peking University.

Rahul C. Shah received a Ph.D degree in Electrical Engineering from University of California, Berkeley. He is currently a principal engineer and senior research scientist in Intel Labs. His research interests are in wireless sensing, communications, wireless networking, location sensing and applied machine learning. Contact him at rahul.c.shah@intel.com.

Hong Lu received a Ph.D degree in Computer Science from Dartmouth College in 2012. He is currently a senior research scientist at Intel Labs. His research focuses on developing mobile sensing system and machine learning techniques for smart phones, tablets, and wearables, which allow a device to observe, understand, or even predict peoples activities, social interaction, context, and life patterns. Contact him at hong.lu@intel.com.

Daqing Zhang (Fellow, IEEE) is a Chair Professor with the Key Laboratory of High Confidence Software Technologies (Ministry of Education), School of Computer Science, Peking University, China, and Telecom SudParis, IP Paris, France. His research interests include context-aware computing, urban computing, mobile computing, big data analytics, pervasive elderly care, etc. He has authored more than 300 technical papers in leading conferences and journals. He received his Ph.D. degree from the University of Rome "La Sapienza", Italy, in 1996. He was the General or Program Chair for more than 17 international conferences, giving keynote talks at more than 20 international conferences. He is the Associate Editor for the IEEE Pervasive Computing, ACM Transactions on Intelligent Systems and Technology, and the Proceeding of ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. He is the winner of the Ten-Years CoMoRea Impact Paper Award at IEEE PerCom 2013 and Ten-Years Most Influential Paper Award at IEEE UIC 2019, the Honorable Mention Award at ACM UbiComp 2015 and 2016, and the Distinguished Paper Award at ACM UbiComp 2021.