RF-based Inertial Measurement

Chenshu Wu, Feng Zhang, Yusen Fan, K. J. Ray Liu
University of Maryland, College Park and Origin Wireless Inc.
{cswu,fzhang15,ysfan,kjrliu}@umd.edu

ABSTRACT

Inertial measurements are critical to almost any mobile applications. It is usually achieved by dedicated sensors (*e.g.*, accelerometer, gyroscope) that suffer from significant accumulative errors. This paper presents RIM, an RF-based Inertial Measurement system for precise motion processing. RIM turns a commodity WiFi device into an Inertial Measurement Unit (IMU) that can accurately track moving distance, heading direction, and rotating angle, requiring no additional infrastructure but a single arbitrarily placed Access Point (AP) whose location is unknown. RIM makes three key technical contributions. First, it presents a spatial-temporal virtual antenna retracing scheme that leverages multipath profiles as virtual antennas and underpins measurements of distance and orientation using commercial WiFi. Second, it introduces a super-resolution virtual antenna alignment algorithm that resolves sub-centimeter movements. Third, it presents an approach to handle measurement noises and thus delivers an accurate and robust system. Our experiments, over a multipath rich area of >1,000 m^2 with one single AP, show that RIM achieves a median error in moving distance of 2.3 cm and 8.4 cm for short-range and long-distance tracking respectively, and 6.1° mean error in heading direction, all significantly outperforming dedicated inertial sensors. We also demonstrate multiple RIM-enabled applications with great performance, including indoor tracking, handwriting, and gesture control.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing**; • **Computer systems organization → Embedded and cyber-physical systems**;

KEYWORDS

Inertial Measurement, Wireless Sensing, Motion Tracking

ACM Reference Format:
Chenshu Wu, Feng Zhang, Yusen Fan, K. J. Ray Liu. 2019. RF-based Inertial Measurement. In *ACM SIGCOMM 2019 Conference (SIGCOMM '19), August 19–23, 2019, Beijing, China.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3341302.3342081

1 INTRODUCTION

Motion measurements are essential inputs for a range of applications such as robot navigation, indoor tracking, and mobile gaming, etc., and have been widely used in robots, drones, automotive, unmanned vehicles, various consumer electronics, and pretty much anything that moves. The mainstream has been using Inertial Measurement Units (IMUs) for motion tracking[1]. The rise in demand of accurate and robust motion tracking, coupled with the increase in smart device production, has been driving the IMU market, which is projected to grow from $15.71 billion in 2016 to $21.74 billion by 2022 [23]. An improvement to motion measurements will profoundly impact a number of systems and applications.

Precise and robust motion measurement, however, is non-trivial. The prevalent IMUs realized by MEMS sensors, namely accelerometers that measure linear acceleration, gyroscopes that calculate angular velocity, and magnetometers that report orientation, are well known to suffer from significant errors and drifts that are non-trivial to be recalibrated [33, 47]. For example, an accelerometer is hardly capable of measuring moving distance due to the noisy readings; magnetometer does not report heading direction and is easily distorted by surrounding environments; while gyroscope experiences considerable drifts introduced by integration especially in long run. These limitations prevent many applications that require accurate motion processing, such as indoor tracking, virtual reality, motion sensing games.

Recent years have witnessed much progress in using radio signals to localize and track targets. Despite novel systems that have led to decimeter or even centimeter location accuracy [13, 14, 43], they only address location while all suffer from significant common limitations that prohibit ubiquitous accurate inertial measurements: (1) They all require one or multiple precisely installed APs, as well as accurate information about their locations and/or orientations. A small error in the APs' geometry information will lead to large location errors. (2) They do not directly measure multiple motion parameters but rather can only determine one or two of them from successive location estimates. Nor can they track in-place angular motion. (3) Many of them face accuracy limitations dictated by frequency bandwidth, antenna amount, and synchronization errors on commercial off-the-shelf (COTS) WiFi, and degenerate or even fail in complex Non-Line-Of-Sight (NLOS) scenarios.

In this paper, we present RIM, an RF-based Inertial Measurement system that measures multiple parameters of object motions, namely moving distance, heading direction, and rotating angle. RIM turns standard COTS WiFi radios into precise IMUs, without support from additional infrastructure or external sensors. RIM does not require large bandwidth, many phased antennas, or multiple APs as reference anchors, nor does it need a priori calibration or fingerprinting of the environment. It has minimal requirements of mere antennas available on COTS WiFi receivers, in addition to a single arbitrarily placed AP as a transmitter, without knowing the

[1]For ease of understanding, we casually refer inertial measurement to motion parameter estimation of moving distance, heading direction, and rotating angle, and use both interchangeably in this paper.

SIGCOMM '19, August 19–23, 2019, Beijing, China

ACM ISBN 978-1-4503-5956-6/19/09…$15.00
https://doi.org/10.1145/3341302.3342081

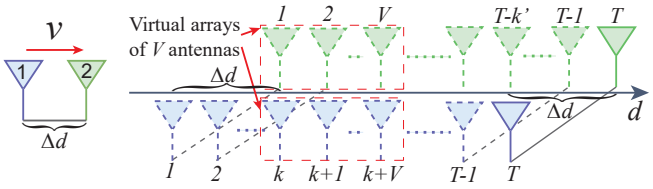


Figure 1: An illustration of virtual antenna alignment. When the antenna array moves, Antenna #1 will retrace the locations Antenna #2 has traveled and will be spatially aligned with a virtual antenna when and only when it arrives at the particular location where Antenna #2 generated that virtual antenna Δt time ago. The moving speed can thus be estimated as $v = \Delta d/\Delta t$.

AP's location or orientation. RIM passively measures the Channel State Information (CSI) of packets transmitted by the AP, without any other support from it. In contrast to many prior indoor tracking proposals that are flawed in NLOS, RIM works anywhere the AP signal reaches, be there LOS or through multiple walls.

RIM leverages an unseen opportunity with MIMO WiFi radios present in most smart hardware. We illustrate the intuition in Fig. 1. When an antenna array moves, one antenna may retrace the trajectory of another and observe the same channel when and only when it arrives at a location traveled by the preceding antenna, allowing for a possibility of "self-tracing" of the array at micro and transient scale. This observation leads to the design of RIM, structured around three components, which thwarts the conventional burdensome "outside-in" tracking solutions and enables precise motion measurement in a ubiquitous "inside-out" system:

(1) Spatial-temporal virtual antenna retracing: Take the two-antenna array as shown in Fig. 1 as an intuitive example. When the array moves, each antenna captures a channel snapshot at every point along its trajectory as if it sets up a *virtual antenna* there. The antenna travels later (*i.e.*, the following antenna, Antenna #1 in this example) will retrace every location the preceding antenna (Antenna #2) has traveled. From the time offset the following antenna takes to be spatially aligned with (*i.e.*, arrive at the location of) a virtual antenna that the preceding antenna emulated, we can estimate the moving speed by dividing the travel distance by the time offset. Here the travel distance is identical to the antenna separation, which is known and fixed and independent of how they move. Extending to 2D antenna arrays (*e.g.*, a circular array as in Fig. 2), we can track the speed in multiple directions designated by different pairs of antennas. That is, we can measure the moving distance and heading direction using a 2D array. The main challenge, however, is to detect the spatial alignment of two virtual antennas with high precision.

(2) Super-resolution virtual antenna alignment: The key insight behind highly accurate antenna alignment is that signals received at different locations undergo diverse reflecting paths and delays, resulting in a unique *multipath profile* for each location (virtual antenna). Achieving super-resolution alignment, however, is non-trivial because: 1) channel measurements on COTS WiFi are considerably noisy, 2) there is only one single measurement associated with one virtual antenna, and 3) it is performed upon measurements from different antennas with hardware heterogeneity. In RIM, we achieve sub-centimeter resolution virtual antenna alignment by three folds. First, we leverage the physics of time-reversal focusing effects in electromagnetic waves and employ an effective metric to distinguish two channel snapshots, which significantly improve location distinction [17, 39]. Second, although a single measurement may not be robustly distinctive, the alignment can be boosted by leveraging a number of virtual antennas, forming a virtual massive array. Yet different from some previous works using virtual arrays for Synthetic Aperture Radar [15], we do not need any information on how the virtual antennas are located in space. Third, we only focus on potential alignment within a short period (*e.g.*, 0.5 seconds) over a small space (*e.g.*, centimeters), for which the channel is unlikely to be changed.

(3) Precise motion reckoning: Built upon virtual antenna retracing and alignment, we devise a novel algorithm to 1) accurately and robustly pinpoint the temporal delays when two antennas are spatially aligned, 2) reliably determine which pair of antennas, among others, are aligned at a specific time, and 3) systematically integrate all information together to output moving distance, heading direction, and rotating angle if there is any.

We implement RIM on COTS WiFi chipsets and deploy it over one floor of busy office space. We evaluate the accuracy of RIM using three antennas available on commodity WiFi Network Interface Cards (NIC) and further extend its capability by designing a hexagonal array that combines two unsynchronized NICs. Experiments show that, with an unmodified WiFi receiver measuring a single unknown AP on a standard channel, RIM achieves a median error in moving distance of 2.3 cm and 8.4 cm for on-desk short-range movement and on-cart long traces respectively, 6.1° degrees mean error in heading direction, and around 30° mean error in rotating angle (corresponding to about 1.3 cm error in rotating distance). A set of case studies are conducted to demonstrate RIM's potential applications, including indoor tracking, handwriting, gesture control, and system integration with inertial sensors. With the promising performance, RIM would upend the way inertial measurement traditionally has been practiced, complementing and perhaps ultimately precluding the need for erroneous inertial sensors.

2 OVERVIEW

RIM enables inertial measurement purely based on RF signals, turning COTS WiFi radios into accurate IMUs. Specifically, it aims at measuring three dimensions of motion parameters as traditional IMUs do, yet at a much finer precision:

- **Moving distance:** The translation distance the target has moved, which is usually coarsely sensed by an accelerometer by step counting [44];
- **Heading direction:** The moving direction, which is very difficult for conventional sensors to measure [28] and thus usually assumed as the device orientation reported by a magnetometer;
- **Rotating angle:** The angle of angular rotation, typically measured by a gyroscope.

RIM estimates all these parameters for 2D motions in a universal scheme termed as *spatial-temporal virtual antenna retracing* (§3.1), and boosts the precision by a novel approach for *super-resolution*

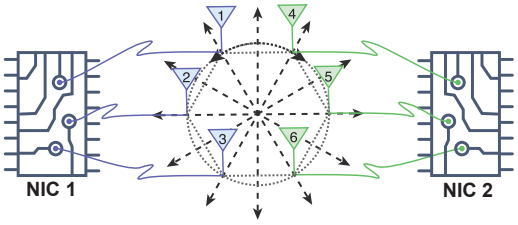


Figure 2: 6-element circular array (which is built by placing together the antennas of two commodity WiFi radios in RIM's prototype) for 2D measurement.

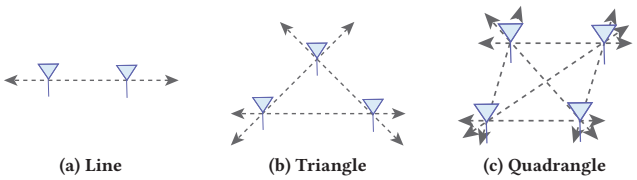


Figure 3: Examples of different antenna arrays. A (a) linear, (b) triangular, and (c) quadrangular array produces 2, 6, and 12 tractable directions at most, respectively.

virtual antenna alignment (§3.2). Then we strive to enable RIM on COTS WiFi, delivering a comprehensive system for accurate and robust inertial measurement (§4).

3 VIRTUAL ANTENNA ALIGNMENT

3.1 Virtual Antenna Retracing

Our key idea to turn a WiFi radio into an IMU lies in a novel scheme for local motion tracking, named *spatial-temporal virtual antenna retracing* (STAR). In this section, we illustrate how STAR enables measuring linear and angular motion, first in 1D and then 2D cases.

1D case: Recall the simplest scenario of a two-antenna array in 1D case, as shown in Fig. 1. When two antennas are moving along the line formed by themselves, one antenna will lead the trajectory while another immediately follows its "footprints". Whichever antenna will take a snapshot of a location it passed through by recording the CSI observations, as if it set up a virtual antenna there. Since the two antennas are moving along the same trace, one after another, the following antenna will continuously encounter the virtual phantoms of the leading one. Considering time t_k in Fig. 1, the blue following antenna arrives at the location where the green leading antenna traveled through at time t_1. By examining the arriving time difference $\Delta t = t_k - t_1$, we can derive the antenna array's traveling speed $v = \Delta d/\Delta t$, where the traveling distance during this time offset is known a priori as the separation distance Δd between the two antennas. Then by continuously aligning the following antenna with the leading antenna's "footprints", we will obtain the real-time speeds along the whole trajectory. The moving direction is, obviously, the orientation of the ray formed by the two aligned antennas.

In one word, the key to estimate motion distance and orientation is to determine (1) whether or not two antennas are aligned, and (2) if yes, what the time delay Δt is. We leave this task to the next section and first present how to resolve motion in the 2D case.

2D case: To measure distance and orientation in 2D space, we resort to 2D arrays of antennas. Similar to the 1D case, each pair of antennas enables distance measurement for the two directions of the line formed by them. Therefore, multiple pairs offer multiple directions in which we can measure moving distances. Fig. 3 illustrates several examples of antenna arrays. As shown in Fig. 3a, a two-element array (or any linear array) only supports two directions in a line. With three antennas arranged in a triangle (Fig. 3b), we can track motion along three lines (each with two moving directions). By adding one more antenna to form a quadrangle (Fig. 3c), we obtain 6 antenna pairs, providing at most 12 directions.

Ideally, m antennas will form $m*(m-1)/2$ lines in 2D space, each corresponding to a pair of antennas, yielding $m*(m-1)$ supported directions that lead to an orientation resolution of $2\pi/(m*(m-1))$. Practically, however, the resolution will be lower since some pairs may be parallel with each other and the corresponding directions become the same. For example, a square array of 4 antennas only has 8 directions, while ideally, a quadrangular array supports 12 directions (Fig. 3c). Fortunately, those parallel antenna pairs can be leveraged for augmented alignment to facilitate distance and angle measurement, as will be detailed later in Section 4.2.

The angle of rotation for angular motion is derived differently from the heading direction. For linear motion in a specific direction, only several pairs (at most 3 in the hexagonal case) of antennas would be aligned. Differently, in the presence of rotation, every adjacent pair will be aligned at the same time since all of them move along the same circle. As a consequence, we can sense rotation by detecting concurrent alignment between all adjacent antennas, and further calculate the rotating angle.

Hexagonal array: In principle, the more antennas are available, the finer distance and orientation resolution we can have. In this paper, considering that most COTS WiFi radios are equipped with 3 antennas, we prototype RIM with a hexagonal array built from two COTS WiFi cards[2], as shown in Fig. 2. Such an array provides 12 different directions in total and thus an orientation resolution of 30°. For each possible direction, there will be at least two pairs of antennas being aligned, making the estimation more robust (§4.2).

Note that our hexagonal design does not require cumbersome phase synchronization across multiple antennas or between the two WiFi NICs (§3.2). We also remark that RIM generally applies to different antenna arrays, especially the upcoming WiFi chipsets with more antennas and shorter wavelengths, which will immediately offer a better resolution in both distance and orientation.

3.2 Super-resolution Antenna Alignment

To put the idea of STAR into practice, however, is a highly challenging task that requires to accurately pinpoint a space-time point that two virtual antennas are aligned with each other, at sub-centimeter precision.

[2]We term this design as hexagonal array instead of the commonly used uniform circular array since it is a physical combination of two arrays rather than a regular phased array.

This task might seem similar to but significantly differs from traditional fingerprint matching [30]. In particular, previous fingerprinting (1) needs *a priori* calibration, (2) requires CSI to be unique over the whole space and stable over a long time, and (3) usually accumulates a number of training samples. In contrast, RIM requires no calibration and consumes only channel measurements in a narrow space (*e.g.*, a few centimeters) and in a transient period (*e.g.*, within 0.5 second). Besides, RIM needs to align two individual antennas, each with only one single channel measurement and expects the largest similarity (meaning that the two antennas are best aligned) to be only observed by the measurements recorded by the two antennas at the same space location, with a millimeter-level resolution.

We introduce two techniques to achieve sub-centimeter resolution in RIM: (1) a similarity measure, *i.e.*, Time-Reversal Resonating Strength (TRRS), for channel samples inspired by the time-reversal focusing effects [17], and (2) an approach to exploit a large number of virtual antennas as a virtual massive array for alignment. In the following, we first present a primer on time-reversal focusing effect, followed by the definition of TRRS and then the enhancement by virtual massive antennas.

Time-reversal focusing effects: Time reversal is a physics phenomenon that the energy of the transmitted signal will be focused in both space and time domains when combined with its time-reversed and conjugated counterpart. It has been studied since the 1950s [3] and later applied to and experimentally verified in the fields of ultrasonics, acoustics, light, and electromagnetism [17, 18].

To put it in the context of the WiFi channel, the received CSI, when combined with its time-reversed and conjugated counterpart, will add coherently at the intended location but incoherently at any unintended location, creating a spatial focusing effect as has been analyzed in [39, 46]. This explains, fundamentally, why multipath profiles using CSI can underpin high-resolution location distinction [5]. Therefore we introduce TRRS, a metric that quantifies the time-reversal focusing effect, as the similarity measure for CSI as follows.

Time-reversal resonating strength: The TRRS between two Channel Impulse Responses (CIRs) $\mathbf{h}_1$ and $\mathbf{h}_2$ is defined as [39]

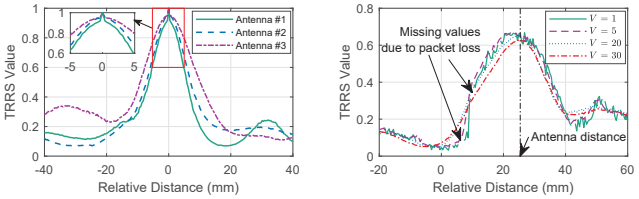
$$\kappa(\mathbf{h}_1, \mathbf{h}_2) = \frac{\left(\max_i \left|(\mathbf{h}_1 * \mathbf{g}_2)[i]\right|\right)^2}{\langle \mathbf{h}_1, \mathbf{h}_1\rangle \langle \mathbf{g}_2, \mathbf{g}_2\rangle}, \tag{1}$$

where $*$ denotes linear convolution, $\langle \mathbf{x}, \mathbf{y}\rangle$ is the inner product between vector $\mathbf{x}$ and $\mathbf{y}$, and $\mathbf{g}_2$ is the time-reversed and conjugated version of $\mathbf{h}_2$, *i.e.*, $\mathbf{g}_2[k] = \mathbf{h}_2^*[T-1-k], k = 0, 1, \ldots, T-1$.

In practice, the frequency domain Channel Frequency Response (CFR) is more often used. Equivalently, the TRRS in Eqn. 1 can be expressed for two CFRs H_1 and H_2 as:

$$\kappa(H_1, H_2) = \frac{|H_1^{\mathrm{H}} H_2|^2}{\langle H_1, H_1\rangle \langle H_2, H_2\rangle}. \tag{2}$$

If H_1 and H_2 are both normalized, then the TRRS becomes the square of their inner product, *i.e.*, $\kappa(H_1, H_2) = |H_1^{\mathrm{H}} H_2|^2$. Obviously, $\kappa(H_1, H_2) \in [0, 1]$, and $\kappa(H_1, H_2) = 1$ if and only if $H_1 = cH_2$ where $c \neq 0$ is any complex scaling factor. Note that while previously many heuristic metrics were used to compare CSI, the TRRS exploits the physics of time-reversal focusing effects in an uncomplicated form.



(a) Self-TRRS (b) Cross-TRRS

Figure 4: Spatial resolution of TRRS. (a) Three antennas are moving at a constant speed and the TRRS for each antenna with respect to itself is calculated. As seen, the TRRS drops immediately (significantly by up to 0.3) when the antenna moves for a few millimeters, and monotonously decreases within a range of about 1 cm. (b) The spatial decay holds for cross-antenna TRRS, especially with virtual massive antennas, although the absolute values decrease.

CSI measured on COTS WiFi is well-known to contain phase offsets, including carrier frequency offset (CFO), sampling frequency offset (SFO), and symbol timing offset (STO) due to unsynchronized transmitters and receivers, in addition to initial phase offset caused by the phase locked loops. As can be seen from Eqn. 2, the impact of the annoying initial phase distortion, which could only be manually calibrated with a wired connection or an external clock previously [43], is eliminated by taking the absolute value in Eqn. 2. We calibrate the other linear offsets by using the sanitation approach employed in [13].

Time-reversal focusing effects will be intensified with larger bandwidths. To facilitate the robustness and uniqueness of TRRS, we exploit spatial diversity attributed by multiple transmit antennas to gain larger effective bandwidth. Specifically, suppose there are N antennas on the AP. The CSI measured on the ith receive antenna at time t is $H_i(t) = \{H_{i,1}(t), H_{i,2}(t), \cdots, H_{i,N}(t)\}$ where $H_{i,k}(t)$ is the CSI between the ith receive antenna and the kth transmit antenna. We then take the average TRRS of the ith and jth receive antenna as

$$\bar{\kappa}\left(H_i(t_i), H_j(t_j)\right) = \frac{1}{N}\sum_{k=1}^{N} \kappa\left(H_{i,k}(t_i), H_{j,k}(t_j)\right). \tag{3}$$

By the above definition, we avoid the need of synchronizing two antennas, but instead take the average of their individually calculated TRRS values.

Virtual massive antennas: Mainstream APs only have a few antennas, limiting the resolution and robustness of the average TRRS in Eqn. 3 to measurement noise. To boost super-resolution alignment, we propose to leverage a number of virtual antennas emulated by the sequence of channel snapshots recorded by a moving antenna, forming a virtual massive antenna array whose size is the number of channel snapshots. As shown in Fig. 1, we extend the multipath profile of an antenna i at time t from a single snapshot $H_i(t)$ to a sequence of samples $P_i(t) = [H_i(t+k), k = -V/2, \cdots, V/2]$, where $H_i(t+k)$ indicates the channel measurement emulating the virtual antenna set up by antenna i at time $t+k$, and V is the number of virtual antennas.

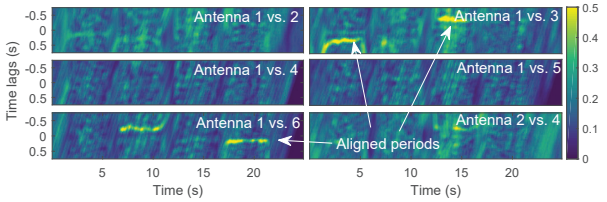


Figure 5: Alignment matrices of a square-shape trajectory. The aligned pairs of the hexagonal array are 1 vs. 3 followed by 1 vs. 6, and then again 3 vs. 1, 6 vs. 1 in turn. Other pairs in parallel with one of the above are not shown.

Accordingly, we calculate the TRRS with V virtual massive antennas as

$$\kappa\left(P_i(t_i), P_j(t_j)\right) = \frac{1}{V}\sum_{k=-V/2}^{V/2} \bar{\kappa}\left(H_i(t_i+k), H_j(t_j+k)\right), \quad (4)$$

where $H_i(t_i-k)$ and $H_j(t_j-k)$ denote the respective virtual antenna placed at certain space location by the ith and jth antenna at time t_i-k and t_j-k. By using the virtual massive antennas, we largely boost the effective bandwidth of the multipath profile and thus enhance the time-reversal focusing effect, or equivalently, attain highly precise antenna alignment. As shown in Fig. 4, the TRRS touches the maximum only when two antennas are closest to each other (precisely aligned) and drops even when they are separated by several millimeters.

TRRS matrix: To pinpoint the precise temporal delay when two antennas i and j are spatially aligned, we match the multipath profile of antenna i against those of antenna j throughout a sliding window. Consider a window of length $2W$, the TRRS vector is calculated as $\mathbf{G}_{ij}(t) = [\kappa(P_i(t), P_j(t-l)), l=-W,\cdots,W]^{\mathrm{T}}$ where l denotes the time lags. Thus if the antennas move for a period of T, we obtain a TRRS matrix

$$\mathbb{G}_{ij} = [\mathbf{G}_{ij}(t_1)\ \mathbf{G}_{ij}(t_2)\ \cdots\ \mathbf{G}_{ij}(t_T)]. \quad (5)$$

Fig. 5 illustrates an example of the TRRS matrix, which is also termed *alignment matrix* hereafter. The window length W should be larger than the expected time delay for two antennas to be spatially aligned. A larger window will incur higher computation overhead and is not needed.

In RIM, we calculate such TRRS matrix for every pair of antennas (Fig. 5). The motion parameters are then estimated by identifying the aligned pairs from the TRRS matrices and continuously estimating the time delays, and accordingly the moving speed, as detailed in the next section.

Deviated retracing: So far we assume that the antenna array is moving along a certain direction in which at least two antennas will be well aligned with each other. In practice, however, the device may be swinging in directions slightly deviated from the exact aligned line, as shown in Fig. 6a, where the array is moving along a direction that deviates an angle of α from their aligned direction, *i.e.*, the horizontal line. Fortunately, we realize that noticeable TRRS peak still exists, albeit weaker, for two antennas that are close enough yet not exactly aligned in case of deviation angles. Hence virtual antenna alignment is still feasible since we only focus on the relative TRRS peaks within a window instead of the absolute values, circumventing the impacts of reduced TRRS values due to deviated retracing.

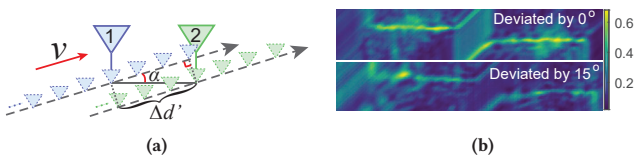


Figure 6: Antenna alignment in case of deviation retracing. (a) Deviated retracing results in (b) much weaker but still evident TRRS peaks. The example shows a forward-then-backward move.

In theory, the width of the TRRS peak without ambiguity is about $\delta = 0.2\lambda$ [46]. Recall Fig. 4, the TRRS peak without ambiguity is about $\delta = 10$ mm, in accord with the theoretical value. Thus, given an antenna separation Δd, the proposed antenna alignment can theoretically tolerate a maximum deviation angle of $\alpha = \arcsin(\frac{\delta}{\Delta d})$, approximately 24° for $\Delta d = \frac{\lambda}{2}$. As shown in Fig. 6b, real-world experiments show that RIM can confidently tolerate as large as 15° deviation, which suffices to cover the complete 2D plane with 6-element circular array. As a result, RIM can track motions in any direction within the plane of the antenna array. For the same reasons, RIM does not require the antenna array to be perfectly leveled on the same plane. As will be shown in our experiments in §6, our designed array with imperfections in antenna arrangement yields good performance.

As shown in Fig. 6a, in case of deviation, the "deviated" antenna separation becomes $\Delta d' = \Delta d \cos\alpha$. Since we have no information about the deviation angle α in advance, we directly let $\Delta d' \approx \Delta d$ in RIM, leading to an overestimated factor of $1/\cos\alpha$ in distance estimation. In terms of our 6-element circular array, the overestimated error will be 1.20% on average, assuming the moving directions are uniformly distributed in [0°, 360°], and reaches the worst of 3.53% when the deviation angle $\alpha = 15°$, which is tolerable in practice. The deviation angle may be resolvable by quantitatively comparing the reduced TRRS to the expected value of perfect alignment, which we keep as future work.

The above discussion assumes that the Rx is moving, listening to a static Tx. RIM also applies to the opposite case when the Tx is moving with a static Rx measuring CSI due to channel reciprocity [39]. In either case, RIM estimates the motion of the moving device, be it the Tx or Rx. This is a useful feature for certain applications, *e.g.*, in drone tracking where the drone may serve as a mobile AP (the Tx) rather than an Rx.

4 MEASURING MOTION

The millimeter resolution antenna alignment underpins inertial estimation in centimeter accuracy in practice. This section presents how to measure motions on this basis. We first examine whether movement exists. If yes, we then attempt to obtain the alignment matrix of each antenna pair and accordingly determine when and which pairs are well aligned. From the aligned pairs, we will derive the distance, heading direction, and angle of rotation if there is any.

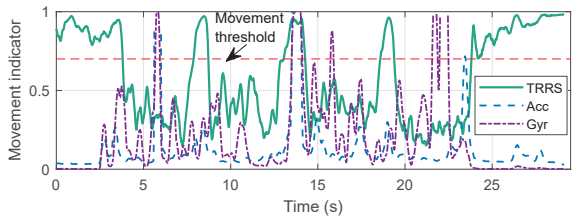


Figure 7: Movement detection. RIM is more robust than accelerometer (Acc) and gyroscope (Gyr) that both fail to detect the three transient stops during movements.

4.1 Detecting Movement

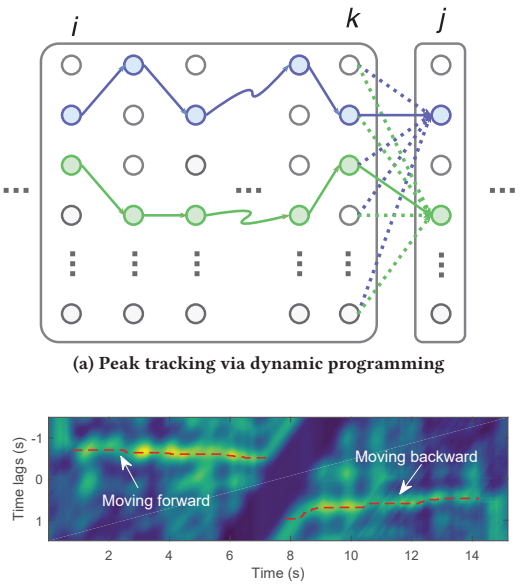
It is straightforward to detect antenna movement from TRRS, since it is, with the virtual massive antennas, highly sensitive to arbitrary location changes. Yet for movement detection, we only need to examine the TRRS for one single antenna based on its own measurements. Specifically, we calculate $\kappa(P_i(t), P_i(t - l_{\text{mv}}))$, the TRRS between the current measurement and the one l_{mv} seconds ago. l_{mv} is the time lag chosen as a conservative period during which location would change by at least millimeters if motion happens. For example, if the antenna is moving at 1 m/s, a time lag of 0.01 second would expect a movement of 1 cm. Fig. 7 illustrates an example of a stop-and-go trace. As seen, there is a clear gap between the TRRSs for movement and non-movement. Thus we can apply a threshold to detect movement, as indicated by the red horizontal line in Fig. 7. The thresholding works generally because the TRRS based on the same antenna always touches close to 1 when it is static and decreases with movement.

4.2 Tracking Alignment Delay

Given an alignment matrix, we need to identify the TRRS peaks of interests that indicate the time lags of antenna alignment. For an ideal case, we can pinpoint the time lags corresponding to the maximum values in each column of the TRRS matrix as the alignment delays. In practice, however, the maximum values could deviate from the true delays due to measurement noise, packet loss, or wagging movements, etc, and thus an approach is needed to robustly track the peak sequence corresponding to the alignment delays. To this end, we propose a novel algorithm based on dynamic programming, which has been widely used to solve pitch tracking problems in signal processing [48]. When applying in our case, however, it still needs elaborate design.

For clarity, we simplify the notation of the TRRS matrix $\mathbb{G}$ for two antennas from time t_1 to t_T as $Q = [\boldsymbol{q}_i]$ where $\boldsymbol{q}_i = [q_{il}], i \in [1, T], l \in [-W, W]$. As shown in Fig. 8a, now suppose we want to search the best path of the TRRS peaks from time point t_i to t_j, denoted as $\boldsymbol{q}_i \rightsquigarrow \boldsymbol{q}_j$. Define a score of the optimal path $\boldsymbol{q}_i \rightsquigarrow \boldsymbol{q}_j$ that ends at point q_{jn} as $S(\boldsymbol{q}_i \rightsquigarrow q_{jn})$. Our approach is to search all $2W$ possible candidate paths $\boldsymbol{q}_i \rightsquigarrow q_{jn}$ that end at time lags $n \in [-W, W]$ and then select the best one among all.

Suppose we already have all the optimal paths from t_i to t_k, each ending at $q_{kl}, l \in [-W, W]$, and the optimal paths from t_k to t_j,



(a) Peak tracking via dynamic programming

(b) An example of peak tracking result

Figure 8: TRRS peak tracking. The alignment peaks are accurately and robustly identified regardless of measurement noises and imperfect retracing.

each starting at q_{kl} and all ending at q_{jn}, then we have

$$S(\boldsymbol{q}_i \rightsquigarrow q_{jn}) = \max_{l \in [-W, W]} \{S(\boldsymbol{q}_i \rightsquigarrow q_{kl}) + S(q_{kl} \rightsquigarrow q_{jn})\}, \quad (6)$$

which stands for the score of the optimal path from $\boldsymbol{q}_i \rightsquigarrow q_{jn}$ since $S(\boldsymbol{q}_i \rightsquigarrow q_{kl})$ and $S(q_{kl} \rightsquigarrow q_{jn})$ are both the scores of the respective optimal paths. When $k = j - 1$, the score of the peak transition between subsequent columns q_{kl} and q_{jn} is calculated as

$$S(q_{kl} \rightsquigarrow q_{jn}) = e_{kl} + e_{jn} + \omega C(q_{kl}, q_{jn}) \quad (7)$$

where e_{kl} and e_{ln} are the TRRS values at q_{kl} and q_{ln} respectively. $C(q_{kl}, q_{jn})$ is the cost for stepping from q_{kl} to q_{jn} and is simply defined as $C(q_{kl}, q_{jn}) = \frac{||l-n||}{2W}$. ω is a negative weighting factor for the cost. The designation of the cost function punishes jumpy peaks. The rationale is that in general cases the moving speed (thus the aligned delays) will not fluctuate too much within successive measurements.

Once we have the scores for paths $\boldsymbol{q}_i \rightsquigarrow q_{jn}$ for any $n \in [-W, W]$, the best path from $\boldsymbol{q}_i$ to $\boldsymbol{q}_j$ can be found as $\boldsymbol{q}_i \rightsquigarrow q_{jn^\star}$ where

$$n^\star = \arg\max_{n \in [-W, W]} \{S(\boldsymbol{q}_i \rightsquigarrow q_{jn})\} \quad (8)$$

The entire path of peaks can then be easily identified by tracing back the previous steps from $q_{jn^\star}$ until the starting point t_i. Fig. 8b depicts an example of the peak tracking results for a back-and-forth movement consisting of two periods of alignment.

In principle, the above peak tracking should be performed on every pair. Two steps, however, are taken to optimize the complexity: (1) Antenna pairs that are very unlikely to be aligned are skipped

(See §4.3); (2) In our implementation with 6-element array, we facilitate it by averaging the alignment matrix of parallel isometric antenna pairs (*e.g.*, antenna pairs (1, 4) and (3, 6), (2, 4) and (3, 5) as in Fig. 2) and conduct peak tracking on the averaged matrix, which is augmented since the two pairs have the same alignment delays.

4.3 Detecting Aligned Pairs

For the sake of robustness, we detect aligned antenna pairs in two steps, a pre-check before peak tracking, and a post-check after that.
Pre-detection: In practice, we perform a pre-detection operation to preclude the unaligned pairs before peak tracking. In particular, for a specific period, we consider only antenna pairs that experience prominent peaks most of the time as prospective candidates and exclude the others that are unlikely aligned. The peak tracking is then only employed on the selected candidate pairs.
Post-detection: After peak tracking, we obtain a path of identified peaks for each antenna pair. We then further examine the extracted paths, accounting for the continuity, TRRS values, smoothness, and orientations they indicate, to ultimately confirm the most likely aligned pairs.

4.4 Putting It All Together

Suppose the ith and jth antennas are detected to be aligned at time t, with a separation distance of Δd_{ij} and an alignment delay of $\Delta l_{ij}(t)$. Then we can measure the following results:
(1) Moving distance: The moving speed $v(t)$ at time t is derived as $v(t) = \Delta d_{ij}/\Delta l_{ij}(t)$. The moving distance can then be simply derived by integrating the instantaneous speed over time, *i.e.*, $d(t) = \int_0^t v(\tau)\mathrm{d}\tau$. Here we approximate $v(t)$ to be constant during the period of $\Delta l_{ij}(t)$, which is reasonable since $\Delta l_{ij}(t)$ is very short, *e.g.*, <0.5s even the moving speed is as low as 0.05m/s (given $\Delta d_{ij} = \lambda/2$) and even shorter with larger speeds. The varying speed will be captured by continuous estimation.
(2) Heading direction: It is straightforward to obtain the moving direction when we know the aligned antenna pairs. Specifically, $\theta(t)$ is assigned as the direction of the ray pointing from antenna i to antenna j if $\Delta l_{ij}(t) \geq 0$, and the opposite direction if $\Delta l_{ij}(t) < 0$.
(3) Rotating angle: At a certain moment, if every adjacent pair of antennas is aligned simultaneously, then a rotation occurs. For a very short window, we can assume that the rotation happens in place. The angle of rotation is estimated from the rotated distance of each antenna as $\Delta\theta = R/r$, where r denotes the radius of the circular array and R is the arc length estimated by the rotating distance. In the case of in-place rotation, we can estimate an individual speed from each pair of adjacent antennas. Thus we can use the average speed for rotating distance calculation. Note that the effective antenna separation for rotation becomes $\frac{\pi}{3}\Delta d$, the arc length one antenna needs to travel to hit another.

The above estimates are all smoothed and then integrated to recover the relative moving trajectory.

5 IMPLEMENTATION

Hardware: We implement RIM using commercial WiFi NICs. Fig. 9 illustrates our hardware equipment. We mainly use Qualcomm Atheros 9k series chipset, which is attached to an Intel Galileo Gen2 microcontroller board equipped with a Bosch Sensortec BNO055 sensor unit. We modify the driver to report CSI on Linux. The Galileo system supports at the most 200Hz sampling rate. To study higher sampling rates, we employ the 802.11 CSI Tool [10] for the Intel 5300 WiFi card equipped on laptops. In our evaluation, we employ two different antenna arrays: a 6-element circular array combining two standard WiFi radios (Fig. 9(a)), and a 3 antenna linear array that is directly available on a COTS chipset. Adjacent antennas are spaced at a half wavelength distance (2.58 cm), which is also the arrangement preferred in commodity APs. Both line dipole antennas and chip antennas are used and tested. For the Galileo device, we run another board with the same NIC as the AP (Fig. 9(d)). While for Intel 5300, a laptop acts as the AP. In both cases, the AP is equipped with three antennas and is set to a broadcast mode on a 5GHz channel.

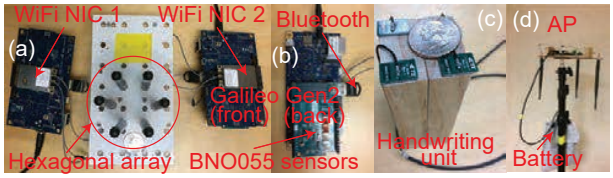


Figure 9: Hardware setup of RIM prototype.

Software: We implement all algorithms in MATLAB, mainly for micro benchmark analysis. For CSI collection on Galileo Gen2, we develop a tool in C++. To assess the application in indoor tracking (§6.3.3), we further build a real-time system in C++, including the CSI collection running on Galileo Gen2, and key algorithm components plus a Windows GUI that displays the tracking results.
Packet synchronization and interpolation: RIM does not require phase synchronization among antennas[3]. Due to potential packet loss, however, it is needed to synchronize CSI measurements (*i.e.*, packets) on two NICs for cross-antenna mapping (and only needed for our implementation using two commercial NICs). In RIM, we accomplish this by letting the AP broadcast packets, using it as a coarse external clock. In particular, two packets with the same sequence number are synchronized, since they are from the same broadcast packet on the AP and thus received simultaneously, ignoring the insignificant propagation delay. In the case of packet loss, a null CSI is inserted.

We would like to point out RIM's packet synchronization is completely different from the requirements of precise phase synchronization of previous works like [13, 26, 43]. All these works utilize phased array antennas for geometrical channel measurements. Instead, we merely need packet-level synchronization, which could be skipped shortly in the future as one NIC will come with more antennas, *e.g.*, 60GHz chipsets.
Minimum initial motion: There is a minimum requirement on the initial moving distance or rotating angle below which RIM cannot measure the motion. The minimum moving distance is, intuitively, the separation distance Δd between the two antennas. From the moment of starting moving, only after a translation of at least Δd will the following antenna "hit" the leading antenna, for the first time. In practice, to reimburse this "blind" period, the

[3]The linear phase calibration based on [13] for calculating TRRS in Eqn. 2 is conducted for individual antenna independently.

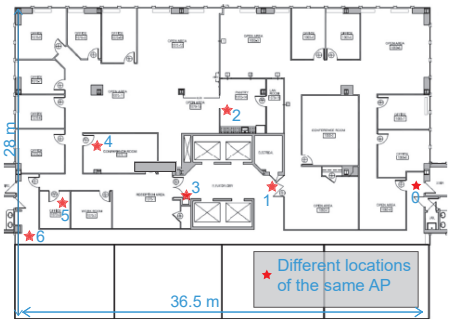


Figure 10: Testbed environment. The AP is tested at different locations, as marked by red stars.

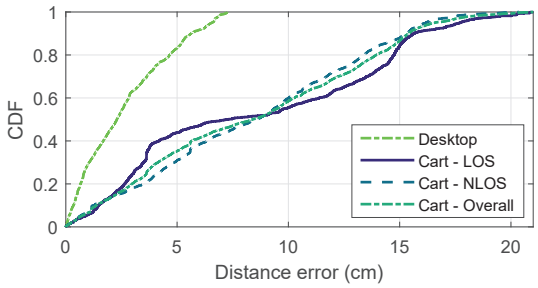


Figure 11: Accuracy of moving distance.

estimated distance will be compensated by Δd. Note that such minimum requirements on the initial motion is different from RIM's motion resolution. Once two antennas are aligned for the first time, RIM will continuously and precisely track motion after that.

6 EVALUATION

6.1 Methodology

We conduct experiments in an office environment to evaluate RIM, including system performance analysis (§6.2) and application study (§6.3).We deploy a single AP (Fig. 9(d)) to cover the experimental areas of $> 1,000$ m^2, as shown in Fig. 10. To demonstrate RIM's advantages in through-the-walls measurements, we by default place the AP at the farthest location #0, a corner of the floor. We conduct motion measurements at different locations over the floorplan (except for some areas where we do not have access), covering areas both near to, and far away (as far as 40 meters) from the AP. During the experiments, the AP keeps broadcasting packets at 200Hz on a 40MHz channel in the 5GHz band.

To obtain ground truth, we set up a camera-based tracking system. Specifically, the target is marked with high contrast color and tracked by the cameras. The motion trace is first calculated in the pixel frames and then converted to 2D world coordinates. The outputs are synchronized with RIM's estimates by the initial point when the target starts moving and are pairwise compared subsequently. There might be slight time offsets, which do not favor our evaluation.

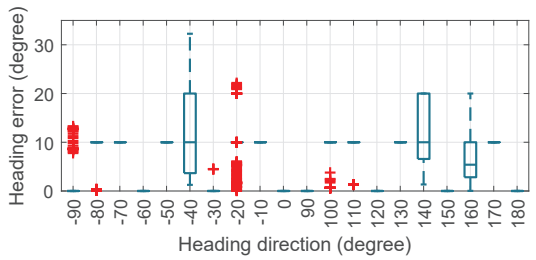
6.2 Micro Benchmarks

We now evaluate the overall performance in estimating moving distance, heading direction, and rotating angle, and study how various factors impact accuracy. Unless otherwise stated, the device is moving at a speed of about 1 m/s, and the lengths of traces for analysis all exceed 10 meters.

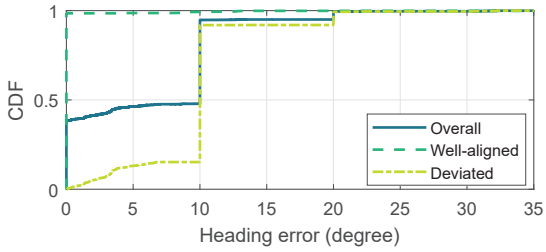
6.2.1 Accuracy of moving distance. To fully understand RIM's centimeter accuracy in moving distance estimation, we use a 3-antenna linear array present in COTS WiFi and move it roughly along lines in two scenarios: 1) Desktop: we move the array on a desk surface for traces around 1 m; 2) Cart: we put the array on a cart and push it straight forward by more than 10 meters in different areas. As shown in Fig. 11, RIM achieves a median error of 2.3 cm for the more stable, thus better aligned, desktop moving. For the less controlled cart movement, it yields 8.4 cm median error. Specifically, it yields 7.3 cm median accuracy in LOS scenarios and, more importantly, a similar accuracy of 8.6 cm in complex NLOS conditions. The 90%tile and maximum errors are under 15 cm and 21 cm, respectively. We do not include results using accelerometer because it easily produces errors of tens of meters [44]. To the best of our knowledge, no existing system can achieve comparable performance under similar settings (*i.e.*, NLOS over a large area using a single unknown AP). Improved tracking accuracy of RIM is due to its novel method in utilizing rich multipath profiles as distinct spatio-temporal virtual antennas.

6.2.2 Accuracy of heading direction. We study RIM's heading direction estimation accuracy by moving the hexagonal array in diverse directions. In particular, we traverse a 90° range with an increased step of 10°, together with each of their opposite directions. For each direction, we move the device for about 1 meter and examine the heading direction estimation errors. As depicted in Fig. 12a, RIM correctly identifies the closest heading direction that it can resolve for most cases, except for a few outliers (*e.g.*, in direction -40° and -20°). Fig. 12b further integrates the results for all directions, which shows that >90% of heading errors are within 10°, with an overall average accuracy of 6.1°. Most of the estimates are either correct without any error (*e.g.*, well-aligned directions) or with 10° errors (*e.g.*, deviated directions), because RIM only resolves a set of discrete directions that are integral multiples of 30°. We will further demonstrate RIM's heading estimation performance for free movements in real applications in the next section. Again, the magnetometer is not compared here because it can not address heading direction.

6.2.3 Accuracy of rotating angle. To evaluate rotating angle accuracy, we attach the hexagonal array on a fan and manually rotate the array for different angles, ranging from 30° to 360°. We repeat each angle for 10 times and calculate the relative errors. As shown in Fig. 13, RIM tracks rotating angles with a median error of 30.1° (about 17.6% in relative error) for the total 80 tests, corresponding to an error of merely 1.3 cm in arc lengths (*i.e.*, moving distances). Unfortunately, gyroscope produces much better results in this case. The currently achieved performance of RIM is limited by that the antenna separation is at the same magnitude as the array radius. Given that the adjacent antenna separation and radius are both half wavelength for the hexagonal array, even a 0.5 cm error in the



(a) Heading errors w.r.t directions



(b) Overall accuracy of heading direction estimation

Figure 12: Accuracy of heading direction.

distance could lead to 11° error in rotating angle. Nevertheless, the results still validate the feasibility of RIM's rotating angle sensing capability, which will be drastically improved by denser and smaller antenna array in the future.

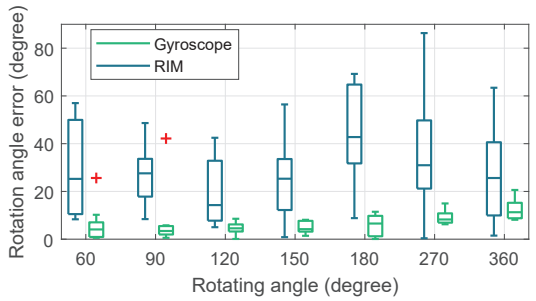


Figure 13: Accuracy of rotating angle.

6.2.4 Coverage and impacts of AP locations. RIM can work anywhere the AP signals can reach. To study whether AP location affects its performance, we place the AP at different locations, as marked in Fig. 10, and perform distance measurements in the middle open spaces. As shown in Fig. 14, RIM achieves consistently high median accuracy of less than 10 cm for any AP location, be it close LOS or far away through multiple walls and pillars. The results show RIM is truly multipath resilient and achieves best-ever coverage, which allows it to work wherever there are WiFi signals.

6.2.5 Impact of accumulative distance. Inertial sensors are known to suffer from accumulative errors over long trajectories. As a relative tracking approach, however, errors may accumulate. Thus it

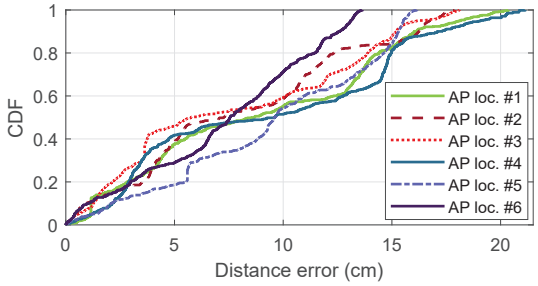


Figure 14: Impact of AP location.

would be interesting to examine how RIM performs with respect to different movement lengths. We move for about 10 m for multiple times and calculate the tracking errors during trace progressing. As shown in Fig. 15, the median errors in moving distance vary from about 3 cm to 14 cm, which do not considerably accumulate over long distances. The performance is further confirmed by longer traces as shown later in Fig. 20. The performance gains attribute to RIM's high precision in speed estimation, which does not drift over time. In practice, however, we note an error correction mechanism may be needed to overcome potential accumulative errors (See an example of tracking in Fig. 21).

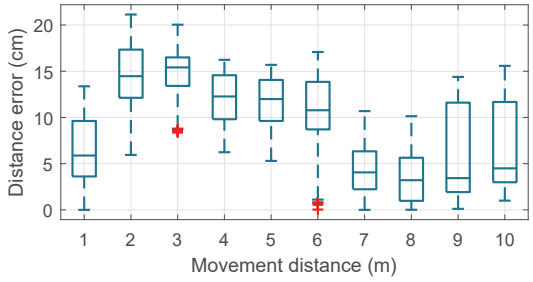


Figure 15: Impact of movement distances.

6.2.6 Impact of sampling rate. Sufficient sampling rate is required to achieve high accuracy in RIM. Thus we downsample the CSI from 200Hz down to 20Hz and rerun the distance tracking. The results are integrated in Fig. 16. As expected, the distance tracking accuracy increases with higher sampling rates. For a moving speed of 1 m/s, a sampling rate of 20Hz or 40Hz is not enough, which results in about a displacement of 5 cm per sample. To ensure sub-centimeter displacement within one sample, at least 100 Hz is needed for a speed of 1 m/s. Higher sampling rate may further improve the accuracy, yet the improvement would be marginal compared to the computation overhead incurred. In general, the target moving speed is the major factor determining the required sampling rate. Faster speeds demand higher sampling rates to gather adequate samples during a movement of Δd. Higher sampling rates, however, are not requested by factors like environmental dynamics.

6.2.7 Impact of virtual antenna number. Using virtual massive antennas is a key to boost RIM's resolution. Thus we study how

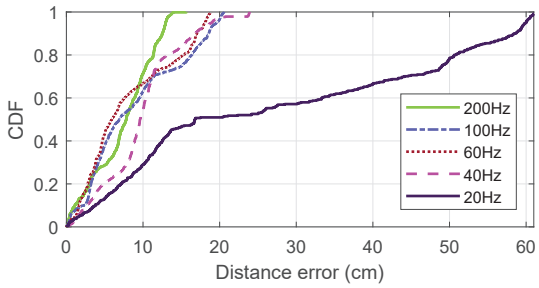


Figure 16: Impact of sampling rate.

the number of virtual antennas impacts accuracy. As shown in Fig. 17, we increase the virtual antenna number V from 1 to 100 and calculate the distance errors. The results show the median accuracy decreases from about 30 cm to 10 cm when V increases from 1 to 5, and further reduces to 6.6 cm when $V = 100$. In practice, a number larger than 30 should suffice for a sampling rate of 200Hz and should be larger for higher sampling rate and smaller for lower.

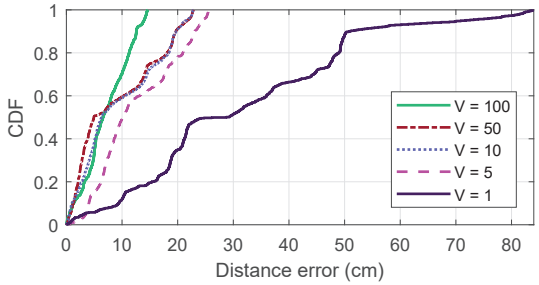


Figure 17: Impact of virtual antenna number.

6.2.8 Robustness to environmental dynamics. Building upon spatial multipath profiles, it is interesting to see if RIM is immune to common environmental dynamics, *e.g.*, walking humans. Considering our experiments, there are always at least two persons doing the experiments, in addition to others performing normal activities around the environment. Imagine when a user moves the device (*e.g.*, pushing a cart), both his/her body and the cart are moving closely to the receiver, altering the multipath distributions. Yet the above results show robustness to these kinds of dynamics. The reasons are two-fold: 1) There are a number of multipaths indoors [8], coming from different directions. As a walking human will only change part of them, the TRRS behavior underpinning RIM still holds. 2) RIM does not rely on absolute TRRS, which may vary with environmental dynamics.

6.2.9 System complexity. The main computation burden lies in the calculation of TRRS. For every sample, RIM needs to calculate TRRS over a window of $2W$, leading to $m \times (m-1) \times W$ values in total for an array of m antennas. Since the MATLAB code is not optimized for speed, we measure the C++ system on Surface Pro, which is equipped with Intel Core i7 4650U and 8 GB RAM. RIM's core modules run in real-time, using around 6% of CPU and about 10 MB of RAM.

6.3 Applications

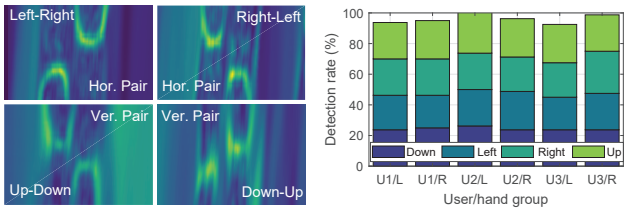
RIM has a range of potential applications in indoor tracking, handwriting, gesture control, movement detection, VR headset tracking, drone tracking, etc. In this section, we evaluate RIM in three application scenarios. Note that existing state-of-the-art in motion tracking [13, 14] may also support some of these applications under their favored settings, yet at considerably high costs. However, we do not compare with them since they do not work in the same conditions as RIM (*i.e.*, single AP with unknown location and NLOS cases with rich multipaths).

6.3.1 Desktop Handwriting. With RIM's fine precision in motion tracking, it supports handwriting on a whiteboard or desk. To show a proof-of-concept scenario, we ask a user to move the antenna array by freely writing some letters on a desk surface. Fig. 18 depicts some examples of the recovered letters, with comparison to the ground truths captured by a camera system. Due to lack of precise timing-synchronization between the trajectories tracked by RIM and the camera, we approximate the tracking error as the minimum projection distance from the estimated location to the trajectory. As seen, RIM accurately reconstructs not only relatively straight segments but also curved strokes in different directions, resulting in recognizable letters. Specifically, the mean error of the trajectories of the letters written in Fig. 18 is 2.4 cm, similar to the errors of straight line trajectories (Fig. 14). Albeit currently the antenna array may be too large to act as a "pen", we demonstrate our technology will directly shape it for writing when smaller antennas with shorter wavelengths become ubiquitous in the near future.

Figure 18: Examples of handwriting using RIM. The trajectories in green are estimates of RIM, while the gray ones are the ground truths captured by a camera.

6.3.2 Gesture Recognition. We demonstrate a gesture control application by integrating RIM into a pointer-like unit, as shown in Fig. 9(c). To make it compact, we use one WiFi NIC with three small chip antennas arranged in an "L" shape. The experiments involve three users, each performing four different gestures (moving towards left/right/up/down and then back) for 20 times with their left and right hand, respectively. In total, we collect 480 actions for testing. As shown in Fig. 19a, RIM will observe speed in one direction in which the user's hand moves towards, immediately followed by a speed in the opposite direction when the hand moves back, from a specific pair of antenna depending on the moving direction. We use this information to detect and identify a gesture.

Fig. 19b illustrates the recognition results. Since all detected gestures are all correctly recognized, we only plot the detection rates. As shown, RIM achieves an average detection accuracy of 96.25% for different gestures and users, with 23 miss detections (4.79%) and 5 false triggers (1.04%) over the total of 480 tests. The



(a) Gesture pattens (b) Detection rates

Figure 19: Gesture recognition. (a) Different gestures produce distinct patterns in alignment matrices on certain antenna pairs. (b) Detection rates for three users with their left/right hands.

false trigger rate is less than the miss detection rate, which is also favored in practical cases because one can simply repeat the gesture for miss detection while a false trigger is annoying. Moreover, the performance is consistently high across different users and actions, and left/right hands. The results demonstrate promising gesture capability that can, *e.g.*, turn a smartphone into a presentation pointer.

6.3.3 Indoor Tracking. We particularly implement two different cases to extensively demonstrate RIM's capability for indoor tracking. During the tracking test, the AP is placed at location #0.

Pure RIM for indoor tracking: We first deploy RIM with hexagonal array as a sole tracking system over the whole floorplan in Fig. 10. The antenna array is put on a cart, which is pushed by a user along different traces. To particularly show RIM in the case of sideway movements (*i.e.*, changing heading direction without turning), which are common in industrial Automated Guided Vehicles (AGVs), we directly move the cart sideway, instead of making a turn, in this experiment. Fig. 20 illustrates two example trajectories, about 36 m and 76 m respectively with multiple sideway movements. As seen, the trajectories are accurately tracked and no significant errors accumulate, even for very long traces over a large area. Note that conventional inertial sensors, gyroscope, and magnetometer, fail to capture such direction changes because there is no turning, meaning that the device orientation keeps the same although the heading direction has changed.

RIM integrated with inertial sensors: RIM has superior advantages in moving distance estimation, even with only three antennas. To fully reveal its potential with COTS WiFi devices with only one NIC, we implement a system by integrating RIM's distance measurement with direction information reported by the gyroscope. Specifically, we use the Galileo Gen2 board equipped with one WiFi NIC of three chip antennas and the BNO055 sensor unit.

We test the integrated tracking system by putting the device on a cart and pushing it throughout the whole floor, as the way a customer pushes a shopping cart in a mall, or a smart home robot moves around. Fig. 21 illustrates an example tracking result, with an initial location and direction given in advance. As seen, while the distances, measured by RIM, are highly accurate, the direction information obtained by inertial sensors may suffer from large errors. To put it into a practical system, we implement a particle filter (PF) to leverage the geometric constraints provided by the digital floorplan to correct errors. The PF will discard every particle that hits a wall and let others survive. The result, with PF based enhancement, gracefully reconstructs the real trajectory, as shown in Fig. 21. Relying on only one single AP with unknown information, the encouraging results demonstrate the great potential of RIM in enabling ubiquitous and accurate indoor tracking system especially for robots and objects, a long-standing challenging problem that has attracted numerous research efforts.

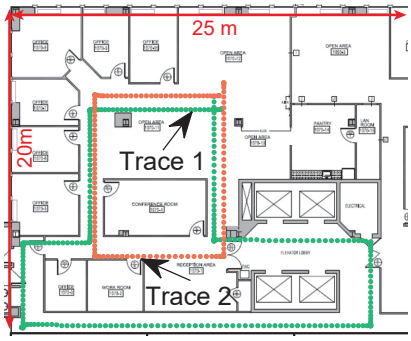


Figure 20: Tracking by sole RIM. Both traces contain sideway movements that inertial sensors do not support.

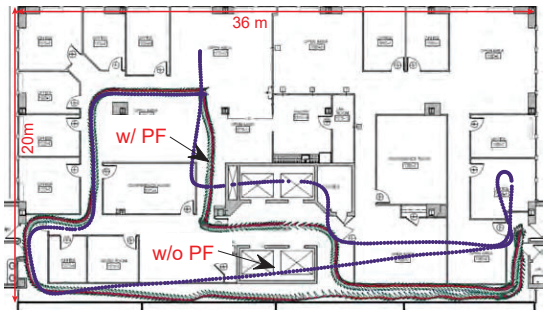


Figure 21: Tracking by RIM integrated with sensors. The results are obtained by fusing the distance estimates by RIM and direction estimates by inertial sensors.

7 DISCUSSIONS AND FUTURE WORK

RIM is an early step towards ubiquitous and precise RF-based inertial measurements, and there is obviously room for continued research in various perspectives.

Antenna array: The current prototype of RIM employs two COTS WiFi cards. As WiFi technology matures with many antennas and at higher frequencies, *e.g.*, 60 GHz, one single radio will be abundant and the size will be smaller to be embedded in mobiles. For the current prototype, RIM with its relatively large form factor is already attractive especially for dedicated interactive gaming devices and various industrial applications, *e.g.*, tracking carts and machines.

Packet loss: Ideally, uniformly sampled CSI offers the best performance of RIM. In practice, RIM can tolerate packet loss to a certain extent by interpolation. However, a relatively clean channel is needed to ensure delightful performance.

Angle resolution: Currently, RIM only exploits discrete directions defined by the antenna array (*e.g.*, 30° resolution with a 6-antenna uniform circular array). Noticing that the TRRS decreases differently with respect to different deviation angles, we believe it is promising to investigate finer-granularity directions in continuous space by leveraging the geometric relationship of adjacent antenna pairs.
Limitation of rotating angle: The current prototype of RIM can only sense in-place rotation with coarse resolution, and is not able to monitor the rotating angle of swinging turns (*i.e.*, move while turn). It remains open to exploring more general and accurate angular motion measurement based on RIM. Furthermore, although we have demonstrated the ability to track movements during turnings in §6.3.3 by integrating inertial sensors, it is interesting to investigate simultaneous translational and rotational motions using RIM only.
3D motion: While RIM achieves better precision and robustness, it cannot measure 3D motion like traditional inertial sensors. Building a specialized 3D array will immediately break the limitation, which may be impractical in some scenarios but still valuable for special-purpose applications like drone tracking. Incorporating existing techniques such as WiBall [46], which is based on TRRS as well, may offer (less accurate) distance estimation in arbitrary directions, without the need of a 3D array. The 3D direction, however, remains open for future research.
Fusing inertial sensors: RIM promotes inertial measurements especially for distance and heading estimation, which traditional inertial sensors can hardly measure. Nevertheless, gyroscope provides a reasonably high accuracy of rotating angles while magnetometer reports absolute orientation. Considering that inertial sensors almost always present alongside with WiFi radios, we keep it as a promising direction to fuse them together, both by applying RIM to calibrate inertial sensors and by incorporating inertial sensors with RIM, which would boost ubiquitous inertial measurements for many applications such as indoor tracking and virtual reality.

8 RELATED WORKS

Related works fall in two areas.
Wireless tracking and sensing: Numerous efforts have been devoted to wireless tracking and sensing during the past decades. Many existing works employ different channel parameters for tracking, such as Angle of Arrival (AoA) [13, 15, 43], Time of Flight (ToF) [9, 29], or their fusion [26, 41]. The latest of them [14] pushes the accuracy to sub-centimeter level. These approaches, however, usually require a large phased array or a large frequency bandwidth, typically together with clear LOS condition, to achieve good performance. Efforts have been made to expand bandwidth by frequency hopping [35, 40] and extend antenna array [42], which however incur extra spectrum or hardware overhead. In addition, the existing works require cooperation across multiple APs (four or five) [13, 15, 43]. A few prior proposals attempt to track using a single AP, which again, still need precise AP location and orientation [29, 35], or achieves decimeter accuracy [46]. Moreover, many of the existing works degrade or fail in NLOS conditions. In contrast, RIM is truly multipath-resilient. It has centimeter motion tracking even in complex NLOS scenarios, and resolves direction simultaneously, using only a single arbitrarily placed AP without knowing any of its information.

Other works leverage fingerprinting of dense APs [1, 27, 30]. [25] employs CIR for movement detection, but does not address motion tracking. Centimeter granularity fingerprinting is studied using CIR [39], which is further enhanced by multiple antennas [5] and by frequency hopping [4]. [37] leverages multipath profiles of RFID channel parameters for fine-grained fingerprinting. These proposals require exhaustive calibration prior to deployment and deteriorate due to temporal dynamics. Differently, RIM does not need any calibration. [38] tracks orientation with an array of RFID tags, which relies on precise phase difference of arrival only available on RFID and does not address moving distance. [31] utilizes phase profiles to determine the relative order of RFID tags, yet does not solve the relative distances nor directions. The emerging 802.11mc [11] and Bluetooth 5.1 [34] standards provide Round Trip Time and AoA measurement respectively. Both of them, however, offer limited accuracy, especially in indoor environments.
Inertial sensing: Inertial sensing and its applications have been widely studied in aerodynamics, robotics, and mobile computing [6, 19, 28, 32, 44]. To reduce measurement errors, various algorithms have been proposed for efficient sensor fusion among accelerometer, gyroscope, and magnetometer [2, 22]. Regarding inertial sensing on mobile and wearable devices, the state-of-the-art orientation is probably achieved by A^3 [47] and MUSE [33]. Compared to orientation, tracking moving distance with inertial sensors is a much more challenging task. Prior works only track coarse-grained inertial distance for short-time motion, by counting steps and estimating step lengths [36, 44]. Low-cost inertial sensors however, are hardly able to track precise distance [12], for which RIM offers a superior complement. Inertial sensors have also been employed for activity recognition and classification, which we envision RIM will enable better. We note that RIM is particularly complementary to conventional inertial sensors and we envision the comprehensive fusion of them in the future.

Many other modalities can also track the location of objects, including computer vision [7, 24], visible light [16, 20], acoustic sensing [21, 45], etc. These technologies usually require special infrastructure or are vulnerable to dynamic ambient contexts, making them less favorable for ubiquitous motion measurements. They do not resolve multiple inertial parameters either.

9 CONCLUSION

This paper presents RIM, a precise inertial measurement system that estimates centimeter-level moving distance, heading direction, and rotating angle using commercial WiFi radios. It works over a large multipath rich area wherever is covered by a single unknown AP that is arbitrarily placed, without the support of additional infrastructure or sensors. By doing so, RIM opens up WiFi-based motion sensing to new applications demanding accurate and reliable motion measurements, such as robot monitoring, VR tracking, mobile gaming, etc. This work does not raise any ethical issues.

ACKNOWLEDGMENTS

We thank our shepherd, Prof. Ben Y. Zhao, and the anonymous reviewers for their constructive comments and feedback. The authors are grateful to the SIG@UMD and Origin Wireless members for discussion and support.

REFERENCES

[1] Paramvir Bahl and Venkata N Padmanabhan. 2000. RADAR: An in-building RF-based user location and tracking system. In *Proceedings of IEEE INFOCOM*.

[2] Billur Barshan and Hugh F Durrant-Whyte. 1995. Inertial navigation systems for mobile robots. *IEEE Transactions on Robotics and Automation* 11, 3 (1995), 328–342.

[3] B Bogert. 1957. Demonstration of delay distortion correction by time-reversal techniques. *IRE Transactions on Communications Systems* 5, 3 (1957), 2–7.

[4] C. Chen, Y. Chen, Y. Han, H. Lai, and K. J. R. Liu. 2017. Achieving centimeter-accuracy indoor localization on WiFi platforms: A frequency hopping approach. *IEEE Internet of Things Journal* 4, 1 (2017), 111–121.

[5] C. Chen, Y. Chen, Y. Han, H. Lai, F. Zhang, and K. J. R. Liu. 2017. Achieving centimeter-accuracy indoor localization on WiFi platforms: A multi-antenna approach. *IEEE Internet of Things Journal* 4, 1 (2017), 122–134.

[6] John L Crassidis, F Landis Markley, and Yang Cheng. 2007. Survey of nonlinear attitude estimation methods. *Journal of Guidance, Control, and Dynamics* 30, 1 (2007), 12–28.

[7] Christian Forster, Matia Pizzoli, and Davide Scaramuzza. 2014. SVO: Fast semi-direct monocular visual odometry. In *Proceedings of IEEE ICRA*.

[8] Wesley M Gifford, William Wei-Liang Li, Ying Jun Zhang, and Moe Z Win. 2011. Effect of bandwidth on the number of multipath components in realistic wireless indoor channels. In *Proceedings of IEEE ICC*.

[9] Wei Gong and Jiangchuan Liu. 2018. SiFi: Pushing the limit of time-based WiFi localization using a single commodity access point. *Proceedings of ACM IMWUT* 2, 1 (2018), 10.

[10] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool release: Gathering 802.11n traces with channel state information. *ACM SIGCOMM Computer Communication Review* 41, 1 (2011), 53–53.

[11] IEEE. 2016. IEEE approved draft standard for information technology–Telecommunications and information exchange between systems - Local and metropolitan area networks–Specific requirements Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. *IEEE P802.11-REVmc/D8.0, August 2016* (2016), 1–3774.

[12] Antonio R Jimenez, Fernando Seco, Carlos Prieto, and Jorge Guevara. 2009. A comparison of pedestrian dead-reckoning algorithms using a low-cost MEMS IMU. In *Proceedings of IEEE WISP*.

[13] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. Spotfi: Decimeter level localization using wifi. In *Proceedings of ACM SIGCOMM*.

[14] Manikanta Kotaru and Sachin Katti. 2017. Position tracking for virtual reality using commodity WiFi. In *Proceedings of IEEE CVPR*.

[15] Swarun Kumar, Stephanie Gil, Dina Katabi, and Daniela Rus. 2014. Accurate indoor localization with zero start-up cost. In *Proceedings of ACM MobiCom*.

[16] Ye-Sheng Kuo, Pat Pannuto, Ko-Jen Hsiao, and Prabal Dutta. 2014. Luxapose: Indoor positioning with mobile phones and visible light. In *Proceedings of ACM MobiCom*.

[17] Geoffroy Lerosey, J De Rosny, A Tourin, A Derode, G Montaldo, and M Fink. 2004. Time reversal of electromagnetic waves. *Physical Review Letters* 92, 19 (2004), 193904.

[18] Geoffroy Lerosey, Julien De Rosny, Arnaud Tourin, and Mathias Fink. 2007. Focusing beyond the diffraction limit with far-field time reversal. *Science* 315, 5815 (2007), 1120–1122.

[19] Mo Li, Pengfei Zhou, Yuanqing Zheng, Zhenjiang Li, and Guobin Shen. 2015. IODetector: A generic service for indoor/outdoor detection. *ACM Transactions on Sensor Networks (TOSN)* 11, 2 (2015), 28.

[20] Song Liu and Tian He. 2017. SmartLight: Light-weight 3D indoor localization using a single LED lamp. In *Proceedings of ACM SenSys*.

[21] Wenguang Mao, Jian He, and Lili Qiu. 2016. CAT: high-precision acoustic motion tracking. In *Proceedings of ACM MobiCom*.

[22] João Luís Marins, Xiaoping Yun, Eric R Bachmann, Robert B McGhee, and Michael J Zyda. 2001. An extended Kalman filter for quaternion-based orientation estimation using MARG sensors. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*.

[23] MarketsandMarkets. 2018. Inertial measurement unit (IMU) market - Global forecast to 2022. https://www.marketsandmarkets.com/Market-Reports/inertial-measurement-unit-market-148851976.html.

[24] David Nistér, Oleg Naroditsky, and James Bergen. 2004. Visual odometry. In *Proceedings of IEEE CVPR*.

[25] Neal Patwari and Sneha K Kasera. 2007. Robust location distinction using temporal link signatures. In *Proceedings of ACM MobiCom*.

[26] Kun Qian, Chenshu Wu, Yi Zhang, Guidong Zhang, Zheng Yang, and Yunhao Liu. 2018. Widar2.0: Passive human tracking with a single Wi-Fi link. In *Proceedings of ACM MobiSys*.

[27] Anshul Rai, Krishna Kant Chintalapudi, Venkata N Padmanabhan, and Rijurekha Sen. 2012. Zee: Zero-effort crowdsourcing for indoor localization. In *Proceedings of ACM MobiCom*.

[28] Nirupam Roy, He Wang, and Romit Roy Choudhury. 2014. I am a smartphone and I can tell my user's walking direction. In *Proceedings of ACM MobiSys*.

[29] Souvik Sen, Jeongkeun Lee, Kyu-Han Kim, and Paul Congdon. 2013. Avoiding multipath to revive inbuilding WiFi localization. In *Proceeding of ACM MobiSys*.

[30] Souvik Sen, Božidar Radunovic, Romit Roy Choudhury, and Tom Minka. 2012. You are facing the Mona Lisa: Spot localization using PHY layer information. In *Proceedings of ACM MobiSys*.

[31] Longfei Shangguan, Zheng Yang, Alex X Liu, Zimu Zhou, and Yunhao Liu. 2015. Relative localization of RFID tags using spatial-temporal phase profiling. In *Proceedings of USENIX NSDI*.

[32] Guobin Shen, Zhuo Chen, Peichao Zhang, Thomas Moscibroda, and Yongguang Zhang. 2013. Walkie-Markie: Indoor pathway mapping made easy. In *Proceedings of USENIX NSDI*.

[33] Sheng Shen, Mahanth Gowda, and Romit Roy Choudhury. 2018. Closing the gaps in inertial motion tracking. In *Proceedings of ACM MobiCom*.

[34] Bluetooth SIG. 2019. Enhancing Bluetooth Location Services with Direction Finding. https://www.bluetooth.com/bluetooth-resources/enhancing-bluetooth-location-services-with-direction-finding/.

[35] Deepak Vasisht, Swarun Kumar, and Dina Katabi. 2016. Decimeter-level localization with a single WiFi access point. In *Proceedings of USENIX NSDI*.

[36] He Wang, Souvik Sen, Ahmed Elgohary, Moustafa Farid, Moustafa Youssef, and Romit Roy Choudhury. 2012. No need to war-drive: Unsupervised indoor localization. In *Proceedings of ACM MobiSys*.

[37] Jue Wang and Dina Katabi. 2013. Dude, where's my card?: RFID positioning that works with multipath and non-line of sight. In *Proceedings of ACM SIGCOMM*.

[38] Teng Wei and Xinyu Zhang. 2016. Gyro in the air: tracking 3D orientation of batteryless internet-of-things. In *Proceedings of ACM MobiCom*.

[39] Zhung-Han Wu, Yi Han, Yan Chen, and KJ Ray Liu. 2015. A time-reversal paradigm for indoor positioning system. *IEEE Transactions on Vehicular Technology* 64, 4 (2015), 1331–1339.

[40] Yaxiong Xie, Zhenjiang Li, and Mo Li. 2019. Precise power delay profiling with commodity Wi-Fi. *IEEE Transactions on Mobile Computing* 18, 6 (2019), 1342–1355.

[41] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. 2019. mD-Track: Leveraging multi-dimensionality for passive indoor Wi-Fi tracking. In *Proceedings of ACM MobiCom*.

[42] Yaxiong Xie, Yanbo Zhang, Jansen Christian Liando, and Mo Li. 2018. SWAN: Stitched Wi-Fi ANtennas. In *Proceedings of ACM MobiCom*.

[43] Jie Xiong and Kyle Jamieson. 2013. ArrayTrack: a fine-grained indoor location system. In *Proceedings of USENIX NSDI*.

[44] Zheng Yang, Chenshu Wu, Zimu Zhou, Xinglin Zhang, Xu Wang, and Yunhao Liu. 2015. Mobility increases localizability: A survey on wireless indoor localization using inertial sensors. *ACM Computing Surveys* 47, 3 (2015), 54.

[45] Sangki Yun, Yi-Chao Chen, and Lili Qiu. 2015. Turning a mobile device into a mouse in the air. In *Proceedings of ACM MobiSys*.

[46] Feng Zhang, Chen Chen, Beibei Wang, Hung-Quoc Lai, Yi Han, and KJ Ray Liu. 2018. WiBall: A time-reversal focusing ball method for decimeter-accuracy indoor tracking. *IEEE Internet of Things Journal* 5, 5 (2018), 4031–4041.

[47] Pengfei Zhou, Mo Li, and Guobin Shen. 2014. Use it free: Instantly knowing your phone attitude. In *Proceedings of ACM MobiCom*.

[48] Qiang Zhu, Mingliang Chen, Chau-Wai Wong, and Min Wu. 2018. Adaptive multi-trace carving based on dynamic programming. In *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*.