

PhaseU: Real-time LOS Identification with WiFi

Chenshu Wu^{*§}, Zheng Yang^{*}, Zimu Zhou[†], Kun Qian^{*}, Yunhao Liu^{*} and Mingyan Liu[‡]

^{*}School of Software and TNLIS, Tsinghua University

[§]Department of Computer Science and Technology, Tsinghua University

[†]CSE, Hong Kong University of Science & Technology

[‡]Department of EECS, University of Michigan

{wucs32, hmilyyz, zhouzimu.hk, qiank10}@gmail.com, {yunhao}@greenorbs.com, {mingyan}@umich.edu

Abstract—WiFi technology has fostered numerous mobile computing applications, such as adaptive communication, fine-grained localization, gesture recognition, etc., which often achieve better performance or rely on the availability of Line-Of-Sight (LOS) signal propagation. Thus the awareness of LOS and Non-Line-Of-Sight (NLOS) plays as a key enabler for them. Real-time LOS identification on commodity WiFi devices, however, is challenging due to limited bandwidth of WiFi and resulting coarse multipath resolution. In this work, we explore and exploit the phase feature of PHY layer information, harnessing both space diversity with antenna elements and frequency diversity with OFDM subcarriers. On this basis, we propose PhaseU, a real-time LOS identification scheme that works in both static and mobile scenarios on commodity WiFi infrastructure. Experimental results in various indoor scenarios demonstrate that PhaseU consistently outperforms previous approaches, achieving overall LOS and NLOS detection rates of 94.35% and 94.19% in static cases and both higher than 80% in mobile contexts. Furthermore, PhaseU achieves real-time capability with millisecond-level delay for a connected AP and 1-second delay for unconnected APs, which is far beyond existing approaches.

I. INTRODUCTION

WiFi technology has fostered a broad range of mobile and ubiquitous computing applications beyond wireless data transmission. The past decade has witnessed the fast conceptualization and continuous revolution of myriad emerging applications, e.g., indoor localization, device-free localization, gesture and activity recognition, etc. These applications rely on careful analysis of radio signal features and Non-Line-Of-Sight (NLOS) signal propagation lurks as a critical concern that cannot be easily staved off.

The severe and fickle attenuation of NLOS propagation deteriorates communication link quality and violates theoretical signal propagation models. Normally, NLOS propagation decreases the stability of received signal strengths (RSSs) and exaggerates the RSS fluctuations [1]. It has been demonstrated that the lack of LOS propagation is a major cause of poor wireless experience since it often leads to subdued signal strengths, high packet losses and low PHY rates [2]. Besides wireless communication, many mobile applications rely even more heavily on the presence of the LOS path. For instance, NLOS propagation induces significant bias in time and power based ranging [3], [4], or produces spurious angle estimation [5]. Even for fingerprinting-based localization, severe RSS fluctuations due to multipath effects substantially challenge the accuracy of location estimation [6] and the maintenance

of a valid radio map [7]. Other applications such as energy harvesting [8], health monitoring [9], device-free positioning [10], time synchronization [11], etc., also depend on the existence of the LOS path and will become less practicable under NLOS propagation.

The awareness of LOS and NLOS propagation acts as a pivotal primitive to combat the adverse impacts of NLOS propagation. For example, with the knowledge of LOS/NLOS conditions, transmitters can tune specific link settings like transmitting power or data rates for high throughput and reliable communication [12]. Location providers can adjust the model parameters or adopt comfortable models to maintain high-quality services [4]. Taking an illustrative example as in Figure 1, we intuitively explain the potential benefits of being aware of LOS and NLOS conditions in various perspectives. Imagine that Bob is surfing the Internet while listening to a talk in a classroom, he would like to connect to AP1 with LOS path rather than NLOS ones. To locate Bob, a LOS-specific model should be applied to AP1 while NLOS-tailored models are preferred for AP2 and AP3. During a coffee break, he walks to the lobby and encounters Alice who wants to share a conference video with him. During this time period, the location provider should continuously adapt the propagation model according to the LOS/NLOS conditions to accurately track Bob's presence. When Bob is outside the classroom, his mobile phone should automatically switch to AP3, which can provide LOS service. Furthermore, Alice could slightly turn to obtain a LOS link for video streaming, which generally achieves better performance. In a nutshell, real-time detection of LOS/NLOS propagation paves the way for the enhancements of wireless and mobile applications.

The vision of real-time LOS identification on commodity WiFi devices, however, entails great challenges. Although many theoretical models have been designed to investigate the distributions of channel parameters like Rician K factor, these statistical models require precise channel profiles from dedicated channel sounders, or assume long period of measurements. Other solutions employ extremely wide-band signals like Ultra Wide-Band (UWB) [13] to explore delay spread or range measurement that only needs one-time measurement, yet often halt at simulation. Unfortunately, current WiFi networks operate with a narrow bandwidth of 20~40MHz, thus unable to resolve multipath propagation indoors. Pioneering works [5], [14] that exploit MIMO techniques with a number

of antennas still require hardware modification, hampering the immediate viability. Consequently, none of existing approaches is directly feasible to WiFi because of the coarse-grained channel measurements under limited bandwidth. Recent innovation [15] has demonstrated the primary feasibility of using amplitude features of PHY layer information for LOS identification, yet at the cost of considerable long-time measurements and contrived slight mobility.

In this work, we attempt to design a *real-time identification scheme for both static and mobile scenarios with only commodity off-the-shelf (COTS) WiFi devices*. Insights and progresses in various perspectives underpin the feasibility of the design goal. (1) The recently exposed PHY layer channel state information (CSI) on commercial WiFi devices reveals multipath channel features at the granularity of OFDM subcarriers [1], which is much finer-grained than the traditional MAC layer RSS. Specifically, we observe the phase information, after appropriate sanitization and integration, tend to be an excellent indicator for LOS/NLOS conditions. (2) Commodity WiFi devices are commonly equipped with multiple antennas (typically two or three nowadays) to support the increasingly popular MIMO techniques. These antennas induce diversity in spatial domain of signal propagation. (3) Even for truly mobile users, inertial sensors embedded in mobile devices contribute to capturing some transitory moments in which devices are static. Signal measurement at such “static” moments can be used for distinguishing LOS from NLOS propagation.

Motivated by these insights, we propose *PhaseU*, a real-time LOS identification system with WiFi. To distinguish LOS and NLOS propagation, we explore and exploit the previously insufficiently explored phase information, which can also be extracted from the CSI reported by commercial WiFi network interface card (NIC), the same as more often used amplitude information. We observe that phase difference over two antennas behaves differently in LOS and NLOS conditions, although the raw phase information itself is not directly usable due to the extreme randomness involved in the measurements. Thus we quantify the distinctions of LOS and NLOS conditions using variance of phase difference over a pair of antennas. Besides space diversity, we also harness frequency diversity over multiple OFDM subcarriers to enhance the identification precision and robustness. To enable PhaseU in mobile scenarios, we notice that there always exist transient moments when the devices are actually motionless when a user is natural moving. Hence we employ inertial sensors that are commonly built in commodity mobile devices to infer such moments and use the corresponding measurements for identification.

We prototype PhaseU and conduct extensive experiments in various indoor environments. Evaluation results show that PhaseU achieves an overall LOS detection rate of 94.35% with a false alarm rate of 5.91%. For mobile scenarios, PhaseU yields a respective LOS and NLOS detection rate of 80.08% and 82.91%. Furthermore, PhaseU achieves LOS identification in millisecond-level delay for a connected AP and 1-second delay for nearby unconnected APs, making it applicable for real-time applications.

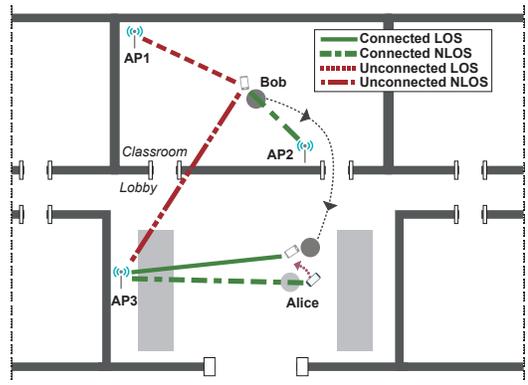


Figure 1. Real-time awareness of LOS and NLOS condition can benefit from various applications of wireless communication and ubiquitous computing.

The main contributions of this work are as follows:

- 1) To the best of our knowledge, we are the first to use phase features of PHY layer information of WiFi to identify the availability of the LOS path in multipath-dense indoor scenarios. And PhaseU is the first real-time LOS identification scheme on commodity WiFi devices.
- 2) Different from applying the direct amplitude and phase information of CSI to diverse environment sensing scenarios, by extracting a new feature of phase difference over antennas from raw CSI, we harness both space diversity and frequency diversity, which advances the state-of-the-art technically. The new feature has demonstrated its high sensitivity and will enable numerous finer-grained sensing applications in addition to LOS identification.
- 3) We prototype PhaseU on commodity WiFi devices and validate its performance in various indoor environments. Experiment results demonstrate that PhaseU consistently outperform previous approaches in both static and mobile scenarios, with respective LOS and NLOS detection rates of around 95% and above 80%.

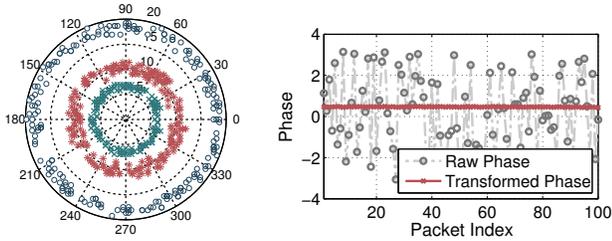
The rest of the paper is organized as follows. We first provide preliminary background in Section II and then introduce the feature extraction in Section III. Section IV details the framework design and the following Section V presents the performance evaluation. We review the related work in Section VI and conclude our work in Section VII.

II. PRELIMINARIES

A. Real-time LOS Identification Problem

In cluttered indoor environments, wireless signals often propagate through multiple paths, where the LOS path may be harshly attenuated or completely obstructed. The LOS identification problem is thus to differentiate the availability of the LOS path from multiple propagation paths. It is typically formulated as a classical binary hypothesis test, where an effective feature plays a central role.

As summarized in our previous work [15], traditional LOS identification techniques mostly use two categories of channel features: Channel Impulse Response (CIR) based and channel



(a) Three groups of random raw phase measurements (b) Random noises are removed in sanitized phase information

Figure 2. Raw phase information and the calibrated version

statistics based. While CIR-based features enable identification with only one channel snapshot, high-resolution CIR is unavailable on COTS WiFi devices. Channel statistics features are built upon certain distributions of the received signal amplitudes and thus usually need a considerable volume of measurements. To enable real-time LOS identification of WiFi, however, triple challenges still reside:

- 1) Commodity WiFi devices fail to support precise CIR measurements due to limited operating bandwidth.
- 2) Existing channel statistics based features require large amount of samples, impeding real-time performance.
- 3) Most LOS identification schemes are designed for stationary scenarios. Even those incorporating slight mobility [15] fail in truly mobile cases.

Before diving into feature extraction in detail, we review channel information available on COTS WiFi infrastructure.

B. Channel State Information

While Received Signal Strength Indicator (RSSI) is the most accessible proxy of channel conditions, it only provides a coarse amplitude estimation for a wireless channel. With slight driver modification, however, PHY layer Channel State Information (CSI) can be revealed to upper layers on off-the-shelf Network Interface Cards (NICs) [1]. CSI contains a set of items and each item represents the channel response information (both amplitude and phase) of an OFDM subcarrier:

$$H(f_k) = \|H(f_k)\|e^{j\angle H(f_k)} \quad (1)$$

where $H(f_k)$ is the CSI at the subcarrier with central frequency of f_k , and $\|H(f_k)\|$ and $\angle H(f_k)$ denote its amplitude and phase, respectively.

Compared with MAC layer RSSI, CSI depicts a finer-grained temporal and spectral structure of wireless links in both amplitudes and phases [1], [4]. In addition, CSI is supported by IEEE 802.11n and subsequent standards.

III. FEATURE EXTRACTION

Fundamentally constrained by the time resolution of COTS WiFi devices, it is infeasible to rely on only one channel snapshot and employ CIR based features for LOS identification. In this section, we extract distinctive statistical features from the largely unexplored phase information, with enhancements harnessing both space and frequency diversity.

A. Exploring Phase Feature

The physical underpinning for statistical features is that the spatial randomness of LOS and NLOS paths differs. NLOS paths typically involve richer reflections, diffractions and refractions. Therefore signals transmitting through NLOS paths often behave more randomly, which manifests in both signal *amplitudes* and *phases*. In this paper, we investigate the potential of phase information for two reasons: First, as most amplitude based features implicitly assume certain distributions (e.g. Rician distribution [16]), large numbers of measurements are necessary for accurate distribution parameter estimation. Second, LOS/NLOS propagation is not the only factor that determines the extent of randomness in received amplitudes. Both obstacle blockage, i.e., NLOS conditions, and long propagation distances can substantially attenuate signal amplitudes. In contrast, phase shifts change periodically over propagation distances, and thus are more robust.

1) *Analysis of Phase Variances*: As discussed in Section II-B, CSI measurements offer the phase information of each subcarrier. The measured phase $\hat{\phi}_i$ for the i^{th} subcarrier can be expressed as:

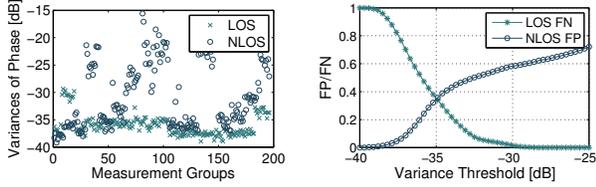
$$\hat{\phi}_i = \phi_i - 2\pi \frac{k_i}{N} \delta + \beta + Z, \quad (2)$$

where ϕ_i denotes the true phase, δ is the timing offset at the receiver, which causes phase error expressed as the middle term, β is an unknown phase offset, and Z is some measurement noise. k_i denotes the subcarrier index (ranging from -28 to 28 in IEEE 802.11n) of the i^{th} subcarrier and N is the FFT size (which equals to 64 in IEEE 802.11 a/g/n). Due to the unknowns listed above, it is infeasible to obtain the true phase shifts with solely commodity WiFi NICs. Figure 2a depicts the raw phase measurements without further calibration. As can be seen, the phases behave extremely randomly.

To mitigate the impact of random noises, we perform a linear transformation on the raw phases, as recommended in [17]. The key idea is to eliminate δ and β by considering phase across the entire frequency band. If the subcarrier frequency is symmetric, we can obtain a linear combination of true phases, denoted as $\tilde{\phi}_i$, from which the random phase offsets have been removed (omitting the small measurement noise Z). Figure 2b illustrates an example of the phase after transformation, which distributes relatively stably as expected.

Similar to the rationale for amplitude based features, the extent of randomness can be revealed in statistics depicting certain order of deviation. For real-time operation, we employ *variance* of the calibrated phase as a candidate feature due to its simplicity. Although we cannot obtain the true phase but a calibrated measurement $\tilde{\phi}_i$, we demonstrate that variances of sanitized phases and true phases differ by only a frequency-related constant multiple (Please refer to Appendix A for details). We thus inspect the relationship of variance of phase to LOS/NLOS conditions using the calibrated phases via a measurement-driven approach.

2) *Measurements of Phase Variances*: We collect 200 groups of measurements under different indoor LOS/NLOS



(a) Variances of phase under different LOS/NLOS scenarios (b) Primary accuracy using variances of phase on single antenna

Figure 3. Variance of phase for LOS identification

conditions, each containing at least 10 seconds of data. Figure 3a plots the variances of the calibrated phases. Each group contains at least 10 seconds of data measured with both TX and RX fixed at a specific location (We will use this dataset throughout this section for all preliminary tests).

Unfortunately, we find that variance of the calibrated phase is insufficient to accurately discern LOS from NLOS propagation. Although the phase variances in NLOS scenarios tend to be larger than those under LOS condition on average, no clear gap can be found between the variances of the two cases. Moreover, as shown in Figure 3b, the prediction accuracy is quite sensitive to the threshold value to achieve stable performance in practice when applying an intuitive threshold-cutting method.

Albeit unable to serve as a distinctive feature for LOS identification, the variance of phases does exhibit different trends under LOS/NLOS propagation. Inspired by the preliminary measurements, we explore more conspicuous phase variance related evidences in the following.

B. Leveraging Space Diversity

A key feature in IEEE 802.11n/ac is to exploit MIMO technology for higher capacity, range and reliability via spatial diversity. It is common to see COTS wireless devices equipped with multiple antennas. In this section, we investigate to fuse the phases of multiple antennas to speed up channel statistics calculation and increase the variance differences between NLOS and LOS conditions.

Similar to Equation 2, we calculate the measured phase difference between two antennas as

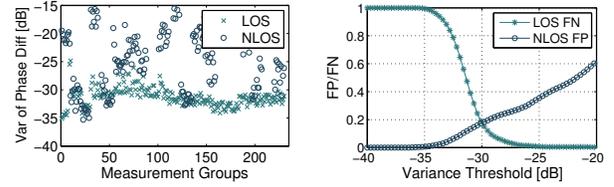
$$\Delta \hat{\phi}_i = \Delta \phi_i - 2\pi \frac{k_i}{N} \Delta \delta + \Delta \beta, \quad (3)$$

where $\Delta \phi_i = \phi_{i,1} - \phi_{i,2}$ is the true phase difference, $\Delta \delta = \delta_1 - \delta_2$ is the corresponding difference of timing offset, and $\Delta \beta = \beta_1 - \beta_2$ is an unknown constant phase difference.

Denote λ as the wavelength, d as the antenna spacing, θ as the direction of arrival, c as the speed of light and T_s as the sample interval at the receiver. Then $\Delta \delta$ can be expressed as

$$\Delta \delta = \frac{d \sin \theta}{c T_s} \leq \frac{1}{2f T_s}, \quad (4)$$

where f is the central frequency, which is around 2.4GHz depending on the operating channel in WiFi networks, while T_s is 50ns. Given $-32 \leq k_i \leq 31$ [1], the phase difference caused by different timing offsets approaches zero, or more specifically, $2\pi \frac{k_i}{N} \Delta \delta \in [-0.0262, 0.0254]$, and thus is negligible in $\Delta \hat{\phi}_i$. As for $\Delta \beta$, although it does vary over time



(a) Variances of phase difference under LOS/NLOS scenarios (b) Primary accuracy using variances of phase difference

Figure 4. Variance of phase difference for LOS identification

due to different uncertain initial phase for each packet, it is possible to attain identical $\Delta \beta$ at different time by shifting the phase difference to be zero mean. Then we can deduce that

$$\sigma_{\Delta \hat{\phi}_i}^2 = \sigma_{\Delta \phi_i}^2. \quad (5)$$

Assume independent signals received at different antennas, which is reasonable for rich scattering environments and antenna space larger than half wavelength [18]¹. Thus $\phi_{i,1}$ and $\phi_{i,2}$ are independent, i.e., $cov(\phi_{i,1}, \phi_{i,2}) = 0$. Then we have the following important inference:

$$\sigma_{\Delta \hat{\phi}_i}^2 = \sigma_{\phi_{i,1}}^2 + \sigma_{\phi_{i,2}}^2. \quad (6)$$

That is, the variances of the phase difference of two antennas is the sum of individual variance on each antenna. [19] derives similar verdict based on a simulated two-antenna system, which is, however, different from commodity WiFi devices and infeasible in practice. Phase difference advances in its easy accessibility from raw phase measurements, requiring no extra transformation or other complex operations. In the following, we primarily validate that variances of phase difference could be a boon to real-time LOS identification by real experiments on COTS WiFi infrastructure.

Figure 4a portrays the variances of phase difference under various LOS/NLOS scenarios using the dataset mentioned above. Compared with Figure 3a, the variances of NLOS conditions are magnified by a larger scale than LOS cases, while variances in both scenarios are amplified. Figure 4b further demonstrates the prediction accuracy of using a naive threshold based method. As can be seen, the result is much better than the case when employing variances of phase on a single antenna (Figure 3b). We argue that variance of phase difference over two antennas proves to be a new applicable feature for LOS identification on commodity WiFi devices.

C. Enhancement via Frequency Diversity

The previous section extracts phase features that characterize the spatial variation of LOS/NLOS propagation. As CSI naturally delineates frequency-selective fading of multipath channels, we investigate to incorporate spectral signatures to attain a more effective and robust feature to characterize LOS and NLOS propagation.

The rationales to exploit frequency diversity for LOS identification are two-fold: 1) Signals experience diverse fading

¹Note that antenna space of greater than half wavelength is also an essential requirement and a basic setting of MIMO which assumes independent channels of different antennas [5].

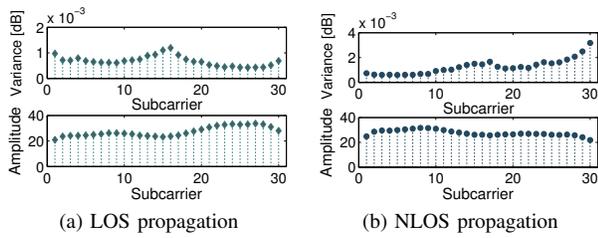


Figure 5. Variances increase with lower amplitudes across the subcarrier frequency in both LOS and NLOS propagation.

levels at different frequency. 2) Signals attenuate differently across the frequency band when penetrating obstacles.

As a result, signals received at different subcarriers exert various disturbances in both envelope and phase [17]. As shown in Figure 5, variances of phase difference and amplitudes both vary over frequency for one CSI measurement. Some key observations are further enlightened on the relationship of variances of phase difference and signal envelopes.

- 1) Signal envelopes vary over the subcarrier frequency band in both LOS and NLOS scenarios.
- 2) Subcarriers that experience severer attenuation suffer from larger variances of phase difference.
- 3) Given the same received power, signals passing via LOS paths experience smaller variances than NLOS paths.

Therefore, both weak LOS signals and NLOS conditions will induce a large variance, while not only LOS condition but also strong signals in NLOS cases experience small variances. To distinguish these scenarios, we propose to suppress the impacts of subcarriers with lower power and magnify the contributions of stronger subcarriers. On this basis, we propose a frequency-selected metric of variances of phase difference as follows, which we term as ρ -factor.

$$\rho = \frac{\sum_{i=1}^n \sigma_{\Delta \hat{\phi}_i}^2 |H(f_i)|}{\sum_{j=1}^n |H(f_j)|} \quad (7)$$

where $|H(f_i)|$ is the mean amplitude of two antennas at the i^{th} subcarrier (since $\Delta \hat{\phi}_i$ is the phase difference of two antennas, we use their mean amplitude as weights accordingly).

To this end, ρ -factor acts as a potential feature that incorporates both space diversity and frequency diversity and we thus utilize it for our real-time LOS identification subsequently.

IV. REAL-TIME LOS IDENTIFICATION

Although we have advanced superior features, significant challenges reside when delivering the features to enable a practical system: 1) How to efficiently measure and process the required information from commodity wireless NICs? 2) How to design a robust and accurate scheme for practical LOS detection? 3) Last but not the least, how to extend LOS identification for truly mobile receivers? In this section, we address these issues and present the design of our real-time LOS identification scheme, called *PhaseU*.

A. Data Measurement

PhaseU collects CSI data from commercial WLAN NICs. After collecting sufficient measurements, PhaseU extracts desired features and infers the current LOS or NLOS conditions.

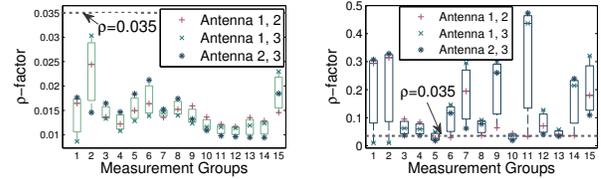


Figure 6. ρ varies over different antenna combinations on both conditions.

The packet intervals and packet quantity are determined according to the accuracy and delay requirements of specific applications and will be discussed in detail in Section V.

As observed in [9], raw CSI data could contain fitful outliers. Hence we filter the raw CSIs through the Hampel identifier [20] to sift outliers before calculating the ρ -factor.

B. Identification

Given a set of CSI samples from N packets, we calculate the variance of phase difference as in Section III and then formulate the following binary hypothesis test with LOS condition H_0 and NLOS condition H_1 .

$$\begin{cases} H_0 : \rho < \rho_{th} \\ H_1 : \rho > \rho_{th} \end{cases} \quad (8)$$

where ρ_{th} denotes a pre-defined threshold. According to our measurements (detailed in Section V), a pre-calibrated unified threshold can fit most scenarios including diverse link lengths, packet numbers and blockage conditions.

C. Selecting Antenna Combination

An increasing number of commercial WLAN devices have been and will be manufactured with more than two antennas, including daily-use wireless routers, ordinary laptops, and edge-cutting pads and smartphones. While PhaseU only requires two antennas, we argue that more antennas can yield more robust and accurate performance.

Given $m(m \geq 2)$ antennas on the receiver, we can derive $\frac{m(m-1)}{2}$ groups of phase differences. As shown in Figure 6, we observe that the variances of phase difference do vary over different antenna combinations. Moreover, the variations fluctuate more severely under NLOS conditions yet always keep low in LOS propagation. Thus we propose to incorporate all possible antenna combinations to enable a robust hypothesis test for LOS identification. Concretely, we extend the hypothesis test in Equation 8 using the median of ρ -factors on different antenna combination:

$$\begin{cases} H_0 : \text{med}(\rho_{i,j}) \leq \rho_{th}, 1 \leq i, j \leq m, i \neq j \\ H_1 : \text{med}(\rho_{i,j}) > \rho_{th}, 1 \leq i, j \leq m, i \neq j \end{cases} \quad (9)$$

where $\rho_{i,j}$ denotes the ρ -factor on antenna i and j .

D. Mobile Scenarios

Previous discussion implicitly assumes a static transmitter-receiver link, which guarantees that the path arriving angles and antenna spacing remain unchanged. Note that the variance of phase is closely related to the changes of propagation

channels. Therefore, in case of a mobile receiver², the phase variance due to LOS/NLOS propagation may be overshadowed by that induced by receiver location change, since the change in receiver’s attitude alters the path arriving angles.

To extend PhaseU to mobile scenarios, we resort to explore a “static” periods during user movements. The feasibility of this strategy arises from three folds:

- 1) For a moving user, there are frequent moments when he/she stops for a while to, e.g., look around, greet somebody, or just check a message on the phone.
- 2) Such “static” moments can be instantly and accurately captured by inertial sensors embedded on most modern tablets and smartphones [7], [21].
- 3) As we will demonstrate in Section V, measurements within a *millisecond*-level period are adequate for PhaseU. Hence, such immediate “static moments” might be long enough for accurate LOS identification.

We thus propose to fortify PhaseU for mobile contexts by adding a motion detection module to capture “static moments” before collecting CSI data. Then only the effective CSIs gathered within such static moments will be used for LOS identification as in stationary cases and the results serve as a LOS/NLOS indicator for adjacent time during a trajectory.

There have been extensive research on human mobility detection via inertial sensors [7], [21]. These techniques employ multi-module sensors, including accelerometer, gyroscope, compass, etc., to infer human mobility such as steps. Since we only need to infer whether the device, instead of human, is in motion, we employ an intuitive threshold-based method. We only utilize the gyroscope since it is more sensitive to device attitude changes. The gyroscope readings are almost zeros when a device keeps stationary, while alters vastly when any motion occurs. Hence a threshold-based method is sufficient and efficient for motion detection in PhaseU.

V. EXPERIMENTS AND EVALUATION

A. Experiment Methodology

Experiment Environments: We conduct the measurements over two weeks in an academic building with corridors and offices, as in Figure 7. Concretely, we collected data in various space including the corridors, one office room, one classroom, and two laboratory rooms. The corridors are enclosed with both bearing walls and non-bearing walls. All rooms are furnished with cubicle desks, computers and other plastic, wooden and metallic furniture. The classroom is further equipped with a metal platform and more desks and chairs.

Data Collection: We collect data along the corridors and in different rooms. Specifically, we perform measurements at around 100 spots, with different environment settings, i.e., diverse TX-RX distances from 1m to 20m, different AP heights including 1m, 1.5m and 2m, various sampling rates, etc. Half of the testing spots are in LOS conditions while

²We only consider receiver mobility since the receiver such as a smartphone is usually attached to a user, while the transmitter is often an AP fixed somewhere such as on the wall.

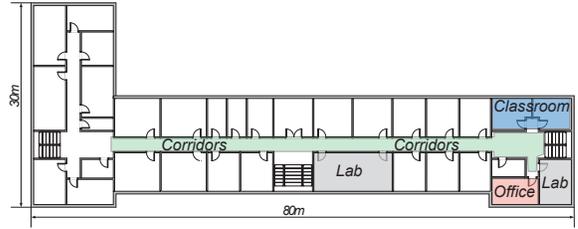


Figure 7. Experiment building (testing areas are highlighted)

the others in NLOS conditions. For each spot, we collect 50 groups of data, each containing 1000 packets. For mobile cases, a volunteer holds a laptop attached with a Google Nexus 7 pad and walks around. The laptop collects CSIs while the smartphone records inertial sensor data. LOS/NLOS propagations are manually marked as ground-truth.

Data are collected from diverse transmitters and receivers. We consider three types of wireless routers (one TP-LINK and one Tenda with single antenna, and one Cisco with multiple antennas installed by the university) as the transmitter operating in IEEE 802.11n AP mode at 2.4GHz. A LENOVO laptop with two antennas and a mini desktop (physical size 170mm×170mm) with three external antennas are used as the receiver pinging packets from the transmitter. Both of them are equipped with Intel 5300 NIC and run Ubuntu 10.0 OS.

Evaluation Metrics: We mainly focus on two metrics to evaluate PhaseU. 1) LOS Detection Rate: The fraction of cases where the LOS condition is correctly identified for all LOS cases. 2) NLOS Detection Rate: The fraction of cases where the NLOS condition is correctly identified for all NLOS cases. (Note that the False Alarm Rate is also interchangeably used, which is simply subtracting the NLOS detection rate from 1.)

To demonstrate the advanced performance of PhaseU, we compare our scheme with two related approaches that both exploit channel statistics of signal envelopes.

- 1) Rician-K factor: The most classical and well-known approach for LOS identification [16]. We employ a practical estimator for Rician-K factor as in [12].
- 2) LiFi: A most recent work in LOS identification, which exploits amplitude features of CSI on WiFi devices [15].

B. Performance

We first report the overall performance of PhaseU and then evaluate the impact of different factors.

1) *Overall Performance:* To quantitatively evaluate the overall performance of PhaseU, we use the Receiver Operating Characteristic (ROC) curves to plot the LOS detection rate against the probability of false alarms. It is a classical graphical view of trade-off between false positives and false negatives over a wide range of thresholds. In general, closer the ROC curve is to the upper left corner, the better performance is indicated. Figure 8 shows the LOS identification performances using 500, 200, and 10 packets, respectively. PhaseU achieves a LOS detection rate of 94.35% with a false alarm of 5.91% using 500 packets. Even with only 10 packets, PhaseU retains high LOS and NLOS detection rates (91.61% and 89.78%). Hence PhaseU can perform accurate LOS identification in 1

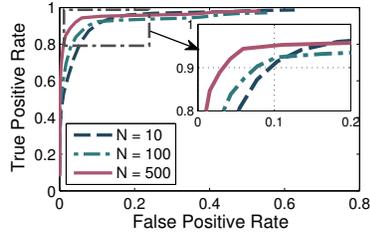


Figure 8. ROC curve of PhaseU

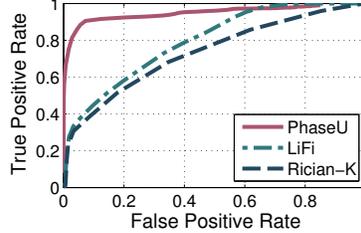


Figure 9. ROC curve comparisons

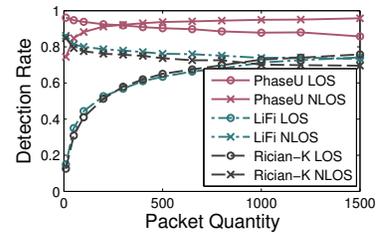


Figure 10. Impact of different packet quantity

second using only beacon packets, which are broadcast at 10Hz by default on COTS APs.

From the ROC curves, we derive a general threshold for ρ -factor that results in balanced LOS and NLOS detection rates. We use 500 packets, the minimum number in LiFi [15], as benchmark quantity for subsequent evaluation. As illustrated in Figure 9, PhaseU greatly outperforms Rician-K and LiFi by up to 20% in both LOS and NLOS detection rates using identical amount of measurements. LiFi only performs slightly better than Rician-K in static scenarios, although it reports reasonable performance in case of mild user-intervened mobility [15].

In the following, we evaluate the impacts of various parameters and demonstrate the benefits of individual components considered in PhaseU design.

2) *Impact of Packet Quantity*: The quantity of packets acts as the most critical factor for real-time LOS identification. We test packet numbers from 10 to 1000. As shown in Figure 10, PhaseU consistently achieves high LOS and NLOS detection rates using different amount of packets (with the same threshold). The average LOS and NLOS detection rates over all cases are 90.84% and 91.01%, respectively. Best balanced detection rates higher than 90% are reached in cases of 200 to 500 packets while the worst case of LOS and NLOS detection rates of 96.12% and 74.47% appears with 10 packets. For comparison, we test the performance of LiFi and Rician-K under the same packet amount range. As shown in Figure 10, PhaseU significantly outperforms previous approaches across all cases, with average improvement of more than 35% and 14% in LOS and NLOS detection performance, respectively. Moreover, both LiFi and Rician-K could achieve balanced performance only with more than 1000 packets, which makes them infeasible for real-time applications. PhaseU reduces the quantity requirement by more than 100x. For the best detection rates of about 90% using 1500 packets of LiFi [15], PhaseU enables similar performance with 15x fewer packets.

We also notice that for both methods, the NLOS detection rate significantly increases while the LOS detection rate slightly decreases with the increasing of packet quantity. This is because that we adopt a variance-based metric for detection. In general, the more packets are involved, the larger the variances are. Thus for a fixed threshold (ρ_{th} in Equation 9), the LOS and NLOS detection rates would certainly exhibit opposite trends over the packet amounts. Fortunately, on the one hand, the drop of LOS detection rate is small compared to the grow of NLOS rate. On the other hand, this unbalanced performance can be tackled by elaborating a packet-quantity-

related threshold.

We evaluate the accuracy over packet quantity instead of time because a receiver can perceive packets from an AP at a wide range of rates. In general, if the receiver is connected to an AP, the packet rate can be up to 1000 packets per second; otherwise, 10 CSI observations can be measured per second from the beacon packets periodically broadcast by each AP. In this sense, we conclude that PhaseU achieves *millisecond*-level real-time LOS/NLOS identification for connected APs as well as *1-second* level identification of unconnected APs nearby, which is far beyond the achievement of previous schemes.

3) *Impact of Obstacle Diversity*: Although we do not explicitly test the performance of PhaseU for each type of blockages, our experiments naturally involve diverse obstacles, including concrete walls, metal platforms and wooden doors. Hence we separately examine the performance of PhaseU in different experimental areas as depicted in Figure 11. There is no significant performance gap among all the cases. The lowest NLOS detection rate of 89.27% appears in case of wooden doors. The metal platform and multi-wall cases slightly exceed others partially because through-metal propagation magnify the difference between LOS and NLOS conditions.

4) *Benefits of Multiple Antenna Combinations*: As depicted in Figure 12, selecting antenna combinations clearly produces better performance than any single antenna pair. Concretely, on our three-antenna device, using only one antenna pair (i.e., antenna 1 and 2, 1 and 3, or 2 and 3) yields best and worst LOS detection rates of 87.44% and 77.52% and NLOS rates of 84.09% and 77.61% (all are the most balanced LOS and NLOS rates). Hence we could summarize that selecting multiple antenna pairs harvests higher performance while PhaseU also works satisfactorily with two-antenna devices. Given that more and more devices are manufactured with three or more antennas, PhaseU is justifiable to provide accurate LOS identification for most devices.

5) *Benefits of Frequency Diversity*: We evaluate the performance improvements brought by frequency diversity in Figure 13. For comparison, we adopt the mean variances of difference of all subcarriers at the average performance of each subcarrier. As is shown, using variances weighted by envelopes of different subcarriers yields marginal performance gain especially on the NLOS detection rate (6% improvement). This result further validates our observation of inverse proportionality between variances of phase difference and CFR envelopes and further demonstrates the necessity of weighting variances against envelopes across difference subcarriers.

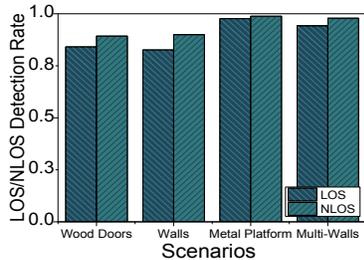


Figure 11. Impact of blockages in 1) office with wooden furniture; 2) corridors with wall blockage; 3) classrooms with metal platform; and 4) rooms separated by multiple walls.

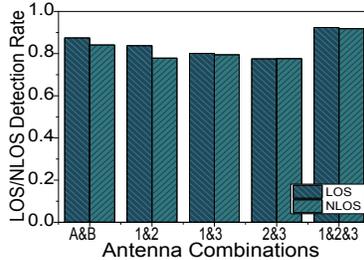


Figure 12. Benefits of multiple antenna combinations. A and B are two antennas combination on the LENOVO laptop while 1, 2, and 3 marks the three antennas on the mini desktop.

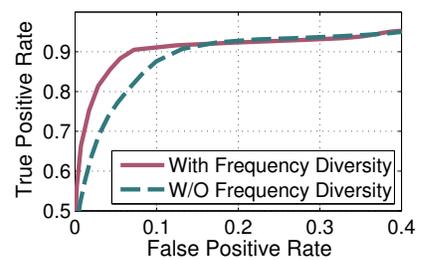


Figure 13. LOS/NLOS identification performance with and without integration of phase information from subcarriers across different frequency.

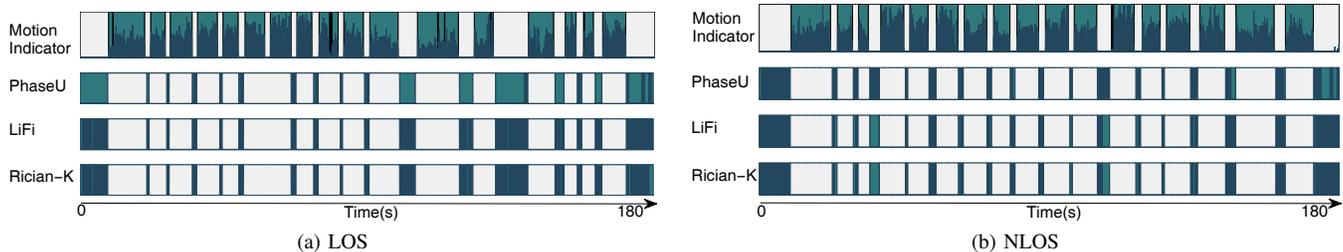


Figure 14. PhaseU achieves remarkable performance in mobile scenarios. For each case, LOS alarms are marked with lighter color and NLOS alarms darker.

C. Performance in Mobile Scenarios

Real-time LOS identification even in mobile scenarios is a spotlight feature of our PhaseU design. To evaluate the performance of PhaseU for mobile users, we let a user walk naturally and stop occasionally and record CSI measurements as well as inertial sensor data along the trajectory. We collect four traces, two LOS and two NLOS conditions. For real-time LOS/NLOS identification, we use 100 packets for each test, i.e., 0.2s if using a packet rate of 500 per second.

We first infer receiver motion by the gyroscope. As shown in Figure 14a and Figure 14b, the gyroscope is extremely sensitive to device motion and thus acts as an accurate indicator for receiver motion. Then for each “static moment”, we employ LOS identification on the corresponding CSIs. Figure 14 illustrates two examples of the identification results. As seen, PhaseU correctly identifies LOS or NLOS conditions in almost all the detected static moments. More specifically, PhaseU successfully predicts LOS condition for 80.08% of the static time and NLOS condition for 82.91%. For comparison, both LiFi and Rician-K fail to identify LOS condition in real-time mobile environments, with respective LOS detection rate of only 6.75% and 10.81%. Most of the time LOS conditions are mistakenly predicted as NLOS. In conclusion, we believe PhaseU is capable of accurate real-time LOS identification even for mobile users, which is far beyond the achievement of previous identification schemes for WiFi.

VI. RELATED WORKS

Related works roughly fall into the following categories.

Exploiting Channel Statistics: Awareness of LOS and NLOS propagation is a primitive in wireless communications and localization, and various channel characteristics have been explored to distinguish the two conditions [22]. Given precise

channel power delay profiles such as high-resolution CIR measurements in UWB systems, metrics depicting both power [23] and delay characteristics [24] have been utilized. For commodity bandwidth constrained systems, researchers turn to analyze the more accessible signal power distributions from multiple packets at the cost of degraded real-time performance and often halt at simulation [16]. The recently PHY layer CSI accessible on COTS WiFi devices has brought about new possibilities since both amplitude and phase information at the granularity of OFDM subcarrier is revealed to upper layers. Prior work such as [15] adopted features of CSI amplitudes in the context of constrained mobility. In contrast, our work harnesses the largely unexplored phase information, and aims at pervasive and real-time LOS identification applicable for both static and truly mobile scenarios.

Leveraging MIMO: The popularity of MIMO technology has extended LOS identification to the spatial dimension. The key insight is that observations from spatially separated locations or antennas potentially magnify the difference in spatial stability between LOS and NLOS paths. By comparing two angular spectra from adjacent antenna arrays [5], peaks with significant angular changes correspond to NLOS paths [5]. However, the calculation of angular spectra requires large number of antenna elements [5] or sophisticated correlation techniques [14]. To further reduce hardware and computation complexity, other researchers explore various non-geometrical features such as phase fluctuation [19]. The closest to our work is [19], which also exploits MIMO as an enhancement when deriving phase features. However, the scheme in [19] is only validated by simulation while ignoring important uncertain noises in practical measurements. Our work takes the similar principle, yet builds upon real measurements, and most importantly, leverages frequency diversity to improve

identification accuracy and investigates inertial sensors to enable identification in mobile contexts and implements a real-time system on COTS devices.

VII. CONCLUSIONS

In this paper, we propose PhaseU, a real-time LOS identification scheme that works in both static and mobile scenarios on commodity WiFi infrastructure. We explore and exploit the phase feature of PHY layer information, harnessing both space diversity and frequency diversity. Accounting for mobile scenarios, we incorporate inertial sensors to infer static moments when the device appears to be motionless for identification. We prototype PhaseU in various indoor spaces. Experimental results demonstrate that PhaseU far outperforms previous approaches in both static and mobile contexts. Furthermore, PhaseU enables real-time capability with millisecond-level delay for a connected AP and 1-second delay for unconnected APs, which is beyond achievement of existing approaches. We envision this work as an important step towards pervasive LOS identification scheme, which paves the way for various WLAN based communication and sensing services.

APPENDIX A PHASE ANALYSIS

Phase Calibration. To mitigate the impact of random noises, we perform a linear transformation on the raw phases, as recommended in [17]. The key idea is to eliminate δ and β by considering phase across the entire frequency band. Firstly, we define two terms a and b as follows:

$$a = \frac{\hat{\phi}_n - \hat{\phi}_1}{k_n - k_1} = \frac{\phi_n - \phi_1}{k_n - k_1} - \frac{2\pi}{N}\delta \quad (10)$$

$$b = \frac{1}{n} \sum_{j=1}^n \hat{\phi}_j = \frac{1}{n} \sum_{j=1}^n \phi_j - \frac{2\pi\delta}{nN} \sum_{j=1}^n k_j + \beta \quad (11)$$

If the subcarrier frequency is symmetric, which indicates $\sum_{j=1}^n k_j = 0$, b can be expressed as $b = \frac{1}{n} \sum_{j=1}^n \phi_j + \beta$. Subtracting the linear term $ak_i + b$ from the raw phase $\hat{\phi}_i$ in Equation 2, we obtain a linear combination of true phases, denoted as $\tilde{\phi}_i$, from which the random phase offsets have been removed (omitting the small measurement noise Z):

$$\tilde{\phi}_i = \hat{\phi}_i - ak_i - b = \phi_i - \frac{\phi_n - \phi_1}{k_n - k_1} k_i - \frac{1}{n} \sum_{j=1}^n \phi_j \quad (12)$$

Note that such linear transformation is not strictly applicable to IEEE 802.11n, where the subcarrier indices are asymmetric, i.e., $\sum_{j=1}^n k_j \neq 0$. Nevertheless, we observe that conducting the transformation on raw phases retrieved by 802.11n devices can still mitigate the randomness to a large extent in practice.

Variance of Calibrated Phase. Since we cannot obtain the true phase but a calibrated measurement $\tilde{\phi}_i$, we demonstrate the relationship between the variances of the calibrated and the true phases. Suppose ϕ_i are i.i.d over frequency, and thus

$$\sigma_{\tilde{\phi}_i}^2 = c_i \sigma_{\phi_i}^2, \quad c_i = 1 + 2 \frac{k_i^2}{(k_n - k_1)^2} + \frac{1}{n} \quad (13)$$

This means the variances of sanitized phases and true phases differ by only a frequency-related constant multiple c_i .

ACKNOWLEDGMENTS

The research of Chenshu Wu is partially supported by NSF China Projects No. 61472098 and 61402338. The research of Zheng Yang is partially supported by NSF China Project No. 61171067. The research of Yunhao Liu is partially supported by the NSF China Major Program No. 61190110.

REFERENCES

- [1] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Predictable 802.11 Packet Delivery from Wireless Channel Measurements," in *ACM SIGCOMM*, 2010.
- [2] A. Patro, S. Govindan, and S. Banerjee, "Observing Home Wireless Experience through WiFi APs," in *ACM MobiCom*, 2013.
- [3] D. Giustiniano and S. Mangold, "CAESAR: Carrier Sense-based Ranging in off-the-shelf 802.11 Wireless LAN," in *ACM CoNEXT*, 2011.
- [4] Y. Zhao, Y. Liu, T. He, A. Vasilakos, and C. Hu, "Fred: Robust rssi-based ranging with multipath effect and radio interference," in *IEEE INFOCOM*, 2013.
- [5] J. Xiong and K. Jamieson, "ArrayTrack: A Fine-Grained Indoor Location System," in *USENIX NSDI*, 2013.
- [6] H. Liu, Y. Gan, J. Yang, S. Sidhom, Y. Wang, Y. Chen, and F. Ye, "Push the limit of wifi based localization for smartphones," in *ACM MobiCom*, 2012.
- [7] Z. Yang, C. Wu, and Y. Liu, "Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention," in *ACM MobiCom*, 2012.
- [8] J. Lin, "Wireless Power Transfer for Mobile Applications, and Health Effects," *IEEE Antennas and Propagation Magazine*, vol. 55, no. 2, pp. 250–253, 2013.
- [9] X. Liu, J. Cao, S. Tang, and J. Wen, "Wi-Sleep: Contactless Sleep Monitoring via WiFi Signals," in *IEEE RTSS*, 2014.
- [10] Z. Zhou, C. Wu, Z. Yang, and Y. Liu, "Sensorless Sensing with WiFi," *Tsinghua Science and Technology*, vol. 20, no. 1, pp. 1–6, 2015.
- [11] Z. Li, W. Chen, C. Li, M. Li, X.-Y. Li, and Y. Liu, "Flight: Clock calibration using fluorescent lighting," in *ACM MobiCom*, 2012.
- [12] C. Tepedelenioglu, A. Abdi, and G. Giannakis, "The Ricean K factor: Estimation and Performance Analysis," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 799–810, 2003.
- [13] I. Guvenc, C.-C. Chong, and F. Watanabe, "NLOS Identification and Mitigation for UWB Localization Systems," in *IEEE WCNC*, 2007.
- [14] K. Joshi, S. Hong, and S. Katti, "PinPoint: Localizing Interfering Radios," in *USENIX NSDI*, 2013.
- [15] Z. Zhou, Z. Yang, C. Wu, W. Sun, and Y. Liu, "LiFi: Line-Of-Sight Identification with WiFi," in *IEEE INFOCOM*, 2014.
- [16] F. Benedetto, G. Giunta, A. Toscano, and L. Vegni, "Dynamic LOS/NLOS Statistical Discrimination of Wireless Mobile Channels," in *IEEE VTC*, 2007.
- [17] S. Sen, B. Radunovic, R. R. Choudhury, and T. Minka, "You are Facing the Mona Lisa: Spot Localization using PHY Layer Information," in *ACM MobiSys*, 2012.
- [18] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An Overview of MIMO Communications—a Key to Gigabit Wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–218, 2004.
- [19] W. Xu, Z. Wang, and S. R. Zekavat, "Non-Line-Of-Sight Identification via Phase Difference Statistics Across Two-Antenna Elements," *IET Communications*, vol. 5, no. 13, pp. 1814–1822, 2011.
- [20] L. Davies and U. Gather, "The Identification of Multiple Outliers," *Journal of the American Statistical Association*, vol. 88, no. 423, pp. 782–792, 1993.
- [21] A. Brajdic and R. Harle, "Walk Detection and Step Counting on Unconstrained Smartphones," in *ACM UbiComp*, 2013.
- [22] C. Gentile, N. Alsindi, R. Raulefs, and C. Teolis, "Multipath and NLOS Mitigation Algorithms," in *Geolocation Techniques*. Springer, 2013, pp. 59–97.
- [23] M. Heidari, N. Alsindi, and K. Pahlavan, "UDP Identification and Error Mitigation in TOA-based Indoor Localization Systems using Neural Network Architecture," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3597–3607, 2009.
- [24] I. Guvenc, C.-C. Chong, F. Watanabe, and H. Inamura, "NLOS Identification and Weighted Least-Squares Localization for UWB Systems using Multipath Channel Statistics," *EURASIP Journal on Advances in Signal Processing*, no. 36, 2008.