

A Survey on Human Profile Information Inference via Wireless Signals

Qiuye He^{1b}, Edwin Yang^{1b}, and Song Fang^{1b}

Abstract—Due to the ubiquitous deployment of wireless infrastructures, the radio signal nature of invisibility, and the elimination of the line-of-sight requirement, it has drawn increasing attention in both academia and industry to infer various human-motion related sensitive information, called *human profile information (HPI)*. The basic idea of these techniques is that varying human profiles (i.e., physiological characteristics and motion patterns) may lead to unique and subtle disturbances in environmental wireless signals, which can be then measured and processed to learn HPI. In this survey, we comprehensively review different categories of existing studies based on (1) how they quantize the motion-induced wireless disturbances in concrete wireless measurements; (2) the signal processing techniques for building wireless HPI inference systems; and (3) practical applications that take advantage of inferred HPI. Also, the survey discusses the emerging challenges and future directions on wireless HPI inference.

Index Terms—Wireless inference, human privacy, wireless channel variation, movement detection.

I. INTRODUCTION

WIRELESS signals are ubiquitous, invisible, and able to penetrate through obstacles. In recent years, there is increasing interest to utilize wireless signals to infer various personal information, such as user identities [1], [2], vital signs [3], [4], [5], private conversations [6], emotional states [7], postures [8], [9], [10], [11], [12], [13], handwriting [14], [15], and keystrokes [16], [17], [18], [19], [20], [21]. We refer to such sensitive information as human profile information (HPI). Specifically, human activity (e.g., breathing, walking, and typing) causes subtle environmental impacts unique to that activity pattern, which can be observed through wireless signals. With sensed HPI, many applications become practical. For example, [22] utilizes changes in wireless channel traces to detect breathing activity to verify the human presence and thus help prevent replay attacks against devices with voice interfaces.

While the growing popularity of wireless sensing techniques has been beneficial to society, this popularity also brings a key source of our security woes, i.e., an adversary may be

able to determine the target user's activity by collecting the corresponding wireless signals and performing HPI inference. For instance, [23] shows that an eavesdropper can silently localize and track individuals in a building from outside walling by listening to ambient WiFi signals with a single smartphone; [21] proposes a wireless inference technique that an attacker can use to infer typed numbers, such as personal identification numbers (PINs) without the aid of any training.

Multiple physical modalities of wireless channel characteristics have been utilized to characterize the environmental disturbance caused by human activities, including received signal strength (RSS) (e.g., [11]), Doppler shift in wireless transmissions (e.g., [8]), channel impulse response (CIR) (e.g., [13]), channel frequency response (CFR) (e.g., [9]), angle of arrival (AoA) (e.g., [15]), and time of flight (ToF) (e.g., [24]). RSS measures the power present in the received wireless signal. CIR and CFR describe how the wireless channel impacts the radio signal that propagates through the channel (e.g., time delay, amplitude attenuation, and phase shift) in the time and frequency domains, respectively. CFR is the Fourier Transform of CIR, and both can be called channel state information (CSI). AoA is utilized to determine the direction in which the wireless signal arrives at the receiver while ToF measures the time taken by a wireless signal to travel from the transmitter to the receiver. A Frequency Modulated Continuous Wave (FMCW) radar is often used to obtain ToF via measuring changes in the signal frequency [3], [7], [25], [26], [27], [28], [29], [30]. Besides, the signal frequency shift induced by human motion can also be measured via a Doppler-based approach [8], [31], [32], [33], [34], [35], [36], [37], [38]. Varying HPI inference algorithms then take advantage of either a single or multiple wireless channel characteristics as input to translate them into target HPI.

Table I compares our survey with related ones on wireless sensing. There are three key differences.

First, existing surveys [39], [40], [41], [42], [43], [44] primarily take advantage of RSS or CSI to achieve inference, while they lack detailed exploration of other notable modalities like ToF, AoA, FMCW, and Doppler shift, all of which are vital for HPI inference systems. Our survey stands out from existing surveys by concentrating on various modalities.

Second, regarding survey context, some surveys (e.g., [39], [40], [45]) put a specific emphasis on applications and future trends. However, such surveys often lack a comprehensive investigation of wireless sensing, particularly in terms of signal processing algorithms and inference models. In our survey, we provide a comprehensive summary and comparison of the data collection, signal preprocessing, feature

Manuscript received 17 February 2023; revised 11 July 2023, 3 November 2023, and 12 January 2024; accepted 26 February 2024. Date of publication 7 March 2024; date of current version 22 November 2024. This work was supported in part by NSF under Grant 1948547 and Grant 2155181, and in part by the University of Oklahoma's Office of the Vice President for Research and Partnerships and the Office of the Provost. (Corresponding author: Song Fang.)

The authors are with the School of Computer Science, University of Oklahoma, Norman, OK 73019 USA (e-mail: qiuye.he@ou.edu; edwiny@ou.edu; songf@ou.edu).

Digital Object Identifier 10.1109/COMST.2024.3373397

TABLE I
SUMMARY OF RELATED SURVEYS ON WIRELESS SENSING

Work	Modalities	Topic	Application Scenarios
Yang: CSUR 13 [39]	RSS, CSI	principles, methods, future directions	indoor localization
Xiao: CSUR 16 [46]	UWB, RFID, Wi-Fi, acoustic	principles, techniques, device-based/device-free systems, future research	indoor localization
Yousefi: ComMag 17 [40]	CSI	different activity recognition systems, evaluation of different methods, challenges	behavior recognition
Liu: COMST 19 [47]	RSS, CSI, FMCW, Doppler shift	techniques, applications, limitations and future trends	intrusion detection, room occupancy monitoring, activity recognition, gesture recognition, vital signs monitoring, user identification, indoor localization&tracking
Ma: CSUR 19 [41]	CSI	signal processing techniques, algorithms, applications, challenges and trends	activity recognition, gesture recognition, human identification, localization, human counting, respiration monitoring, WiFi imaging
Al-qaness: Sensors 19 [42]	RSS, CSI	CSI-based approaches, CSI sensing methodology, challenges and future suggestions	activity recognition, motion detection, localization
Wang: Access 19 [43]	CSI	principles, general methods, applications, issues and future directions	Daily behavior recognition, falling detection, hand gesture recognition, crowd counting, user authentication, respiration monitoring
Liu: Sensors 20 [48]	RFID, FMCW, Wi-Fi, visible light, LoRa, acoustic, LTE	model (Doppler, Fresnel zone, FMCW, AoA, mD-Track), signal processing, application, challenges and future trends	activity recognition, counting, detection, tracking
Wang: JCDE 21 [44]	CSI	signal models, signal processing, application, advantages, limitations and future trends	localization and tracking, daily behavior recognition, respiration detection
Tan: JIOT 22 [45]	CSI, AoA/AoD, ToF, Doppler shift	signal models, datasets and tools, application, challenges and future trends	activity recognition, object sensing, localization
Our survey	RSS, CSI, AoA, ToF, FMCW, Doppler shift	wireless measurements, wireless HPI inference techniques, applications, challenges and trends	user identification, intrusion detection, indoor user localization&tracking, person counting, gesture tracking, vital signs monitoring, keystroke recognition

extraction, and inference design for various wireless HPI inference systems. The data collection section covers the hardware platforms utilized to obtain wireless measurements; we examine signal preprocessing, comparing techniques such as interpolation, outlier and noise reduction, and signal separation; we discuss feature extraction methods in the time domain, frequency domain, and time-frequency domain. Also, we categorize wireless HPI inference algorithms into traditional modeling based and machine learning based approaches, and further discuss their specific advantages and limitations. Moreover, our survey sheds light on potential future research trends that aim to enhance existing wireless sensing capabilities and uncover corresponding new opportunities.

Third, considering application scenarios, each of the existing surveys [39], [40], [45] only gives a brief review of specific scenarios such as indoor localization [39], [46] and behavior recognition [40]. On the other hand, some surveys [42], [44], [45] narrow their scope to three principal application scenarios without delving into a discussion of other broader potential applications. On the contrary, our survey provides a holistic topic on wireless HPI inference, exhibiting a systematic structure, comprehensive comparisons of applications, and an insightful view of future trends. Specifically, we summarize nine main applications related to HPI inference, encompassing (i) user identification, (ii) intrusion detection, (iii) indoor user localization and tracking, (iv) person counting, (v) gesture tracking, (vi) vital signs monitoring, (vii) keystroke recognition, (viii) integration into real-world systems, and (ix) communication efficiency enhancement.

Our main contributions are summarized as follows.

- We present a systematic and comprehensive review of HPI inference via wireless signals, including basic principles, system structure, detailed technical components, as well as extensive comparisons of existing studies.
- We classify existing approaches on wireless HPI inference based on varying critical aspects, including utilized wireless modalities, extracted feature properties, and application domains.
- On observing wireless sensing technology advancement, we discuss the future trends and challenges of wireless HPI inference, including the popularity of Internet-of-Things (IoT) devices, integration with machine learning techniques, and increasing adoption of millimeter wave (mmWave) communications.

Figure 1 illustrates the main structure of the survey. In Section II, we discuss multiple wireless measurements that can be utilized as the HPI source. The general pipeline for wireless HPI inference system and its practical applications are presented in Sections III and IV, respectively. The collected raw wireless measurements are fed into the data preprocessing module. Various features could be extracted from the preprocessed data, and can further lead to the target HPI by correspondingly designed inference algorithms. The summarization and categorization of existing wireless HPI inference techniques shed light on future research directions, as discussed in Section V.

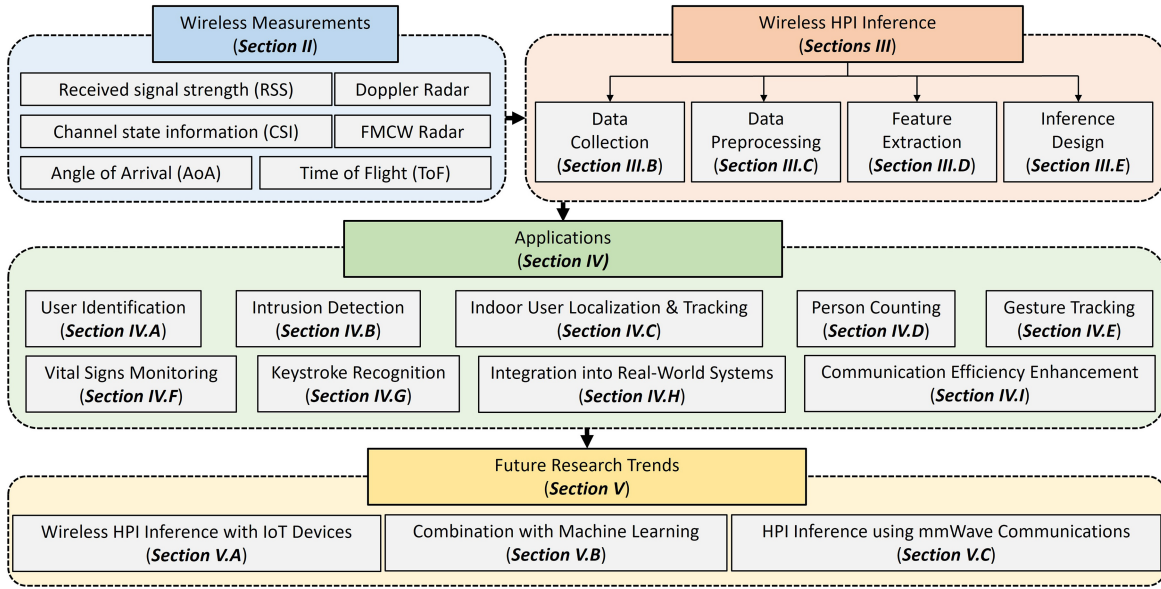
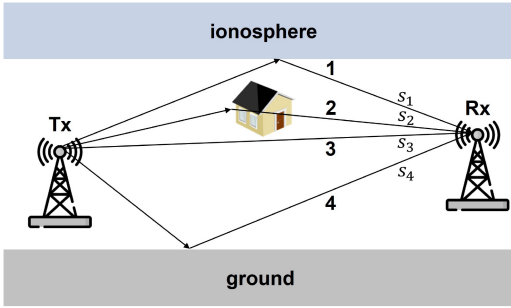


Fig. 1. The basic structure of the paper.

Fig. 2. Multipath example: except going through the direct path, the signal sent from the transmitter (Tx) is also reflected by the ionosphere, the building, and the ground; the received signal is the combination of four multipath components (i.e., $s_1 + s_2 + s_3 + s_4$).

II. WIRELESS CHANNEL CHARACTERISTICS

In this section, we first introduce the concepts of the multipath effect and then give the common physical modalities of wireless channel characteristics. Finally, we present the prevalent algorithms for estimating these modalities.

A. Multipath Effect

A wireless signal usually propagates in the air along multiple paths due to reflection, diffraction, and scattering. As a result, a receiver receives multiple copies of the signal from different paths, each of which has a different delay due to the path it traverses. The received signal can be denoted with the sum of these time-delayed signal copies.

Figure 2 shows an example of a multipath channel. Each path imposes a response (e.g., time delay, magnitude attenuation, and phase shift) on the signal traveling along it, and the superposition of all responses between two nodes is referred to as a *channel impulse response (CIR)* [49], [50], [51], which quantifies the effect of the multipath environment in wireless communications. Mathematically, to fully characterize the

individual paths, the CIR can be denoted as

$$h(\tau) = \sum_{i=1}^L a_i e^{-j\theta_i} \delta(\tau - \tau_i), \quad (1)$$

where a_i , θ_i , and τ_i are the magnitude, phase, and time delay of the i^{th} path, respectively. L is the total number of multipath components, and $\delta(\tau)$ is the Dirac delta function. The channel impulse response is actually the superposition of multiple component responses, each characterizing the distortion that each path has on the multipath component.

The multipath effects of different wireless links are different, and thus their corresponding channel impulse responses also differ [50], [51], [52]. As mentioned earlier, human activities cause subtle environmental impacts and thus will cause changes in wireless channel characteristics. To infer HPI, existing research efforts take advantage of various wireless measurements, which are introduced in Section II-B.

B. Various Wireless Measurements

Wireless channel characteristics between two wireless devices are unique to environmental disturbance and can be quantified by the CIR of the channel. In addition, the following wireless measurements are also utilized to estimate environmental changes and thus infer corresponding HPI:

- *Received signal strength (RSS)*: it is a measurement of the power present in a received wireless signal.
- *Channel state information (CSI)*: it represents how wireless signals propagate from the transmitter to the receiver at certain carrier frequencies along multiple paths [41], and is formed by the *channel frequency responses* measured from the subcarriers.
- *Angle of Arrival (AoA)*: it is often utilized to determine the direction in which a wireless signal arrives.
- *Time of Flight (ToF)*: it demonstrates the time a wireless signal takes to travel from the transmitter to the receiver.



Fig. 3. A taxonomy of HPI inference system.

- *Frequency Modulated Continuous Wave (FMCW) radar*: it can measure the differences in frequency between the transmitted and the received signals.
- *Doppler radar*: it measures the frequency change of the reflected signal using the Doppler effect.

In summary, various wireless measurements serve specific purposes: RSS indicates the average power in a received wireless signal across its entire power bandwidth, CSI provides detailed subcarrier-level amplitude and phase information, AoA determines the angle of the received signal, and ToF measures the distance of signal sources. Also, FMCW utilizes frequency modulation for range and velocity estimation, and the Doppler effect is leveraged to measure the velocity of moving objects. Figure 3 gives a summary of existing wireless HPI inference techniques.

C. Obtaining Wireless Measurements

These wireless measurements have distinct advantages and limitations, and Table II comprehensively compares them.

1) *Deriving RSS*: As aforementioned, a wireless signal sent from the transmitter usually propagates to the receiver through multiple paths, and each path has a different impact on the transmitted signal. Therefore, the distorted signal that arrives at the receiver is the combination of signals via these paths and can be represented as [70]

$$V = \sum_{i=1}^L V_i e^{j\psi_i} + n, \quad (2)$$

where n is the additive noise, V_i and ψ_i are the amplitude and phase of the i^{th} multipath component respectively, and L is the total number of these components. The received power is thus $\|V\|^2$. Therefore, RSS, i.e., the received power in decibels (dB), can be expressed as $RSS = 10\log_{10}(\|V\|^2)$.

Relationship Between RSS and Transmission Distance: The received power is related to the amplitude attenuation (i.e., path loss), which increases exponentially with the distance of the propagation path. RSS thus can be used to estimate the distance between the transmitter and the receiver, denoted

TABLE II
COMPARISON OF WIRELESS MEASUREMENTS

Measurements	Derived Metric	Granularity	Specialized Hardware	Use Case
RSS	signal strength	coarse-grained	No	positioning system
CSI	channel properties	fine-grained	Yes	indoor localization, gesture recognition
AoA	angle of signal source	fine-grained	Yes	direction measurement, indoor tracking
ToF	distance between object and signal source	fine-grained	Yes	distance estimation, depth sensing
FMCW	range and velocity estimation of moving object	fine-grained	Yes	velocity estimation, micro-motion measurement
Doppler shift	velocity estimation of moving object	fine-grained	Yes	speed detection, precise localization

as d . Specifically, the relationship between RSS and d can be denoted with the log-distance path loss model [71], i.e.,

$$\overline{PL}(d) = PL(d_0) + 10 \cdot \gamma \cdot \log_{10}\left(\frac{d}{d_0}\right) + X_\sigma, \quad (3)$$

where $\overline{PL}(d)$ is the total path loss in decibels for the transmitter-receiver separation distance d , $PL(d_0)$ denotes the path loss (in dB) from the transmitter to the reference distance d_0 , γ is the path loss exponent (which indicates how fast path loss increases with distance), and X_σ is a zero mean log-normally distributed random variable with standard deviation σ in decibels, reflecting the shadowing effects.

By utilizing the relationship between RSS and transmission distance, one of the most popular applications of RSS has been in the field of indoor localization [55], [72], [73]. Previously collecting RSS data at preset locations allows for distance calculation based on the current signal and reference locations, enabling target position estimation. Moreover, changes in the multipath signal can cause variations in RSS, making it useful for other HPI inference applications, including intrusion detection [74], [75], [76], person counting [55], [70], [77], [78], and vital sign monitoring [56], [79], [80], [81].

Pros: RSS can be easily measured without requiring specialized hardware.

Cons: RSS measurement is prone to be affected by environmental factors such as multipath fading and obstacles, resulting in inaccuracies for RSS-based distance estimation.

Use case: While RSS can be easily obtained in commodity WiFi devices without requiring additional hardware, it has limitations in various applications due to the coarse-grained information it provides. Particularly, it cannot be used for fine-grained HPI inference, such as detecting finger-level gestures [82] or tracking hand trajectory [15].

2) *CSI Estimation:* The orthogonal frequency-division multiplexing (OFDM) technique is widely used in various modern wireless communication systems (e.g., 802.11a/g/n/ac/ad). OFDM utilizes multiple subcarrier frequencies to encode a packet. As mentioned earlier, we can utilize the channel frequency responses measured from the subcarriers to denote the CSI of OFDM. For an OFDM system with N subcarriers, let $H(f_j, t)$ ($j \in \{1, 2, \dots, N\}$) denote the channel frequency response at time t , where f_j is the frequency of the j^{th} subcarrier. The channel frequency response is usually estimated with a pseudo-noise sequence which is publicly known [49]. Specifically, a transmitter transmits a publicly known pseudo-noise sequence $X(f_j, t)$ over the j^{th} subchannel to the receiver, who estimates the channel frequency response from $X(f_j, t)$ and the received, distorted

signal $Y(f_j, t)$. We can thus compute $H(f_j, t)$ by $H(f_j, t) = \frac{Y(f_j, t)}{X(f_j, t)}$. Each channel frequency response is a complex value representing both amplitude and phase components. Therefore, $H(f_j, t)$ can be rewritten with $|H(f_j, t)|e^{j\angle H(f_j, t)}$, where $|H(f_j, t)|$ and $\angle H(f_j, t)$ denote amplitude and phase, respectively. The CSI of this OFDM system can be then denoted with a vector of channel frequency responses, i.e., $[H(f_1, t), H(f_2, t), \dots, H(f_N, t)]^T$, where $(\cdot)^T$ is the transpose operator.

Since the received signal reflects the constructive and destructive interference of multipath signals, a certain human activity creates a unique multipath environment and thus generates a unique pattern in the time series of CSI values, which can be used for recognizing human identity (e.g., [66], [83]), moving humans (e.g., [84], [85]), and various human activities, such as finger gesture [82], [86].

Pros: CSI provides fine-grained information about the wireless channel, including amplitude, phase, and frequency response, allowing for high-accuracy indoor positioning and tracking applications.

Cons: It normally requires sophisticated hardware to extract CSI, such as Intel 5300 NIC [87], Atteors 9580 NIC [88], and Universal Software Radio Peripheral (USRP) platforms. However, CSI measurements, particularly corresponding phase information, can be impacted by imperfections in hardware components. Additionally, it typically requires more processing and computational resources to analyze CSI streams compared to RSS data.

Use case: Nowadays, with the availability of advanced hardware, an increasing number of studies adopt CSI rather than RSS, to achieve superior performance for various human activity sensing applications, including intrusion detection [85], gesture tracking [89], and keystroke recognition [21].

3) *Calculation of AoA:* The AoA at the receiver represents the direction of the incident signal arriving at the antenna array. It can be calculated by comparing the phases of the CSI values obtained from multiple antennas. The CSI phase changes linearly by 2π for every carrier wavelength λ (i.e., $\frac{c}{f}$, where f is the signal frequency and c denotes the speed of light) along the path from the transmitter to the receiver [15].

We assume that there are M incoming signal paths. Correspondingly, the M incident signals $s_1, s_2, \dots, s_M$ arrive at the antenna array from directions $\theta_1, \theta_2, \dots, \theta_M$, respectively. As shown in Figure 4, for the i^{th} signal ($i \in \{1, \dots, M\}$), two adjacent antennas spaced d apart would introduce a phase difference of $\frac{2\pi \cdot f \cdot d \sin \theta_i}{c}$. For the whole antenna array, we can then define these phase shifts relative

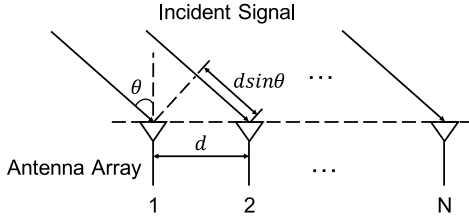


Fig. 4. An incident signal arrives at an array of N evenly spaced antennas with an angle θ (i.e., AoA), where the distance between two adjacent antennas is denoted with d .

to the first antenna as the following steering vector,

$$\mathbf{a}(\theta_i) = \left[1, e^{-\frac{2\pi \cdot f \cdot d \sin \theta_i}{c}}, \dots, e^{-\frac{2\pi(N-1) \cdot f \cdot d \sin \theta_i}{c}} \right]^T. \quad (4)$$

Given all M incoming signal paths, we can construct the $N \times M$ steering matrix as $\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_M)]$. Thus, the received signal at each antenna can be expressed as the superposition of all M incoming signal paths:

$$[x_1, x_2, \dots, x_N]^T = \mathbf{A}[s_1, s_2, \dots, s_M]^T + \mathbf{n}, \quad (5)$$

where x_j ($j \in \{1, \dots, N\}$) denotes the received signal at the j^{th} antenna, and $\mathbf{n}$ is the noise vector. Rewriting Equation (5) in a compact matrix form yields

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{n}, \quad (6)$$

where $\mathbf{X}$ is a vector consisting of received signals at N antennas, and $\mathbf{S}$ is a vector consisting of transmitted signals from M sources.

When the number of antennas exceeds that of the incoming signal paths (i.e., $N > M$), the conventional multiple signal classification (MUSIC) algorithm [90] can be then applied to estimate the matrix $\mathbf{A}$, from which the AoAs can be deduced. The underlying principle of the MUSIC technique is that the eigenvectors of $\mathbf{X}\mathbf{X}^H$ corresponding to the eigenvalue zero if they exist, are orthogonal to the steering vectors in $\mathbf{A}$ [61], where $(\cdot)^H$ denotes the Hermitian transpose operator.

Besides, a naive AoA estimation method can be developed when we consider a signal source at position s , and a pair of antennas a_1 and a_2 separated by d at the receiver. d_{s,a_1} and d_{s,a_2} denote the distances from the signal source to the two antennas respectively. Let Δd and $\Delta\Phi$ denote the distance difference and the measured phase difference between the received signals at the two antennas, respectively. Thus, we have $\Delta d = |d_{s,a_1} - d_{s,a_2}| \approx d \sin \theta$, where θ is the angle of arrival, and $\Delta\Phi = 2\pi \cdot (\frac{\Delta d}{\lambda} - k)$, where k can be any integer in $[-\frac{\Delta d}{\lambda} - \frac{\Delta\Phi}{2\pi}, \frac{\Delta d}{\lambda} - \frac{\Delta\Phi}{2\pi}]$. As a result, we obtain $\theta = \arcsin(\frac{\Delta\Phi \cdot \lambda + 2\pi \cdot k \cdot \lambda}{2\pi \cdot \Delta d})$.

Figure 5 shows a simple application scenario, where the target user (with a mobile device sending out wireless signals) can be localized with the locations of the two nodes and the corresponding AoAs θ_1 , and θ_2 . AoA has also been utilized to achieve other HPI-related applications, e.g., tracking human motion [37], [62], [68], and occupancy detection [91].

Pros: AoA provides information about the angle or direction from which a signal arrives at the receiver, enabling source localization and tracking.

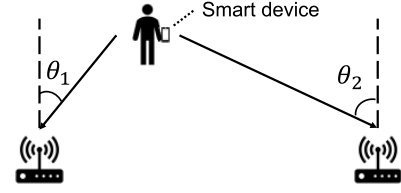


Fig. 5. AoA-based localization.

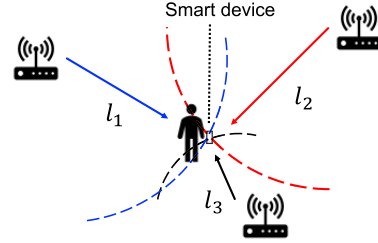


Fig. 6. Trilateration in two-dimensional space.

Cons: AoA measurements require antenna arrays or specialized hardware to accurately estimate angles. Meanwhile, they can be affected by multipath fading and reflections, which may introduce inaccuracies.

Use case: AoA measurements are commonly used in scenarios where the direction or location of a signal source needs to be determined. They find applications in estimating user positions [37] and tracking hand trajectories of drawing letters [15]. Moreover, AoA measurements are often combined with other wireless measurements to provide comprehensive and detailed information on the environment, enhancing human motion sensing capabilities. For instance, mD-Track [92] utilizes multiple sensing techniques together, including AoA, Time of Flight (ToF), Doppler shift, and other information, to track individuals in indoor scenarios.

4) *Calculation of ToF:* The ToF (denoted with Δt) can be utilized to calculate the corresponding propagation distance l or phase shift $\Delta\phi$, and we have $l = c \cdot \Delta t$ and $\Delta\phi = 2\pi \cdot f \cdot \Delta t$, where c is the speed of light and f denotes the carrier frequency of the transmitted signal. Consequently, there are extensive research efforts in using ToF to achieve localization [24], [61]. Figure 6 illustrates the principle of ToF-based localization in two-dimensional space, where a receiver (e.g., a smart device) receives signals from the three transmitters and measures corresponding ToFs. With each ToF, the receiver obtains corresponding propagation distance l_i ($i \in \{1, 2, 3\}$). With the location information of the three transmitters and the technique of trilateration, the receiver can be localized. Accordingly, we draw three circles with each transmitter as a center and the corresponding l_i as a radius. The three circles intersect at one point, which is the position of the receiver.

The accuracy of measured ToF usually depends on three key factors: the time synchronization between the pair of transmitter and receiver, the signal bandwidth, and the sampling rate. First, if a transmitter and receiver pair is used to measure one-way ToF, which equals the difference between the time

of signal transmission and the time of signal arrival. Thus, an accurate ToF estimation requires time synchronization between the two parties [93]. Second, the temporal resolution of ToF can be expressed as $\frac{1}{2B}$, where B denotes the signal bandwidth [94]. Thus, a low signal bandwidth would reduce the resolution. Third, a low sampling rate degrades the ToF resolution as the signal may arrive at the receive antenna during the sampling intervals. Besides, the existence of non-line-of-sight (NLoS) paths between the transmitter and the receiver may also introduce errors in ToF measurements, as a longer path would lead to an increase in propagation time.

Human activity occurring between the transmitter and the receiver brings changes in the propagation distance of the transmitted signal. The ToF would vary accordingly. Thus, with the variation of ToF, it is possible to infer HPI, such as human motion [62] and vital signs [69].

Pros: ToF measures the time it takes for a signal to travel from the transmitter to the receiver, offering precise distance information, particularly in line-of-sight scenarios. Furthermore, ToF is less susceptible to the effects of multipath fading, resulting in more reliable distance estimations.

Cons: ToF measurements often necessitate a high sampling rate in a device's Analog-to-Digital Converter (ADC) to achieve fine range resolution. Alternatively, they can be obtained via specialized hardware such as ultra-wideband (UWB) transceivers. However, UWB transceivers rely on precise synchronization for accuracy, as synchronization errors may introduce inaccuracies.

Use case: ToF measurements are often selected when accurate distance or range information is required. For example, both WiTrack [29] and WiTrack2.0 [57] use ToF measurements to achieve centimeter-level localization accuracy.

5) *FMCW Radar Measurement:* Radio signals travel fast at the speed of light (i.e., 3×10^8 m/s). It thus requires system clocks with high resolution to directly measure ToF. When ToF is tiny, the frequency shift (i.e., the inverse of ToF), would be a more reliable metric. An FMCW radar provides a practical solution to indirectly measure ToF through corresponding frequency shifts. Specifically, in an FMCW radar system, a chirp signal (modulated in frequency) is transmitted. When the signal hits the target, it is reflected to a receive antenna. The frequency difference between the received signal and the transmitted signal increases with reflection time (i.e., ToF), which is linearly proportional to the distance between the reflector and the radar and changes as the reflector moves.

Figure 7 demonstrates the concept of the sawtooth-modulated FMCW technique. The red line shows the carrier frequency of the transmitted signal which sweeps linearly with time, and the blue dashed line demonstrates the carrier frequency of the received signal with time. The frequency shift Δf between the transmitted and received signals increases with the time shift (i.e., ToF) Δt between them. Let w denote the slope of the sweep. Thus, with the knowledge of w and measured Δf , the ToF Δt can be easily obtained, i.e., $\Delta t = \frac{\Delta f}{w}$. The distance between the transmitter and the receiver can be then calculated as $\frac{c \cdot \Delta t}{2}$. Consequently, the FMCW technique has been extensively adopted to infer HPI,

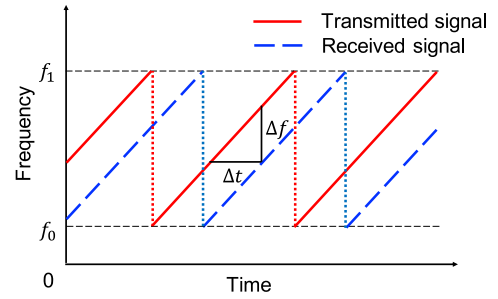


Fig. 7. Transmitted and received signals for an FMCW radar.

including human moving [38], gesture [57], [58], [95], and vital signs [3], [26].

The resolution of an FMCW radar system (i.e., the minimum measurable change in location) equals $\frac{c}{2B}$ [29], where B is the total sweep bandwidth. Thus, to obtain a high resolution, a large sweep bandwidth is required.

Pros: An FMCW radar can achieve high distance resolution due to its large bandwidth, and is highly sensitive to small changes in object position, enabling the estimation of tiny vibration frequencies (e.g., breathing and heartbeat).

Cons: FMCW radar systems require specialized hardware and signal processing techniques to ensure accurate measurements. They are primarily utilized for relative velocity and motion measurements, rather than precise distance estimation.

Use case: FMCW radar systems are widely employed in human motion sensing applications, particularly when high resolution and accurate distance measurements are required [28], [96], [97]. For instance, RF-Capture [28] combines FMCW and antenna arrays to track human motions by estimating the distance and direction between humans and antennas. Similarly, both RF-Pose [96] and Marko [97] adopt FMCW radar equipped with two antenna arrays (horizontal and vertical) to achieve higher localization accuracy by separating signals from different spatial locations.

6) *Doppler Radar Measurement:* In contrast to an FMCW radar, which transmits wireless signals modulated in frequency, a Doppler radar sends a wireless signal with a single frequency. According to the Doppler theory, a target with time-varying movement but zero net velocity will reflect the signal, whose phase is modulated in proportion to the displacement of the target [98].

Figure 8 illustrates the principle of the Doppler effect. The signal sent by the transceiver reflects when it hits the body of the target user. If the target is stationary, the reflected signal will have the same frequency f_0 as the transmitted signal. The Doppler effect occurs when the target moves. If the target moves away from the transceiver with a speed v , the signal is stretched out, which results in increased wavelength and decreased frequency f_1 of the reflected signal. We thus have $f_1 = f_0 - \Delta f$, where Δf ($\Delta f > 0$) is the frequency change caused by the Doppler effect. On the other hand, if the target moves towards the transceiver with a speed v , the signal is compressed, which leads to a shortening of the wavelength and thus increased frequency $f_2 = f_0 + \Delta f$. The frequency change Δf is a function of the target speed v , transmit frequency f_0

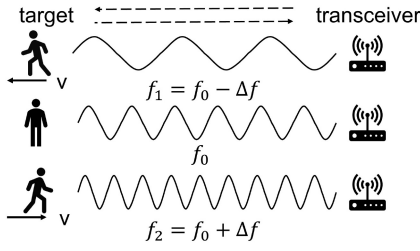


Fig. 8. The Doppler effect.

and speed c of light, i.e., $\Delta f = \frac{2 \cdot v}{c} \cdot f_0$. Thus, by tracking Δf , the movement of the target can be inferred.

With a Doppler radar, we can achieve detecting people moving behind walls [38], gesture recognition [8], and vital signs monitoring [60], [99].

Pros: A Doppler radar has the ability to cover a relatively large area, allowing for monitoring multiple individuals simultaneously or tracking movements in a broad space.

Cons: Doppler shift may not provide highly accurate distance information, limiting its applicability in certain applications that require precise localization.

Use case: Doppler shift has been successfully utilized in numerous human motion sensing applications [8], [100]. It is particularly well-suited for scenarios where relative velocity measurements are desired, prioritizing the detection and analysis of human gestures [8] or vital signs [101], [102] rather than precise position or distance information.

Each wireless channel characteristic has its own set of advantages and constraints. RSS is easily accessible and suitable for coarse tasks, but it may not provide the level of detail required for fine-grained sensing. CSI offers detailed insights into the channel, but it requires physical layer (i.e., PHY layer) access. ToF and AoA provide precise timing and angle information, but they typically require advanced hardware for accurate measurements. FMCW and Doppler measurements offer high-resolution speed and distance data, but they need specialized hardware and complex processing techniques. A comprehensive study [103] demonstrates CSI measurements provide the most robust respiratory rate estimates compared to CIR and RSS. However, the decision on which characteristic to employ depends on the specific requirements of the wireless HPI inference task. Factors to consider include the precision and accuracy needs, the available hardware platforms, and the complexity of the environment.

III. WIRELESS HPI INFERENCE

We first provide an overview of wireless HPI inference systems, and then detail each involved technical component.

A. System Overview

We consider a general scenario, where a user can utilize a transmitter and receiver pair to infer HPI (a transceiver can act as both the transmitter and the receiver in some cases). To create a radio environment, the transmitter transmits the wireless signal, which would be affected by human activities in the physical environment. The receiver receives the signal

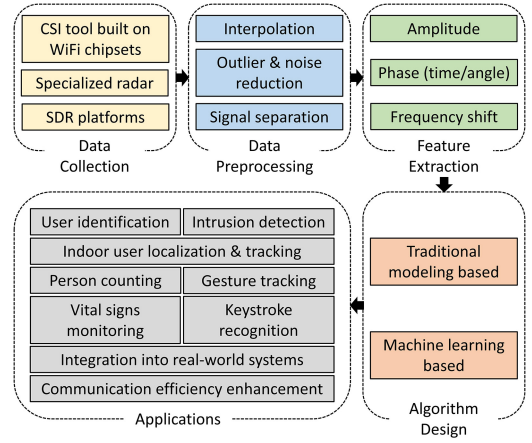


Fig. 9. General overview of wireless HPI inference techniques.

from the wireless channel and computes one or multiple target wireless measurements, as discussed in Section II-B. The collected raw measurements are then fed into the module of data preprocessing, which aims to compensate for packet loss, remove outliers, noise, and uncorrelated components from the data, and help to reduce computational complexity. Next, the unique features (e.g., amplitude, phase, and frequency shift) are extracted from each segment observed in the last step. By inputting these features into a traditional modeling based or machine learning based algorithm, various HPI inference applications can be achieved. Figure 9 presents the general overview of HPI inference techniques via wireless signals.

B. Data Collection

With corresponding hardware, raw wireless measurements can be obtained:

- RSS can be measured with commercial off-the-shelf (COTS) WiFi chipsets [73], [75], [80], [104].
- CSI is available on certain COTS network interface cards (NICs), e.g., Linux 802.11n CSI tool built on an Intel WiFi Link 5300 NIC with a custom modified firmware [64], [87], [105], Atheros CSI tool built on top of ath9k which is an open source Linux kernel driver supporting Atheros 802.11n WiFi chipsets [88], [106], and Nexmon CSI extractor on multiple Broadcom WiFi chipsets [107], [108].
- AoA and ToF can be derived from CSI (e.g., [109]).
- It requires specialized equipment, such as FMCW radar [3], [26] or Doppler radar [60], [102], to make the corresponding technique work.

Besides, wireless HPI inference systems can also be implemented using software-defined radio (SDR) platforms such as Universal Software Radio Peripherals (USRPs) or a Wireless Open Access Research Platform (WARP).

C. Data Preprocessing

The original wireless measurements are inherently noisy and may contain empty (null) values and outliers (i.e., samples that significantly differ from neighboring ones). To improve the data reliability, a phase of data preprocessing is applied.

1) *Interpolation*: Packet loss may occur due to weak signals in certain links caused by NLoS connections. To make all traces have the same sample rate and counter the sampling jitter of the wireless measurement, the measurement data should be interpolated. Linear interpolation can be used to estimate the missing measurement values and obtain evenly spaced measurements in the time domain [59], [110].

2) *Outlier and Noise Removal*: The imperfect wireless measurements can be caused by environmental change induced by activities irrelevant to the target human activity, radio signal interference, or hardware imperfection.

Outlier Removal: Hampel filter is a classical technique to detect and remove outliers in a given series [111], [112]. With this filter, any point within a sliding window falling out of the closed interval $[\mu - \gamma \cdot \sigma, \mu + \gamma \cdot \sigma]$ is treated as an outlier, where μ and σ are the median and the median absolute deviation of the series, and γ is a pre-determined scalar threshold. Each identified outlier is then replaced with the corresponding median.

Noise Removal: The speed of human activity (e.g., hand movement) is often low and the induced signal changes usually lie at the low end of the frequency spectrum. To remove high-frequency noises, a low-pass filter (e.g., Butterworth filter) is a natural choice [1], [16], [54], [82], which has a maximally flat frequency response in the pass band. Though a Butterworth low-pass filter can easily remove out-band noise greatly, it cannot completely eliminate the noise as it has a slightly slow fall-off gain in the stop band. To further remove the noise, different strategies can be taken.

The weighted moving average filter [65], [82] is widely used to further remove random environmental noise from the signals. Specifically, a wireless measurement M_t at time t is averaged by the equation:

$$\bar{M}_t = \frac{1}{\sum_{i=1}^k i} \times (k \cdot M_t + (k-1) \cdot M_{t-1} + \dots + 1 \cdot M_{t-k+1}), \quad (7)$$

where k decides to what degree the current value is related to the historical records. Also, the median filter [1], [54] is able to smooth out the signal while preserving the edges.

Besides, Principal Component Analysis (PCA) can be applied to remove noise from the signals by leveraging correlated variations in wireless measurements of different subcarriers [16]. PCA-based noise reduction can remove the uncorrelated noisy components which can not be removed via traditional low-pass filtering. Meanwhile, PCA has another function of dimension reduction by identifying the most representative components influenced by the target human activity (which discloses the user's HPI) [19].

3) *Signal Separation*: We can further analyze the filtered wireless data stream to understand how the target human activity and the accompanying micro motion impact the observed signals, and thus develop techniques to separate them.

Fast Fourier Transform (FFT) is the most common method that converts a signal from its original domain (often time or space) to a representation in the frequency domain. Short-time Fourier Transform (STFT) is a sequence of Fourier Transforms (FTs) of a windowed signal. It splits a longer time

signal into shorter segments with the same length and then measures the FTs independently for shorter segments. Thus, it provides a time-frequency relationship that describes how frequency components of a signal vary over time, whereas the standard FT just provides the frequency information averaged over the entire signal time interval. We can thus distinguish different types of motion via their respective frequencies in the frequency domain.

Different from the traditional frequency analysis such as FT, Discrete Wavelet Transform (DWT) is the time-frequency analysis which has a good resolution at both of the time and frequency domains [19]. In DWT, the first step splits the original signal into an approximation coefficient vector and a detail coefficient vector; this splitting is applied recursively in several steps (i.e., levels) to the approximation coefficient vector to obtain finer details from the signal [16], [19], [56]. We can use the approximation coefficients to compress the original waveforms to reduce computational costs. To achieve the desired compression using DWT, appropriate wavelet and scaling filters should be selected [16], [19].

D. Feature Extraction

After preprocessing the raw measurements, we often observe a strong correlation between the resultant data stream and the corresponding HPI. The phase of feature extraction is thus applied to choose appropriate features that can uniquely represent the specific HPI. Table III summarizes different features that existing wireless HPI inference studies utilize. In general, we can extract the following three types of features:

1) *Features in Time Domain*: The basic time-domain features include basic statistics such as maximum, minimum, mean, skewness, kurtosis, mean absolute deviation (MAD), variance, standard deviation (STD), and mean crossing rate. Such features represent the variation pattern of the preprocessed signal over time. Various research efforts (e.g., [55], [75]) utilize the mean, variance, or STD of RSS/CSI to distinguish simple scenarios, such as whether there is user motion in an area. For identifying specific human activities (e.g., walking, sitting, jumping, and falling), customized features may be proposed. For example, [65] calculates the Local Outlier Factor (LOF, the ratio of average local densities of one object's neighbors to the local density of the object) to obtain anomaly CSI amplitude patterns, which indicate human activities; [137] uses the state transition of the CSI phase difference for separating fall activities.

Some of the above features cannot be directly used as distinctive features as they can be easily changed by other factors rather than just by the target human activity. For example, the sample maximum/mean/variance can be affected by the transmit power. Instead, we can extract features such as the maximum eigenvalue and the second maximum eigenvalue using a cross-correlation matrix of preprocessed data [64], [82], [119], [145]. Such features are independent of power control and would be distinguishable for different human activities. Thus, these features would facilitate the inference of human activities (e.g., gesture or motion).

TABLE III
SUMMARY OF EXTRACTED FEATURES

Type	Modality	Features	Existing Wireless HPI Inference Work
In time domain	RSS	basic statistics	[54], [55], [70], [72], [75], [78], [104], [113]–[118]
		correlation matrix based	[119]
	CIR	basic statistics	[37], [58]
		basic statistics	[64], [82], [108], [112], [120]–[130]
	CSI	correlation matrix based	[37], [59], [63], [64], [84], [85], [91], [105], [131]–[136]
		customized	[65], [137]
		correlation matrix based	[15], [37], [53], [61], [68], [69]
In frequency domain	ToF	correlation matrix based	[61], [92]
	RSS	FFT/energy/entropy peaks	[56], [79], [80]
		basic statistics	[138]
	CSI	correlation matrix based	[139]–[141]
		FFT/energy/entropy peaks	[4], [67], [142]
	FMCW	basic statistics	[58]
		FFT/energy/entropy peaks	[3], [26]
In time-frequency domain	Doppler	Doppler frequency	[60]
	RSS	time-domain features + frequency-domain ones	[118]
	CSI	spectrogram (of three dimensions: time, frequency, and amplitude)	[10], [137]
		time-domain features + frequency-domain ones	[91], [120], [143]–[145]
		waveform shape/image	[146], [147]

2) *Features in Frequency Domain*: Different from that in the time domain, the basic statistics based features in the frequency domain mainly contain carrier frequency offset (CFO), the spectrogram magnitude, energy, entropy, difference, and percentile frequency components (PFC). Such features reveal the changes in the frequency domain of wireless signals. For example, the study [138] exploits CFO to secure wireless transmission for privacy protection and authentication with an accuracy of 93.2% on average.

With the Fast Fourier Transform (FFT) technique, we can convert the signal in the time domain to its frequency spectrum. The spectrum influenced by the periodic human activity (e.g., breathing) normally has a strong component (i.e., FFT peak) close to the frequency of the activity [56]. In addition, peaks of power spectral density (PSD) are always employed to analyze periodic human activities. PSD depicts the power distribution of temporal CSI measurements on each subcarrier in the time-frequency domain. A strong periodic signal generates a peak at the frequency corresponding to its period in PSD. For example, in a two-person scenario, the first two highest peaks present the breathing rates of these two persons in PSD [4], [67], [142].

Besides, we can extract features such as energy (which denotes total energy in all frequencies), and entropy (which measures the impurity in the signal) in the frequency domain. Also, with Doppler or FMCW radar, we can obtain features such as Doppler frequency or frequency shift. Such features can be correlated with HPI, e.g., vital signs [60].

3) *Features in Time-Frequency Domain*: After obtaining the time-frequency profile of the preprocessed signal, we can extract some features exhibiting properties in both time and frequency to fully describe the specific HPI. For example, WiWho [143] leverages the features of CSI signal in both the time domain and frequency domain to identify a person with an average accuracy of 92% to 80% from a group of 2 to 6 people respectively. Also, Short-time Fourier transform (STFT) is one common method to finish the extraction,

which transforms the temporal waveforms to spectrograms so that waveforms can be analyzed in the time-frequency domain [10], [110], [128]. In addition, the study [147] transforms CSI measurements from multiple channels into a radio image (with time as the x-axis and channel as the y-axis), and then extracts color and texture features from the image for HPI inference. Another work [146] chooses CSI waveform shapes as an activity feature that contains both time and frequency information. It can then classify four exercise activities (e.g., dumbbell lifting, deep squatting, kicking, and boxing) with average recognition accuracies of 97.8% and 91.2% in LoS and NLoS scenarios, respectively.

E. Inference Design

With extracted features, we can then utilize either a traditional modeling based or machine learning based algorithm, as outlined in Table IV, to initiate the inference procedure, i.e., mapping each unrecognized wireless measurement segment (or feature) to the corresponding HPI.

1) *Traditional Modeling Based*: Traditional modeling based algorithms formalize relationships between variables in the form of mathematical equations. We can thus directly infer target HPI by applying corresponding equations. In the following, we introduce several popular models for inferring HPI.

Fresnel Zone (FZ): In the context of wireless communication, as shown in Figure 10, Fresnel Zones refer to concentric ellipses with the transmitter (TX) and receiver (RX) at two focal points, and denote regions of different wireless signal propagation strengths between TX and RX. For a given radio wavelength λ , each ellipse can be obtained by ensuring

$$|TX, U_n| + |RX, U_n| - |TX, RX| = n\lambda/2, \quad (8)$$

where U_n is a point in the n^{th} ellipse and $|p_1, p_2|$ denotes the Euclidean distance between the two points p_1 and p_2 . The innermost ellipse is the first FZ, representing the region through which the direct LOS signals can pass. The n^{th} (when

TABLE IV
SUMMARY OF TRADITIONAL MODELING BASED AND MACHINE LEARNING BASED INFERENCE ALGORITHMS

Type	Modality	Model	Work
Traditional modeling based	RSS	threshold model	[76], [78], [152]
		peak-valley detection based model	[56], [79]
		optimization based model	[70], [73], [153], [154]
	CSI	Fresnel Zone	[129], [142], [155]–[157]
		threshold model	[19], [86], [122], [123], [133]
		peak-valley detection based model	[4], [67], [105], [112], [158]
		optimization based model	[62], [150]
	AoA	optimization based	[15], [37], [53], [61], [68], [159]
	ToF	optimization based	[61], [92]
	FMCW	optimization based model	[3], [26]
Machine learning based	RSS	unsupervised learning	[78]
		supervised learning	[54], [114], [116], [118], [160]
	CSI	unsupervised learning	[4], [67], [161]
		semi-supervised learning	[86], [162]
		supervised learning	[136], [144], [163]
		deep learning based	[135], [163]–[166]
	FMCW	supervised learning	[58]
	Doppler	supervised learning	[101]

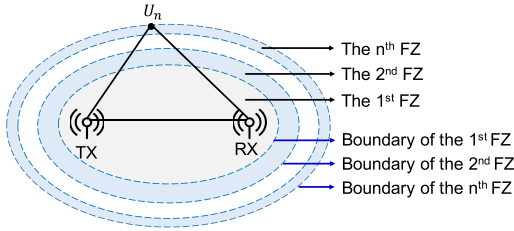


Fig. 10. Illustration of Fresnel Zones.

$n \geq 2$) FZ corresponds to the region between the $(n - 1)^{\text{th}}$ and n^{th} ellipses.

The received signal at RX is a linear combination of reflected and LOS signals. The distance difference ΔD (i.e., $n\lambda/2$) between the two paths generates a phase difference of $2\pi \cdot \Delta D / \lambda = n\pi$ between the two signals. As the phase shift introduced by the reflection is π , the total phase difference $\Delta\phi$ between reflected and LOS signals equals $(n + 1)\pi$. Thus, if n is even, we obtain $\Delta\phi \bmod 2\pi = \pi$, causing the two signals arrived at RX to have opposite phases and destructively interfere with each other. In contrast, we have $\Delta\phi \bmod 2\pi = 0$ if n is odd, i.e., both signals have the same phase and constructively interfere with each other to form a boosted signal. The FZ model can thus help reveal the signal change pattern (i.e., sensitive or insensitive) in each subcarrier (with different waveforms) caused by human activities, such as respiration [142].

Threshold Model: A threshold model is any model where a threshold value, or a set of threshold values, is utilized to distinguish value ranges, each of which discloses a specific pattern of the target HPI. In a simple case, RSS or amplitude of CSI is in a different range when the target user is in different status (e.g., moving or static). By setting appropriate thresholds for extracted features, we can do a fine-grained analysis of the identified human activity. Also, we can utilize similarity analysis to recognize different signal patterns. The signal similarity can be quantified by a commonly used method, such as Euclidean distance, dynamic time warping (DTW) distance [148], or Earth Mover's Distance

(EMD) [149]. Specifically, we first measure the similarity between the collected data stream with a reference signal, and then set up a similarity threshold (usually based on empirical data) to determine whether two signals are similar or not.

Peak-valley Detection based Model: Peak-valley detection is widely used for recognizing periodic human activities such as respiration and walking. The wireless measurements (e.g., CSI amplitudes, RSS) may exhibit a periodic change pattern over time, corresponding to periodic motion. Such a pattern can be identified by peak-valley detection in the time or frequency domain. For example, the studies [4], [67] detect the peaks (especially peak-to-peak intervals) of CSI amplitudes in the time domain to calculate the breathing cycle as the time series of corresponding CSI amplitudes presents a sinusoidal-like periodic changing pattern when the user breathes, and in the frequency domain, [112] identifies the breathing rate within a segment of a time signal as the location of the peaks of the FFT of this segment.

Optimization based Model: Some HPI interference problems (e.g., human trajectory tracking) can be formulated as an optimization problem, i.e., a process of optimizing an objective function concerning some variables in the presence of constraints on those variables. Reference [150] achieves indoor localization of users carrying WiFi-enabled devices by modeling constraints of the physics of wireless propagation and solving the established objective function. Also, the Kalman filter is a widely used technique to dynamically estimate unknown variables with a series of temporal measurements [151], which recursively applies two phases: prediction and correction. The prediction phase uses the state estimate from the previous time step to generate a state estimate for the current time step. In the correction phase, the obtained prediction result will be combined with current observation information to refine the state estimate. For example, WiDeo [62] builds an optimization problem with variables (i.e., AoA, ToF, and signal strength measurements) and utilizes the Kalman filter to update parameters (e.g., the

current position of the object, velocity) of the motion model.

The pros and cons of traditional modeling based algorithms are summarized as follows:

Pros:

- 1) They may require minimal or no training data, model training, and ground truth annotation.
- 2) The inputs and outputs are based on cause-and-effect relationships. Thus, the outcomes are derived from well-understood principles.
- 3) They are typically characterized by low computational costs and more efficient implementation.

Cons:

- 1) They need more effort to build appropriate models and select optimal model parameters, requiring domain-specific knowledge for setup and adjustments.
- 2) They are often designed for specific tasks, and thus may not be suitable for tasks that require adapting to evolving data or capturing intricate patterns in large datasets.
- 3) They could become inefficient or inaccurate for very complex systems.

Use Case: They are mainly employed in applications that involve estimation, where there's a critical need for accurate numerical value predictions.

2) *Machine Learning Based:* Machine learning based methods involve two steps, namely, training and testing. The training step builds a classification model for the target HPI using supervised, semi-supervised, or unsupervised learning. In the testing step, an observed feature for an unknown piece of HPI is then matched within the classifier from the training step to determine which piece of HPI it corresponds to.

Supervised Learning: Supervised learning refers to the training process performed with labeled training data. By recording each piece of specific HPI (i.e., a gesture or keystroke) and the corresponding feature, a training model for classification can be built. A variety of supervised learning algorithms have been used for HPI inference, including

- **Decision Tree (DT):** DT uses a tree-like model of decisions and their possible consequences. It can output simple if-else classification models that are useful in understanding the importance of different CSI features for recognizing different types of motions [145].
- **Hidden Markov Model (HMM):** An HMM consists of a discrete-time, discrete-state Markov chain among states and an observation (probability function) of each state. The states are hidden, and only the observations are available. With a sufficient number of representative training samples of an activity, an HMM can be built to implicitly model all of the main sources of variability inherent in the activity [139]. HPI inference can be regarded as a probability problem and thus achieved with HMMs (e.g., [54], [124], [139], [140]).
- **K-Nearest Neighbor (kNN):** Recent work has demonstrated success in applying kNN classifiers to obtain HPI [16], [89], [167]. KNN decides on the assignment of a new signal in the feature space leveraging majority voting of k -nearest neighbors (k is a positive integer).

- **Support Vector Machine (SVM):** SVM is a classical discriminative classifier formally defined with a separating hyperplane, and it can be used to distinguish different patterns of a specific HPI, e.g., human presence and absence [84], [132]. Also, with the training data that has only one class (i.e., "normal" class), one-class SVM, which is an unsupervised and extended algorithm of SVM, can be then used to infer the properties of normal cases and from these properties can predict which examples are unlike the normal examples. For example, [65] utilizes a one-class SVM to distinguish failing from the other movements.

Unsupervised Learning: With unsupervised learning, where only output data rather than labeled sets of input-output pairs are given, some work can still successfully infer HPI. K -means clustering is one of the popular unsupervised machine learning algorithms. It partitions n observations into k ($k \leq n$) sets in which each observation belongs to the cluster with the nearest mean (i.e., cluster center). References [4], [67] utilize the k -means clustering technique to separate the breathing rates of two users when they breathe simultaneously. Reference [78] applies k -means clustering to generate the crowd density levels and types based on the collected RSS from the training phase.

Semi-supervised Learning: Semi-supervised learning falls between unsupervised learning (without any labeled training data) and supervised learning (with completely labeled training data). Such methods can use readily available unlabeled data to improve supervised learning tasks when the labeled data are scarce or expensive [168]. As an example, [162] develops a semi-supervised learning approach leveraging the non-linear regression model to accurately estimate the number of participants with significantly reduced training efforts.

Deep Learning: The performance of the above conventional machine learning algorithms heavily relies on the input feature (i.e., representation of the given wireless measurements). However, it is difficult to select the right set of features sometimes. Deep learning provides a solution that can automatically learn feature representation from raw data. Specifically, we introduce two typical deep learning models for HPI inference.

- **Deep Neural Networks (DNN):** DNN can extract features from preprocessed wireless measurements without manual and subjective feature selection and has been used for various HPI inference, e.g., tracking human movements [135], intruder detection [163], and recognizing different types of human activities [164].
- **Convolutional Neural Network (CNN):** CNN, as a specialized kind of neural network, is able to implicitly extract features from raw data, and has also been successfully applied in HPI inference. For example, by feeding CSI measurements to a CNN, [165] can track the user with the target device, [166] can classify sign gestures, and [163] can detect an intrusion or monitor an independently living elderly person.

The pros and cons of machine learning based algorithms are outlined below:

Pros:

- 1) They require minimal or no demand for signal processing.

- 2) They can handle multi-dimensional data and complex decision boundaries.
- 3) They are capable of automatically identifying relevant features, especially with deep learning techniques.
- 4) They can predict unseen data when the model is properly trained.

5) They apply to a diverse range of tasks.

Cons:

- 1) They typically require significant amounts of labeled data.
- 2) They have the overfitting risk of performing poorly on unseen data if not properly regularized or validated.
- 3) Complex models are difficult to interpret and explain.
- 4) Especially with deep learning, a training phase can be computationally expensive.
- 5) Models are susceptible to inheriting biases embedded in their training datasets.

Use Case: They are primarily utilized for tasks that involve the extraction of patterns or insights from data to make predictions, recommendations, or decisions.

Both traditional modeling based and machine learning based algorithms have their strengths and weaknesses, and the choice often depends on the specific task, available resources, and domain knowledge. In many modern solutions, a hybrid approach combining aspects of both can often yield optimal results. For example, WiWho [143] employs a peak-valley detection algorithm to construct the step cycle. Subsequently, user identification is achieved using a decision tree based machine learning classifier, leveraging both step and walk analysis. BodyScan [125] exploits a threshold-based algorithm to determine if a user is stationary or engaged in an activity. After that, a Support Vector Machine (SVM) is leveraged to classify five common activities of daily living, including walking, brushing teeth, typing on a phone, shaking hands, and typing on a keyboard.

This section provides a general framework of human motion sensing, illustrating how wireless signals are processed and used to accomplish specific sensing tasks. It is crucial to understand that as the task complexity increases, a combination of various techniques is often employed to ensure improved performance. Additionally, more advanced approaches, such as transfer learning or multimodal sensing, may be integrated into the framework to tackle more complex tasks.

IV. APPLICATIONS

In this section, we present various applications leveraging wireless HPI inference, including user identification, intrusion detection, indoor user localization & tracking, person counting, gesture tracking, vital signs monitoring, and keystroke recognition. For each application, we introduce corresponding motivation and techniques with different wireless measurement modalities (e.g., RSS and CSI).

A. User Identification

User identification is the process of verifying the identity of a person. It plays a central role in securing personal devices and protecting personal privacy. Also, user identification can

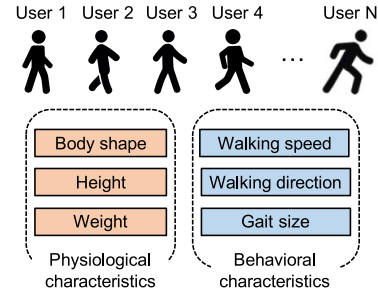


Fig. 11. Different persons may have different physiological and behavioral characteristics.

help develop new applications in smart homes, offices, and transportation. For example, with user identification, a smart building can identify a user walking along the corridor. As a result, when the user approaches her/his office, the door can be opened automatically. Also, the driver authentication system may improve security and automatically make in-vehicle driver-specific adjustments (e.g., temperature and seat) [141].

Traditional user identification systems include biometric-based or knowledge-based systems. The former technique often requires setting up extra devices (e.g., cameras, and fingerprint scanners). For the latter, users may find difficulties in remembering their PINs or passwords. Recent work can leverage already available wireless signals (e.g., WiFi and cellular) to extract the physiological and behavioral characteristics of a person, as shown in Figure 11. Physiological traits refer to the physical functions of a human, such as body shape, height, and weight, while behavioral attributes are based on the behavior of a person, such as walking speed, walking direction, and gait size. As such characteristics are unique and hard-forged, automatic user identification can be achieved. Therefore, wireless user identification systems neither require any hardware deployment nor need users to memorize the secret. Table V summarizes existing wireless user identification methods.

1) *RSS-Based:* RSS-based approaches infer user identification via analyzing the pattern in RSS measurements. Different users may have different physical constraints such as moving areas and walking paths, leading to different RSS sequences. Motivated by this observation, the study [54] analyzes and exploits the characteristics of RSS sequences generated by mobile devices, and extracts useful features from different users for user identification. This system achieves an average identification accuracy of 90.83% with a false negative rate of 1.11% and a false positive rate of 17.22%.

2) *CSI-Based:* The intuition behind CSI-based user identification methods lies in the idea that CSI measurements can reveal human gait. Each person's gait, such as their walking style, is unique and characterized by differences in limb (hand and feet) movement patterns and velocity [173]. The wireless signals reflected by a walking human generate distinctive variations in the CSI at the receiver. Motivated by this observation, [10] proposes to utilize CSI spectrograms to characterize fine-grained walking patterns and, consequently, recognize humans, with an average accuracy of 92.31% in

TABLE V
SUMMARY OF EXISTING USER IDENTIFICATION TECHNIQUES VIA WIRELESS SIGNALS

Work	Modality	Activity	Implementation	Scenario	Average Identification Accuracy
Cheng: Mobihoc 16 [54]	RSS	walking, stationary	TX: RB912UAG-2HPnD router, RX: mobile devices	office, laboratory, dormitory	90.83%
Wang: UbiComp 16 [10]	CSI	walking	TX: NetGear JR6100 WiFi router, RX: laptop equipped with Intel 5300 NIC	laboratory	92.31% for 10 subjects, 79.28% for 50 subjects
Zeng: IPSN 16 [143]	CSI	walking	1 TX: Asus RT-AC66U 802.11n WiFi router with 3 omnidirectional antennas, 1 RX: laptop equipped with an Intel 5300 8 NIC and 3 omnidirectional antennas	three rooms with varying sizes and layouts	92% to 80% from a group of 2 to 6 subjects, respectively
Zhang: DCOSS 16 [1]	CSI	walking	1 TX: Netgear R7000 with 3 antennas, RX: laptop equipped with an Intel WiFi link 5300 NIC and 3 antennas	corridor	93% to 77% for a group of 2 to 6 subjects, respectively
Chen: IMWUT 17 [120]	CSI	walking	a TX-RX pair: each designed node consisting of 1 microphone, 1 Intel 5300 NIC, and 1 omnidirectional antenna, controlled by HMB mini-computer	laboratory, corridor, meeting room	92% to 82% from a group of 2 to 6 subjects, respectively
Liu: TMC 17 [169]	CSI	moving, stationary	1 TX: Linksys E2500, 2 RX: each is a laptop equipped with Intel 5300 NIC	laboratory, apartment	98.4% for 2 users
Shi: Mobihoc 17 [128]	CSI	walking, stationary	a TX-RX pair: each is a laptop equipped with an Intel 5300 card	office, apartment	91.2% for 11 subjects (office); 92.4% for 5 subjects (apartment)
Li: INFOCOM 20 [170]	CSI	arm gestures	a TX-RX pair	—*	96.74% for user identification accuracy
Wang: JIOT 22 [171]	CSI	walking, daily activities (e.g., jumping)	1 TX: TP-Link N750 routers with 1 antenna, 1 RX: TP-Link N750 routers with 3 antennas	a lab, a cubic office	For a group of 2 to 15 subjects: 99.24% to 88.94% (lab); 99.18% to 87.69% (cubic)
Lin: TOSN 23 [172]	CSI, AoA, ToF	hand and arm activities	1 TX: TP-Link TL WR886N router, 1 RX: a laptop equipped with Intel 5300 NIC	indoor, corridors, outdoor	92.6% for multi-activities, 97.1% under multi-scenes

*— denotes unspecified.

identifying a person from a group of 10 subjects. WiWho [143] and WiFi-ID [1] analyze both time and frequency domain features of CSI measurements to identify a person's walking gait for user identification. By comparison, WiWho achieves an average accuracy ranging from 92% to 80% for a group of 2 to 6 people, while such values of average accuracy for WiFi-ID are 93% and 77%. These techniques [1], [10], [143], however, only capture human walking gait patterns and apply them to a small group of people (i.e., 2 to 10). Moreover, WFID [66] exploits a novel feature of subcarrier-amplitude frequency (SAF) and achieves the best accuracy of 93.1% and 91.9% for groups of 6 and 9 subjects, respectively. This is based on three human activities including standing still, marching on the same spot, and walking. A recent work [128] extracts unique physiological and behavioral characteristics inherited from people's daily activities including both walking and stationary activities (e.g., operating appliances). It achieves an average accuracy of 91.2% with 11 subjects in an office environment and 92.4% with 5 subjects in an apartment environment.

Challenges: User identification via wireless signals still presents notable challenges. The existing systems work in a controlled environment for a limited number of users. However, environmental variations and individual differences can significantly impact signal propagation and the accuracy of user identification. But in real life, it is impossible to retrain a model for a new environment and new users. In this way,

the performance of these systems may degrade when applied to unseen environments and more users.

B. Intruder Detection

To secure residential and commercial properties, intruder detection systems have attracted a great deal of interest. Traditional methods primarily utilize cameras [177] or passive infrared (PIR) sensors [178] to detect intruders who enter the target home or office. The cameras identify an intruder via image processing algorithms and the PIR sensors detect an intruder via a few infrared rays emitted from the intruder. However, the camera-based techniques may not work under the influence of poor lighting or NLoS environment, while PIR-based techniques are insensitive to very slow motions or objects in standing mode [179]. Also, the performance of camera-based or PIR-based methods depends on the installation location and the number of sensors or cameras [74].

Figure 12 presents the typical structure of a wireless intruder detection system (e.g., [75], [180]), consisting of transmitters or access points (APs) and receivers or monitoring points (MPs). The system operates in two phases. In the first phase (training), the system utilizes RSS/CSI received at the MPs during periods without human motion to construct a silence profile. In the second phase (monitoring), the system analyzes the collected readings at the MPs to determine whether there is intruder activity based on the information

TABLE VI
SUMMARY OF EXISTING INTRUDER DETECTION TECHNIQUES VIA WIRELESS SIGNALS

Work	Modality	Implementation	Scenarios	Accuracy
Youssef: MobiCom 07 [72]	RSS	2 TX: each is a Cisco Aironet 350 Series AP, 2 RX: each is a device with Orinoco Silver card	—*	Controlled environment: 100% TPR, 100% TNR
Moussa: PerCom 09 [76]	RSS	2 TX: each is a Cisco 1130 AP, 2 RX: each is a laptop with Orinoco Gold card	laboratory	Controlled environment: 100% TPR; 100% TNR; real environment: TPR: 90%
Kosab: PerCom 12 [75]	RSS	4 TX: each is Cisco Aironet 1130AG series AP, 3 RX: each is a laptop equipped with D-Link AirPlus G+ DWL-650+ Wireless NIC	office, home	F-score: at least 0.93
Xiao: ICPADS 12 [64]	CSI	TX: TL-WR941ND router with 3 antennas, RX: a laptop equipped with a three-antenna Intel WiFi Link 5300	laboratory, corridor	TPR>70% when FPP≤1% (lab); TPR>90% when FPR is around 9% (corridor)
Zhou: INFOCOM 13 [174]	CSI	TX: TP-LINK TLWR741N wireless router, RX: laptop equipped with Intel 5300 NIC	conference hall, laboratory	8% FPR and 7% FNR in 4 directions
Qian: ICPADS 14 [84]	CSI	TX: TP-link wireless router with 1 antenna, RX: mini PC with 3 antennas	laboratory, offices and classrooms	TPR: 97%, TNR: 98%
Wu: JSAC 15 [105]	CSI	TX: TP-LINK TL-WR741N wireless router, RX: a laptop equipped with Intel 5300 NIC	classroom, laboratory	TPR: 95%, TNR: 95%
Zhu: JSAC 17 [59]	CSI	TX: TL-WR742N with 1 antenna and TL-WR841N with 2 antennas, RX: a laptop equipped with Intel 5300 NIC	meeting room, office	TPR: 99%, TNR: 99%
Ikeda: VTC 08 [74]	AoA	TX: Dipole antenna, RX: 8-element linear array	room equipped with metal walls and partition	With 8-element antenna array and optimum thresholds: 0 FPR and 0 FNR
Wang: TVT 21 [175]	AoA	TX: a laptop equipped with Intel 5300 NIC and 1 antenna, RX: a mini PC equipped with Intel 5300 NIC and 3 antennas	meeting office, room	accuracy of 94.3% for glass occlusion, 93.1% for brick wall
Ni: JSEN 22 [176]	FMCW	TI 77 GHz FMCW radar IWR1443 EVM and DCA1000 EVM	corridor	accuracy of 88.59% in identifying up to 5 subjects (1 known user and 4 intruders)

*— denotes unspecified.

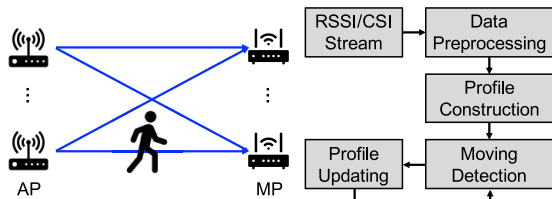


Fig. 12. Typical structure of an intruder detection system via wireless signals.

gathered in the training phase. Additionally, it continuously updates its silence profile to adapt to environmental changes.

The following metrics are often utilized to evaluate the performance of such a system.

- **True Positive Rate (TPR) & False Negative Rate (FNR):** TPR (also referred to as *recall*) is the probability that an intruder is successfully detected, while FNR equals the probability that an intruder passes the system without being detected (i.e., $FNR = 1 - TPR$).
- **True Negative Rate (TNR) & False Positive Rate (FPR):** TNR is the probability that a static environment (i.e., with no intruder) is correctly classified, while FPR equals the probability that the system determines incorrectly that there is an intruder in a static environment (i.e., $FPR = 1 - TNR$).
- **Precision:** This is the fraction of cases where there is indeed an intruder among all cases where the

system detects an intruder (correctly and incorrectly),
i.e., $Precision = \frac{TPR}{TPR+FPR}$.

- **F-score:** F-score (F_1 score) provides a single value to measure the system effectiveness, and is defined as the harmonic mean of the corresponding precision and recall values, i.e., $F_1 = \frac{2}{\text{recall}^{-1} + \text{precision}^{-1}}$. It ranges from 0 to 1, and a higher precision or recall leads to a higher F-score.

Table VI compares the performance of existing studies using wireless signals to achieve intruder detection.

1) *RSS-Based*: The study [72] firstly defines the concept of Device-free Passive (DfP) localization and introduces a DfP localization system, which can be used to track intruders. This system works by monitoring and processing changes in RSS to detect changes (i.e., intruder movements) in controlled environments with a 100% TPR and a 100% TNR, while [76] points out that the performance of such a system degrades significantly when tested in a real environment. Reference [76] then develops an alternative algorithm based on the maximum likelihood estimator (MLE) and can achieve 90% recall in a real environment. Reference [75] presents another DfP localization system, which achieves an F-score of at least 0.93 and outperforms the previously proposed techniques [72], [76] in terms of robustness and accuracy.

2) *CSI-Based*: Due to the multipath effect in indoor environments, RSS is often susceptible to the measurement itself. Consequently, the slow dynamic can be easily hidden by the inherent RSS variance, which may lead to miss detection [64].

Compared with coarse-grained RSS, CSI enables better performance of intruder detection as fine-grained PHY layer information. Based on the insight that CSI maintains temporal stability in a static environment while exhibiting burst patterns when motion takes place, [64] proposes a CSI-based motion detection system, and tests them in two different scenarios. In the lab environment, for a FPR less than or equal to 1% the detection rate would be greater than 70%; in the corridor environment, the detection rate would be greater than 90% when FPR is around 9%. Reference [174] exploits the multipath components as signatures to detect human presence in a reliable and omnidirectional manner, with an average FPR of 8% and an average FNR of 7% in 4 directions. Reference [84] leverages full information (both amplitude and phase) of CSI and is the first to incorporate meaning phase information for intruder detection, achieving a TPR of 97% and a TNR of 98%. Reference [105] also takes advantage of both amplitude and phase information of CSI to detect moving targets. Also, it detects static person by detecting rhythmic human breathing, which induces repetitive changes in received signals.

Usually, walls may cause severe signal attenuation. The study [59] shows with experiments that [84] and [105] suffer from great performance degradation in through-the-wall (TTW) scenarios, and also proposes a novel scheme for device-free TTW detection of moving humans.

3) *AoA-Based*: Except RSS and CSI, AoA can also be utilized for intruder detection. For example, [74] exploits the direction-of-arrival (DoA, i.e., AoA) of incident signals on an antenna array to detect indoor events such as intrusion. It can achieve $FPR = 0$ and $FNR = 0$ with an 8-element antenna array and optimum thresholds of the built cost function.

Challenges: Detecting intruders in indoor settings becomes challenging when faced with significant noise and occlusions, potentially leading to false alarms or missed real threats. Also, an advanced intruder may mimic authorized users' signals to bypass the detection systems, resulting in new threats.

C. Indoor User Localization & Tracking

The proliferation of wireless communication and smart devices with sensing capabilities has given rise to a growing interest in location-aware systems or services, such as personal navigation, mobile tourist guides, and mobile advertising. The technique of user localization and tracking plays a pivotal role in achieving those systems or services. In outdoor environments, mobile devices are often localized with the Global Positioning System (GPS) technology, while in indoor environments, GPS signals are severely attenuated, and thus the GPS positioning fails. Instead, wireless signals have been widely adopted to pinpoint mobile devices or users in indoor environments. Table VII presents the comparison of such studies. Since user movement near wireless communication links may cause variance in received wireless signals. Such variance can be thus used to infer the user's location change.

1) *RSS-Based*: Reference [104] introduces a radio-frequency (RF) based system called RADAR, which combines

RSS measurements and signal propagation modeling to localize and track users inside buildings. The median error distance of RADAR is 2 to 3 meters. Unlike RADAR which relies on extensive measurement to map the RF environment, [150] proposes a configuration-free indoor localization scheme that uses existing WiFi infrastructure rather than a pre-deployment effort to localize mobile devices. Usually, locating interior movement from the outside of a building can help people (e.g., emergency responders and policy) make life-saving decisions. Accordingly, [70] leverages RSS variance to track the location of a person or object behind walls, without the need for an electronic device to be attached to the target.

The aforementioned approaches [70], [104], [150] focus on single-target tracking, while [152], [153] realize real-time multiple target tracking with RSS measurements of wireless packets exchanged between the sensors in the network. Specifically, [153] achieves a root-mean-squared error (RMSE) tracking accuracy of approximately 0.3 m for a single target, 0.7 m for two targets, and 0.8 m for three targets, while the highest average tracking error for [152] is 0.45 m with two targets, 0.46 m with three targets, and 0.55 m with four targets.

2) *CSI-Based*: Indoor positioning systems based on RSS are easily affected by the temporal and spatial variance due to the severe multipath effect in indoor environments. To alleviate the multipath effect at the receiver, a great number of research efforts (e.g., [63], [109], [122], [130], [182], [185], [186]) leverage more reliable CSI to achieve user localization and tracking. For example, [63] utilizes temporal stability and frequency diversity of CSI to design a passive indoor localization system. Unlike approaches (e.g., [63], [130]) which often require the dense deployment of APs, [109] proposes a single-AP indoor localization system, which can capture the user's location with a mean error of 2.3 m.

In addition, [122] presents a model-based device-free localization system without requiring any explicit pre-deployment effort or exhaustive fingerprint collection. The proposed system can achieve a median accuracy of 0.5 m and 1.1 m in LoS and NLoS scenarios, respectively. Furthermore, [182] builds a theoretical model that geometrically quantifies the relationships between CSI dynamics and the user's location and velocity, achieving decimeter-level accuracy, with a median location error of 25 cm given initial target positions and 38 cm without them.

3) *AoA-Based*: AoA has also been utilized for indoor user localization and tracking. For instance, [53] proposes a fine-grained indoor location system leveraging AoA to locate indoor wireless clients, with a median 23 cm localization error.

4) *FMCW-Based*: As mentioned earlier, FMCW can help to measure ToF. Reference [29] WiTrack tracks the 3D motion of a user using radio reflections that bounce off the body. Based on ToF to measure the distance, on average, WiTrack localizes the center of a human body within a median of 10 to 13 cm in the x and y dimensions, and 21 cm in the z dimension. Moreover, WiTrack2.0 [57] is improved to localize static users due to breathing and multiple users. Specifically, it can localize up to five people simultaneously with a median accuracy of 11.7 cm in the x/y dimensions.

TABLE VII
COMPARISON OF USER LOCALIZATION & TRACKING STUDIES

Reference	Modality	Scenarios	Implementation	Accuracy
Bahl: INFOCOM 00 [104]	RSS	second floor of a 3-story building (43.5 m $\times$ 22.5 m)	3 base stations: each is Pentium-based PC running FreeBSD 3.0 equipped with a wireless adapter	median error: 2-3 m
Chintalapudi: MobiCom 10 [150]	RSS	small building with 48 APs; large building with 156 APs	WiFi APs	median error: 2 m (in small building); 7 m (in large building)
Wilson: TMC 11 [70]	RSS	through-wall environment with 34 nodes	*_	average error: 2 ft (tracking), 1.5 ft (localization)
Nannuru: TMC 13 [153]	RSS	an indoor area with 24 nodes; a lab with 24 nodes; a through-wall environment with 28 nodes	each node is system-on-chip (SoC) TI CC2530 device	RMSE: 0.3 m for 1 target, 0.7 m for 2 targets, and 0.8 m for 3 targets
Bocca: TMC 14 [152]	RSS	an open indoor environment with 30 sensors; one bedroom with 33 sensors; an office with 32 sensors	each node is TI CC2531 USB dongle node	average tracking error: 0.45 m, 0.46 m, and 0.55 m for 2, 3 and 4 targets, respectively
Abbas: PERCOM 19 [181]	RSS	122 APs in a 37 m $\times$ 17 m floor, 59 APs in 14.5 m $\times$ 4.5 m apartment	TX: WiFi AP, RX: Android phones	average accuracy: 2.64 m (floor), 1.12 m (apartment)
Wu: INFOCOM 12 [130]	CSI	chamber; laboratory with 3 APs; lecture hall; corridor with several APs deployed in multiple rooms	TX: TPLINK TL-WR941ND router, RX: a laptop Intel WiFi Link 5300 NIC	median accuracy: 0.45 m (laboratory), 1.2 m (lecture hall), 1.2 m (corridor)
Xiao: ICDCS 13 [63]	CSI	laboratory; lobby (both with 2 pairs of APs and detecting points)	*_	percentage of correct localization attempts: 90% - 98%
Mariakakis: MobiSys 14 [109]	CSI, ToF	office environment with a single AP	HP MSM 460 AP using Atheros 9590 chipset	mean error: 2.3 m
Wang: MobiCom 16 [122]	CSI	home; part of a library; classroom (each with 4 APs and 7 clients)	each is a laptop equipped with Intel 5300 NIC	median error: 0.5 m (LoS); 1.1 m (NLoS)
Qian: Mobihoc 17 [182]	CSI	1 one-antenna transmitter and 2 three-antenna receivers in a 4 m $\times$ 4 m area	each is a mini-desktop equipped with Intel 5300 NIC	median error: 25 cm and 38 cm (with and without initial positions respectively)
Xiong: NSDI 13 [53]	AoA	office with 6 APs	TX: WARP equipped with 4 radio front ends and 4 omnidirectional antennas, RX: Soekris boxes equipped with Atheros 802.11g radios	median error: 23 cm
Tong: TNET 21 [183]	CSI, AoA, ToF	4 APs in 100 m^2 corridor, 5 APs in 255 m^2 laboratory	each is mini-PC equipped with Intel 5300 NIC	achieves 80% localization errors within 0.3 m
Ding: JIOT 22 [184]	CSI, AoA	5 m $\times$ 3 m $\times$ 4 m laboratory, 8 m $\times$ 2 m $\times$ 4 m corridor	each is a device equipped with Intel 5300 NIC	mean error: 0.8 m (laboratory), 1.1 m (corridor)
Adib NSDI 14 [29]	FMCW	VICON room	FMCW front-end operating as a daughterboard for the USRP software radio	median error: 10 cm, 13 cm, and 21 cm along the x , y and z dimensions
Adib NSDI 15 [57]	FMCW	VICON room	a single FMCW radio with multiple antennas	median error: 11.7 cm in each of the x/y dimensions (≤ 5 people)

*_ denotes unspecified.

Challenges: In indoor user localization & tracking systems, wireless signals will distort due to the complex indoor environments and dynamic changes. Moreover, collecting a comprehensive fingerprint database for the entire indoor area can be labor-intensive and time-consuming, and keeping this database updated with changing indoor conditions adds further complexity. Another issue is to distinguish signals from users in multi-user scenarios, in which it is challenging to accurately localize and track each user's movement.

D. Person Counting

The person counting (i.e., counting the number of persons) in public places (e.g., shopping centers and airports) plays a significant role in pervasive applications, including human safety monitoring, traffic control, public area surveillance, user-oriented services, and marketing analysis. Traditional camera- or device-based methods have limitations. Specifically, camera-based methods may bring privacy concerns, and the performance will be poor in NLoS scenarios or low-light environments; device-based ones are costly and inconvenient, as they require users to physically carry Radio

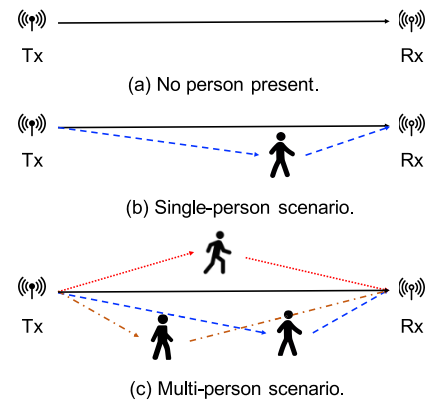


Fig. 13. Impact of persons on wireless signal transmission.

Frequency Identification (RFID) tags, smartphones, or other sensors. Wireless sensing is promising to achieve person counting due to the wide adoption of wireless infrastructure via either coarse-grained RSS or fine-grained CSI. Table VIII illustrates the comparison of such studies.

TABLE VIII
COMPARISON OF THE EXISTING WIRELESS PERSON COUNTING STUDIES

Reference	Modality	Activity	Implementation	Scenarios	Num of Persons	Accuracy
Nakatsuka: ICMU 08 [77]	RSS	moving, static	1 TX: AirStation AP WLA-2-G54C, 1 RX: a laptop	width of the space is 4.5 m	0-10	1.5 (average number difference)
Xu: IPSN 13 [55]	RSS	moving, static	Chipcon CC1100 radio transceiver	13 TX and 9 RX in the offices (150 m^2), 12 TX and 8 RX in the open floor (400 m^2)	1-4	84% (office), 86% (floor)
Yuan: JSEN 13 [78]	RSS	moving, static	16 TelosB nodes	room (18 $\times$ 18 m^2)	0-10+	94% (crowd static), 86% (crowd moving)
Xi: INFOCOM 14 [117]	CSI	moving	1 TX and 3 RX: each is a laptop equipped with Intel 5300 NIC	indoor, outdoor	0-15	>98% is less than 2 persons (indoor), >70% is less than 2 persons (outdoor)
Guo: SenSys 17 [162]	CSI	moving	a TX-RX pair: each is a laptop equipped with Intel 5300 NIC	activities room (16 m $\times$ 10 m), meeting room (13 m $\times$ 7 m), classroom (12 m $\times$ 5 m)	0-10	>90%
Choi: ACCESS 22 [187]	CSI	moving, static	4 pairs of ESP32 nodes	meeting room (5.5 m $\times$ 5.5 m), meeting room (11 m $\times$ 5.5 m)	0-10	MAE: 0.35 (5 people in a small-sized room), 0.41 (10 people in a medium-sized room)
Guo: TVT 22 [188]	CSI	walking	TX: TP-Link WDR7500 router with 1 antenna, RX: mini terminal equipped with Intel 5300 NIC and 3 antennas	laboratory (6 m^2), meeting room (5 m^2)	0-10	91.27% (laboratory), 88.76% (meeting room)
Khan: JIOT 23 [189]	CSI	moving, static	each is a device equipped with Intel 5300 NIC	research office (6 m^2), conference room (5 m^2)	0-7	96% (research office), 93.4% (classroom)
Ren: JIOT 23 [190]	FMCW	walking	TI IWR6843ISK radar and DCA1000EVM	outdoor open area	1-4	average probability of true positive: 98%

1) *RSS-Based*: It is observed that the RSS measurements tend to have very little variance in an environment where there is no person present between the transmitter (TX) and the receiver (RX), as shown in Figure 13 (a). However, the RSS value would have large variations when there is a person between TX and RX. Meanwhile, different numbers of persons lying between TX and RX would generate varying impacts on the TX-RX wireless channel [55], [77], [78], leading to distinctive RSS changes. Figures 13 (b) and 13 (c) illustrate examples of the corresponding single-person and multi-person scenarios, respectively. Generally, the relationship between the number of persons and the RSS average/variance can be modeled with a linear approximation formula [77], which can be used to estimate crowd density. Also, the study [55] verifies the feasibility of using the profiling RSS data collected with only a single subject present to count multiple subjects in the same environment with no extra hardware or data collection. To obtain precise estimation results, [78] presents a novel RSS-based Wireless Sensor Network (WSN) application using RSS to estimate different densities in different subareas by the K-means algorithm. These RSS-based approaches, however, require the extensive deployment of sensor nodes and pre-building of the RSS fingerprint database, causing extremely high costs and training efforts.

2) *CSI-Based*: CSI is highly sensitive to environmental variation, and its variation is able to indicate the change in the number of persons in the monitored area. Particularly, the study [117] identifies a stable monotonic function to theoretically characterize the relationship between the number

of moving people and the CSI variation, and proposes a metric PEM (Percentage of nonzero Elements) in the dilated CSI matrix to indicate the crowd size. Another study [162] develops a semi-supervised learning approach using the non-linear regression method to accurately estimate the number of participants leveraging CSI readings in indoor venues, which can significantly reduce training efforts. It also proposes a human density estimation approach based on the analysis of the CSI variance histogram across different WiFi subcarriers.

Challenges: Person counting systems are capable of estimating the number of persons in the monitored area. However, when there are multiple users overlapped by each other, these systems fail to accurately differentiate users. Also, it is not easy to deal with diverse environmental and user conditions.

E. Gesture Tracking

Gesture tracking is an active and important topic due to its wide application in a variety of areas, such as healthcare, elder care, and smart homes. In general, as shown in Figure 14, we divide gestures discussed in existing studies into three categories: (1) macro-movement, which are gestures involving whole body movement, such as walking, running, sitting down, standing up, picking up, cooking, and opening a door; (2) micro-movement, which are small gestures involve a portion of a body, such as hand/finger/mouth movements; (3) sudden-movement, which happens accidentally, such as falling.

Traditional gesture recognition approaches often utilize audio, vision, or wearable sensors. Such methods may

TABLE IX
COMPARISON OF MACRO-MOVEMENT GESTURE TRACKING STUDIES

Reference	Modality	Implementation	Scenarios	Gesture	Accuracy
Sigg: MoMM 13 [160]	RSS	RSSI-based system based on INGA sensor nodes	seminar room, corridor	5 activities	86.4%
Wang: MobiCom 14 [9]	CSI	TX: a laptop with Linksys E2500, RX: a laptop equipped with Intel WiFi Link 5300 card	apartments	11 in-place activities, 9 walking activities	> 96% TPR with < 1% FPR
Wang: MobiCom 15 [139]	CSI	TX: NETGEAR JR6100 and TP-Link TL-WDR7500, RX: a laptop equipped with Intel 5300 NIC	lab, apartment	9 activities	>96%
Wei: IPSN 15 [191]	CSI	a pair of WASP nodes	apartment	8 activities	97%
Wu: UBICOMP 16 [129]	CSI	1 TX: WiFi AP, 3 RX: Gigabyte Bxi3H-5010 Brix mini-PCs with Intel 5300 wireless NICs and external omnidirectional antennas	empty student room, activity room, office room	8 walking directions	median: < 10°
Fang: MobiSys 16 [125]	CSI	TX: Espressif ESP8266, RX: HMB Wearable Unit (RX)	lab, outdoor	5 activities	72.3% (controlled environment); 60.2% (real world)
Gao: TVT 17 [147]	CSI	a TX-RX pair: each is a computer equipped with Intel 5300 NIC	laboratories	8 activities	>92.3%
Wang: JSAC 17 [140]	CSI	TX: NETGEAR JR6100 and TP-Link TL-WDR7500, RX: a laptop equipped with Intel 5300 NIC	lobby, office, apartment	9 activities	96%

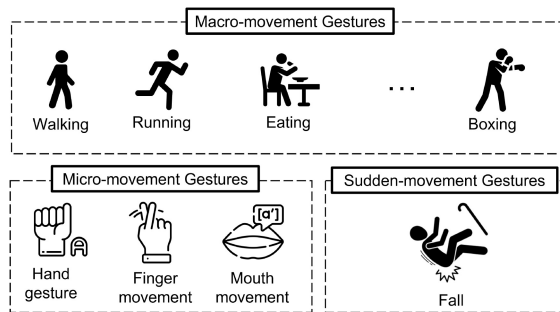


Fig. 14. Human gesture tracking using wireless signal.

have some limitations in practice. For example, audio-based approaches are vulnerable to ambient acoustic noise and they are thus not desirable to recognize micro-movement gestures; vision-based methods require line-of-sight (LOS) scenarios with good lighting conditions; wearable sensor-based strategies need users to physically wear these sensors, resulting in inconvenience and extra costs. Wireless-based techniques, however, are able to address the above challenges and achieve high accuracy in tracking various types of human gestures. Tables IX, X, and XI present the corresponding comparison of these wireless gesture tracking studies.

1) *RSS-Based*: Human gestures may distort surrounding wireless signals, causing distinguishable fluctuation patterns on RSS measurements.

Macro-movement: Each whole-body gesture has its particular way of being performed by a human, inducing a unique impact on environmental wireless signals. Such an impact would be reflected by RSS readings captured by nearby wireless receivers. For example, [160] proposes a device-free gesture recognition system by leveraging RSS fluctuations caused by human macro-movements. Specifically, it first extracts 17 empirical features (e.g., highest signal peak and median signal strength) from the captured RSSs. With such features, it then develops a *k*NN classifier to recognize

four regular macro-movement gestures (i.e., lying, standing, walking, and crawling) with an average accuracy of 86.4%.

Micro-movement: RSS can also be utilized to identify different micro-movement gestures. For example, the study [114] exploits the RSS fluctuation to identify 11 gestures, while the recognition accuracy is as low as 51%. To improve the accuracy, [115] leverages a directional antenna to recognize action-oriented gestures from the standard American Sign Language (ASL). It can then classify up to 25 fine-grain gestures with an average accuracy of 92% in the office scenario and 84% in the car scenario. Another study [11] further improves the recognition accuracy of 7 hand gestures to 96%.

2) *CSI-Based*: Due to the low resolution and limited sensing capability of RSS measurements, it is often difficult to achieve fine-grained activity recognition. Instead, recent studies propose to exploit CSI measurements for better gesture recognition performance.

Macro-movement: By matching CSI patterns against activity profiles, the study [9] distinguishes a set of in-place (e.g., cooking, washing dishes, bathing, studying, eating, and sleeping) and walking activities. It achieves an average detection rate of over 96% and an average false positive rate of less than 1%. Also, based on theoretical models of CSI dynamics and human activities, the studies [139], [140] build a human activity recognition and monitoring system, which can achieve an average accuracy of greater than 96% to classify a large set of macro-movement gestures.

Micro-movement: With fine-grained CSI information collected from COTS devices, [126] is the first to achieve micro-movement gesture recognition. It can classify 4 micro-movement gestures with an average accuracy of 92% in the LoS scenario and 88% in the NLoS scenario. For finger movement, [86] recognizes 8 typical finger gestures (i.e., zoom in/out, circle left/right, swipe left/right, and flip up/down) with over 93% recognition accuracy. In comparison, another study [82] achieves up to 90.4% average accuracy for recognizing 9 digits finger-grained gestures from the standard

TABLE X
COMPARISON OF MICRO-MOVEMENT GESTURE TRACKING STUDIES

Reference	Modality	Implementation	Scenarios	Gesture	Accuracy
Sigg: PerCom 14 [114]	RSS	TX: standard AP, RX: modified firmware on a Nexus One phone running Cyanogen mod 7.2	office, lecture room	11 gestures	51% (11 gestures), 72% (4 more disparate ones)
Melgarejo: UBICOMP 14 [115]	RSS	a TX-RX pair: each is WARP v3 board connected to a RE14P directional patch antenna	office, car	25 wheelchair-based gestures	92% (office), 84% (car)
Abdelnasser: INFOCOM 15 [11]	RSS	TX: 2 Cisco Linksys X2000 APs (apartment) or 3 Netgear N300 APs (floor), RX: a laptop	apartment, floor	7 hand gestures	87.5% (1 AP), 96% (3 AP)
He: ICCCN 15 [126]	CSI	TX: TP-Link TL-WR882N router with 3 antennas, RX: a laptop equipped with Intel 5300 NIC and 2 antennas	laboratories	4 gestures	92% (LoS), 88% (NLoS)
Tan: MobiCom 15 [86]	CSI	TX: a laptop connected with LINKSYS E2500 N600 Wireless Router, RX: a laptop equipped with an Intel WiFi Link 5300	office, home	8 finger gestures	>93%
Li: UBICOMP 16 [82]	CSI	TX: TP-Link TL-WDR4300 wireless router, RX: a desktop equipped with Intel 5300 NIC and 3 omnidirectional antennas	laboratory, dormitory	9 finger-grained gestures	90.4% (9 digits finger-grained gestures), 82.67% (continuous number text input)
Wang: TMC 16 [6]	CSI	TX: TP-Link TL-WDR4300 wireless router, RX: a desktop equipped with Intel 5300 NIC and 3 antennas; 4×4 MIMO based on 4 USRP N210 boards and XCVR2450 daughterboards	office	tongue, lips, and jaws motion related to 14 syllables and 33 words	91% (1 user), 74% (< 3 users)
Virmani: MobiSys 17 [89]	CSI	TX: a laptop equipped with an Intel 5300 WiFi NIC and 2 omnidirectional antennas, RX: TP-Link N750 with 3 antennas	room	6 gestures	91.4%
Ma: IMWUT 18 [166]	CSI	a TX-RX pair: each is a laptop equipped with Intel WiFi Link 5300	lab, home	276 sign gestures	94.81%
Tian: ACCESS 18 [134]	CSI	a TX-RX pair: each is a laptop equipped with Intel 5300 NIC	indoor	9 gestures	>96% (one-hand) gestures, 95% (two-hand) gestures
Meng: JIOT 22 [192]	CSI, AoA	1 TX: a commercial WiFi device, 1 RX: a laptop equipped with Intel 5300 NIC	meeting room	4 gestures	>96%
Adib: SIGCOMM 13 [37]	AoA	MIMO system consisting 2 TX and 1 RX: each is USRP N210 software radios with SBX daughter boards connected to LP0965 directional antennas	conference room	2 gestures	100% (≤ 5 m), 93.75% (6-7 m), 75% (8m)
Sun: MobiCom 15 [15]	AoA	TX: 802.11 devices, RX: a laptop equipped with Atheros 9590 chipset and 3 antennas	apartment, office, cafeteria	hand trajectory of drawing letters, words, and sentences	91%
Pu: MobiCom 13 [8]	Doppler	TX: USRP-N210 equipped with a XCVR2450 daughterboard with 1 antenna, RX: multiple USRP-N210s with 5 antennas	office, two-bedroom apartment	9 gestures	94%
Wang: THMS 21 [193]	Doppler	2 K-LC2 short range radars with the ST-100 radar development board	bedroom, living room, conference room	12 gestures	93.5%
Lien: TOG 16 [58]	FMCW	designed Soli FMCW and impulse radars	*—	4 gestures	92.1%
Palipana: IMWUT 21 [194]	FMCW	TI IWR1443	open, office, restaurant, factory, through-wall	9 gestures (easy), 12 gestures (complex)	96.6% (easy), 95.1% (complex)

*— denotes unspecified.

American Sign Language (ASL). For mouth movement, [6] enables Wi-Fi signals to “hear” a user’s talks by employing the CSI changes caused by the motion related to the tongue, lips, and jaws. It can achieve an average accuracy of 91% for a single individual and 74% for three people talking simultaneously. The above systems only work effectively for stationary users, whose position or orientation in the given environment does not change significantly during the training and test phases. However, the user configuration (i.e., the position or orientation information) may affect the patterns of CSI changes. To address this problem, [89] proposes a novel translation function that can generate virtual samples

of a given gesture using a real sample of that gesture collected from the user. It thus requires the user to provide training samples in only one configuration and can recognize 6 gestures irrespective of the user’s configuration. It also shows that when a user’s configuration at testing is not the same as that at training, the gesture recognition accuracy can reach 91.4%.

Sudden-movement: Existing studies propose wireless-based real-time, non-intrusive, and low-cost fall detection systems [65], [110], [137], [195]. Among them, [195] is the first to use both the phase and amplitude of CSI to detect falls in indoor environments with a fall detection precision of

TABLE XI
COMPARISON OF SUDDEN-MOVEMENT GESTURE TRACKING STUDIES

Reference	Modality	Implementation	Scenarios	Accuracy
Zhang: ICOST 15 [195]	CSI	TX: TP-Link WDR5300 Router with 1 antenna, RX: a laptop with 2 internal antennas equipped with an Intel WiFi Link 5300 card	office, meeting room	TPR 89% with FPR 13%
Wang: TMC 17 [65]	CSI	TX: TP-LINK TL-WDR7500 wireless router, RX: a desktop equipped with Intel 5300 NIC	chamber, laboratory, dormitory	TPR 94% with FPR 13%
Wang: TMC 17 [137]	CSI	1 TX: TP-Link WDR5300 Router with one antenna, one or more RX: laptop/mini-pc equipped with an Intel WiFi Link 5300 card	office, apartment, meeting room, hall	sensitivity 91% and specificity 92%
Palipana: IMWUT 17 [110]	CSI	a TX-RX pair: each is a laptop equipped with an Intel 5300 card and 2 external omnidirectional antennas	apartment, laboratory, bathroom, toilet	93% (pre-trained), 80% (environmental change)
Yang: TMC 22 [196]	CSI	1 TX: router equipped with an Intel 5300 card and 3 omnidirectional antennas, 1 RX: device equipped with an Intel 5300 card and 3 omnidirectional antennas (only 1 used)	bathroom, bedroom, corridor, kitchen, laboratory	false alarm rate 5.7%, missed alarm rate 3.4%
Chen: TMC 23 [197]	AoA	1 TX: a laptop equipped with Intel 5300 NIC and 1 antenna, 2 RX: each is a laptop equipped with Intel 5300 NIC and 3 antennas	laboratory, office, classroom, kitchen, bathroom	84.31%
Li: ICCASP 22 [198]	FMCW	TI AWR1843	indoor	accuracy 98.8%, precision 93.6%

89% and a false alarm rate of 13% on average. Later, another work [65] provides the theoretical analysis to prove that falling can be distinguished using wireless signals from other human motions (i.e., resting, sitting down, walking, and standing up), and it achieves a 94% TPR (True Positive Rate) with a 13% FNR (False Negative Rate).

To further enhance the performance, the study [137] finds that the difference in CSI phase between two antennas is more sensitive to fall than CSI amplitude, and also the sharp power profile decline pattern of the fall in the time-frequency domain can be used for accurate fall detection. It then achieves 91% of sensitivity and 92% of specificity. Similarly, another study [110] utilizes the conventional Short-Time Fourier Transform (STFT) to extract time-frequency features. It then uses a sequential forward selection algorithm to single out features that are resilient to environmental changes while maintaining a higher fall detection rate, achieving an average fall detection accuracy close to 80% when the environment changes while 93% for a pre-trained system.

3) *AoA-Based*: The study [37] presents Wi-Vi, a wireless technology that uses AoA to track motion. It is the first to implement a gesture recognition system through walls and does not require the human to carry any wireless device. Wi-Vi can detect simple body gestures with 100% accuracy at all distances less than or equal to 5 m. Instead of simple gesture recognition, another AoA-based study [15] can continuously track the hand's trajectory for in-air handwriting of letters, words, and sentences with an average letter and word recognition accuracy of 95% and 91%, respectively.

4) *Doppler-Based*: Different body movements may result in different positive and negative Doppler shift patterns. The unique Doppler shifts profile can be thus used for gesture tracking. For example, the study [8] extracts Doppler shifts in LOS, NLOS, and through-the-wall scenarios to detect and

classify a set of nine gestures with an average accuracy of as high as 94%.

5) *FMCW-Based*: FMCW radar has also been applied to gesture recognition. Since FMCW radar takes up to 1.79 GHz bandwidth instead of the 20 MHz bandwidth of WiFi devices, it can achieve higher time resolution and better performance. For example, the study [58] is the first end-to-end FMCW radar sensing system specifically designed for tracking and recognizing fine hand gestures. It can achieve 92.1% accuracy over the 1000 test gestures.

Challenges: Accurate recognition and tracking of gestures present significant challenges. One such challenge is that the signal variation induced by micro-movement tends to be much smaller compared to marco-movement and sudden-movement. To address it, existing systems often place the transmitter and receiver in close proximity to the target, which limits the small-scale scenario. Alternatively, directional antennas and beamforming techniques can be exploited to remove the noise and interface, also reducing the sensing area.

F. Vital Signs Monitoring

Vital signs, including breathing rate/volume and heart rate, may disclose a person's essential body function, e.g., sleep quality, stress level, and health condition. Traditional methods to monitor vital signs need special equipment (e.g., ambulatory cardiorespiratory monitor [199]) and require a patient to wear dedicated sensors, which are costly and intrusive. For example, polysomnography is often used to diagnose various sleep disorders [200], and the average cost of overnight polysomnography is around \$1,000 to \$2,000 per night [201]. Other portable devices (e.g., Masimo's MightySat fingertip pulse oximeter [202]) are less expensive, but they are also intrusive as they require users to wear them and may make people feel uncomfortable. Additionally, some systems (e.g., [203]) utilize

TABLE XII
COMPARISON OF EXISTING WIRELESS VITAL SIGNS MONITORING STUDIES

Reference	Modality	Implementation	Scenarios	Activity	Accuracy
Patwari: TMC 13 [81]	RSS	a network of 20 MEMSIC TelosB wireless sensors	clinical room	lying on the bed	RMSE: 0.07-0.42 bpm
Kaltiokallio: IPSN 14 [79]	RSS	a TX-RX pair: each node equipped with Texas Instruments CC2431)	bedroom	lying on the bed	MAE: 0.12 bpm
Abdelnasser: MobiHoc 15 [56]	RSS	TX: Cisco Linksys X2000, RX: Samsung Galaxy S4 Mini mobile/Samsung Galaxy Note II devices/HP EliteBook laptop	apartment, floor (through-wall scenario)	holding the device, standing	<1 bpm for 1 user, 0.9 bpm for 3 users, >96% for apnea detection
Liu: TMC 15 [112]	CSI	3 TX-RX pairs: TP-LINK WR740 wireless routers (TX), desktop equipped with Intel 5300 NIC (RX)	bedroom	lying on the bed	>85% for 6 sleep postures, >80% for change sleep postures detection, 82.1% for apnea detection
Wang: UBICOMP 16 [142]	CSI	TX: TP-Link WDR5300 with one antenna, RX: mini-PC equipped with an Intel WiFi Link 5300 card and one antenna	office room	sitting in a chair, lying on a bed	Fresnel model is verified and the theoretical foundation is provided
Wang: TIST 17 [204]	CSI	TX: a desktop equipped with Intel 5300 NIC, RX: a laptop equipped with Intel 5300 NIC	laboratory, through-wall scenario, corridor	standing, sitting, sleeping	<0.5 bpm (96% estimation) for 1 user, <0.5 bpm (93% estimation) for 2/3 users
Liu: JIOT 18 [67]	CSI	TX: a laptop connected to TP-Link TL-WDR4300, RX: a laptop equipped with Intel 5300 NIC	laboratory, apartment	lying on a bed	<0.5 bpm (80% estimation) for 1 user, <1 bpm (90% estimation) for 2 users, >98% for 4 sleep postures detection, <4 bpm (90% estimation) for heartbeat estimation
Zhang: THMS 23 [205]	CSI	a TX-RX pair: each is a mini PC with Intel Link 5300 NIC	room	lying on a bed	96.887% for respiration and 94.708% for heartbeat estimation
Guo: TOSN 23 [206]	CSI	a TX-RX pair: each equipped with 3 omnidirectional antennas	laboratory	sitting in a chair, standing	0.1 bpm for respiration estimation
Gu: TIM 09 [100]	Doppler	instrument-based radar transceiver consisting of Agilent spectrum analyzer E4407B, Agilent vector signal generator E8267C, and Agilent vector signal analyzer 89600S	laboratory	sitting in a chair	100% (at the optimal detection point), <50% (at the null detection point)
Nguyen: INFOCOM 16 [101]	Doppler	radio transceiver (developed from a iMotion radar [208]), a radar navigator	hospital	lying on a bed	90%-95.4% for breathing volume estimation
Zhao: TIM 17 [102]	Doppler	designed a heterodyne CW Doppler radar with digital-IF receiver	laboratory	sitting in a chair	100% for respiration estimation, >90.2% for heartbeat estimation
Juan: TMTT 23 [208]	Doppler	designed MIMO CW radar	room	sitting in a chair	<1.2 bpm for respiration rate estimation, <3 bpm for heartbeat estimation
Adib: CHI 15 [3]	FMCW	FMCW radar [29]	office	sitting in a chair	median: 99.3% for respiration and 98.5% for heartbeat estimation (1 user), ≥97.3% for respiration and ≥98.7% for heartbeat estimation (3 users)
Li: TIM 22 [209]	FMCW	MMWCAS-RF-EVM 77GHz mmWave FMCW radar	corridor	sitting in a chair, standing	average error 1% for respiration

cameras to track the movements of the user's chest for breathing rate estimation. Such camera-based approaches, however, can be negatively affected by a low-light environment, and meanwhile, raise user privacy concerns.

Vital signs inference via wireless signals (e.g., [67], [81], [112]) has drawn increasing attention because of its low-cost, non-invasive, and easy-to-deploy properties. Table XII summarizes and compares existing such studies.

1) *RSS-Based*: Wireless channels are highly sensitive to tiny body movements caused by breathing and heartbeats. Consequently, the corresponding RSS measurements fluctuate over time with these movements and thus can be leveraged to infer vital signs. For example, the study in [81] shows that breathing induces sinusoidal variation in the measured RSS on a link. With RSS measurements from 20 links simultaneously, it can achieve breathing rate estimation with root-mean-square error (RMSE) between 0.07-0.4 breaths per minute (bpm) by applying the maximum likelihood estimation

(MLE). Another study [79] later utilizes just a single TX-RX pair to measure RSS to estimate a user's breathing rate with a mean absolute error (MAE) of 0.12 bpm. Also, [56] can estimate different breathing rates within 1 bpm error and detect abnormal breathing situations (e.g., apnea) with more than 96% accuracy.

2) *CSI-Based*: As aforementioned, RSS represents coarse channel information while CSI represents fine-grained channel information, consisting of subcarrier-level information. As a result, CSI is more sensitive to detecting breathing activity and the CSI-based approaches can capture breathing from a distance [4], [67], [103], [112], [142], [156], [159], [204], [210], [211]. Off-the-shelf WiFi devices can continuously collect CSI around a person to infer the person's breathing rate in different sleeping positions [112]. Also, the study [142] is the first to introduce the Fresnel zone model for respiration rate detection, and provides the general theoretical foundation to explore the feasibility of breathing rate detection by correlating

TABLE XIII
COMPARISON OF WIRELESS KEYSTROKE INFERENCE TECHNIQUES

Reference	Modality	Implementation	Scenarios	Type	Accuracy
Oligeri: WiSec 20 [214]	RSS	SDR equipped with VERT2450 (omnidirectional)/Aaronia HyperLOG 60350 (directional antenna)	proximity, behind-the-wall, remote	keyboard (A-Z)	>70% (harsh conditions), >90% (normal conditions)
Ali: MobiCom 15 [16]	CSI	TX: TP-Link TLWR1043ND WiFi router, RX: a laptop with Intel Link 5300 WiFi NIC	*—	keyboard (A-Z, 0-9)	97.5% detection rate, 96.4% for recognizing single keys, 93.5% for recognizing continuously typed sentence
Li: CCS 16 [19]	CSI	laptop equipped with Intel 5300 NIC with 1 directional antenna and 2 omnidirectional antennas	*—	keyboard (0-9), PIN pad (0-9)	1-digit (top 3 candidates): 89.0% (Samsung) and 95.0% (Xiaomi), 6-digit (top 100 candidates): 80%
Li: ACCESS 17 [215]	CSI	TX: a smartphone equipped with 2 antennas, RX: a laptop equipped with Intel 5300 NIC with 3 antennas	*—	PIN pad (0-9)	> 83% (key recognition)
Fang: CCS 18 [17]	CSI	USRP X300 with 40 MHz bandwidth CBX daughterboards	*—	keyboard (A-Z)	95.3% (word recognition)
Yang: CCS 22 [21]	CSI	USRP X300 with LP0965 (directional)	private office, public cafeteria	PIN pad (0-9)	> 52% of 6-digit PINs with less than 100 attempts

*— denotes unspecified.

CSI and one's breathing depth, location, and orientation. Based on the observation that the variation of the CSI phase is more robust than amplitude to various environmental factors (e.g., the link distance and nearby obstacles), TensorBeat [204] leverages a tensor decomposition method and CSI phase variations to monitor breathing rate. Another study [103] empirically reveals CSI provides more robust estimates of breathing rate compared with RSS.

3) *Doppler-Based*: Doppler radar systems have been proposed to achieve breathing detection [31], [32], [33], [34], [35], [100], [101]. For example, the study [101] exploits 2.4 GHz Doppler radar to capture the breath volume based on a phase-motion demodulation algorithm, achieving a median accuracy from 90% to 95.4%. Another work [102] later utilizes the synchrosqueezing transform (SST) based signal processing method to extract high-resolution instantaneous vital sign rates, and obtains accuracy of 100% for respiration estimation and over 90.2% for heartbeat estimation. According to the Doppler theory, a target with time-varying movement but zero net velocity will reflect the signal, whose phase is modulated in proportion to the displacement of the target [98]. A stationary person's chest and stomach can be thus regarded as a target. However, such Doppler radar based techniques suffer from the null point problem, which significantly degrades the measurement accuracy [34], [212], [213].

4) *FMCW-Based*: Doppler-based methods do not have a good way to eliminate the influence of moving objects in the front or behind the target, while an FMCW radar can separate the radio signal reflections from different objects and measure the breathing rate and heartbeat with high accuracy [3], [25], [26]. The breathing-induced body movement changes the signal reflection time. By analyzing such changes, the breathing rate can be extracted. For instance, Vital-Radio [3] can track the vital signs of multiple users simultaneously, and achieves median accuracy in measuring breathing and heart rates of 99.3% and 98.5% respectively for one single user. For three users, breathing and heart rate monitoring accuracy is around 98%. However, high resolution (i.e., the minimum measurable change) requires a large sweep

bandwidth B as the resolution equals $\frac{C}{2B}$ [29], where C is the speed of light.

Challenges: Many existing systems require users to be stationary for accurate measurements, which limits the utility in real-world applications where users often move. Also, the performance is closely related to the user's location and body orientation. Moreover, privacy and security concerns arise when collecting and processing sensitive health information.

G. Keystroke Recognition

People inevitably type sensitive information, such as personal identification numbers (PINs) and private messages, into computer systems via a keyboard for authentication, storage, or transmission under many practical and sometimes public scenarios. In recent years, many studies have shown the success of leveraging wireless signals to infer keystrokes (e.g., [16], [17], [19], [21], [216]). Compared with other side-channel keystroke inference attacks (such as video-based [217], [218] or vibration-based [219], [220]), wireless-based techniques have three advantages: (1) wireless signals are ubiquitous and invisible, causing wireless-based keystroke inference techniques easy to set up and less likely to arouse suspicion; (2) they are non-invasive as there is no need to pre-install malware on the victim's device; (3) unlike video-based or sensor-based attacks, they do not require the victim to be in line-of-sight or close proximity of the wireless transceivers. Table XIII compares some technical parameters and the performance of existing wireless keystroke inference studies.

1) *RSS-Based*: BrokenStrokes [214] detects the presence of specific keywords in long keystroke sequences by only eavesdropping on the keyboard-dongle communication links. In detail, it converts the received signal strength peaks to inter-keystroke timings, and then to keywords, achieving over 70% accuracy in the harshest conditions and over 90% accuracy in normal operating conditions.

2) *CSI-Based*: The study [16] uses fine-grained CSI to perform keystroke inference, with 96.4% recognition accuracy for classifying single keys. Another study [19] utilizes CSI to identify a 6-digit password with a success rate of 80%.

Also, the work [215] uses acceleration and microphone sensors of smartphones to determine the keystroke period in which CSI is collected for keystroke inference. It can achieve a key recognition accuracy above 83%. Most of those wireless-based keystroke inference techniques (e.g., [16], [19], [216], [221]) still require a training process to pre-label the observed wireless signal sample with the corresponding keystroke. Reference [17] removes the training process by exploring the context correlation which is strictly constrained by the spelling and grammar of the English language, achieving high accuracy of 95% for an input of 150 words. However, it cannot be used for inferring numbers, in which digits are usually randomly combined. A recent study [21] requires neither labor-intensive training nor contextual information, and identifies a new type of attack that can compromise numerical keystrokes by exploiting spatiotemporal features of keystroke-disturbed wireless signals.

Challenges: Keystroke recognition is further complicated by the presence of ambient noise and interference from other environmental objects or individuals, as well as the diversity in typing styles among different individuals. Furthermore, practical implementation of keystroke recognition systems often requires high sampling rates, which may not be supported by all wireless devices or may lead to high power consumption.

H. Integration Into Real-World Systems

Numerous practical implementations of HPI can be found across various domains. We introduce several associated real-world systems and platforms.

Biometric Identification Systems: These systems utilize HPI to uniquely identify individuals, including fingerprint scanner [222], iris camera [223], facial recognition [224], or voice recognition [225], [226]. They analyze unique biological or behavioral attributes to grant authorized access to the system. For example, Apple Face ID [224] employs the TrueDepth camera system to accurately map the geometry of a user's face for authentication. Similarly, Samsung [223] leverages iris recognition based on the unique shape and pattern of the iris to identify users.

Security Surveillance Systems: These systems incorporate multiple sensors such as motion sensors, image sensors, or microphones, for intruder detection. They analyze human behavior to recognize unauthorized individuals, trigger alerts or alarms for prompt responses, and identify various suspicious activities [227], [228], [229], [230], [231], [232].

Indoor Positioning and Navigation Solutions: Companies like IndoorAtlas [233] and Estimote [234] provide indoor positioning and navigation solutions for accurate indoor location tracking in various applications, including retail, healthcare, and transportation. IndoorAtlas [233] enhances positioning performance on top of existing WiFi or beacon deployment, while Estimote [234] focuses on proximity sensing and location tracking using Bluetooth Low Energy (BLE) beacons.

People Counting and Crowd Management Solutions: Retailers, event organizers, and transportation hubs commonly utilize people counting systems to estimate crowd density, manage queues, and optimize space utilization. Axis

Communications [235] and Hikvision [236] provide systems consisting of surveillance cameras and advanced algorithms to count individuals and track their movements in real time.

Gesture Recognition Systems: Gesture recognition technology is applied to a wide range of industries, including gaming, Virtual Reality (VR), Augmented Reality (AR), and smart IoT devices. Notably, Microsoft Kinect [237] utilizes depth sensors based on ToF to measure the distance of objects from the camera, enabling skeletal tracking in gaming applications. It allows users to control games and navigate menus using natural interactions. Also, Ultraleap 3Di [238] and Intel RealSense cameras [239] offer precise hand and finger tracking with high accuracy and low latency, interpreting gestures and translating them into commands or interactions with digital content.

Healthcare Monitoring Devices: Wearable devices, such as fitness trackers and smartwatches, offered by Fitbit [240], Apple [241], and Garmin [242], have gained popularity for enhancing health and well-being. These devices may use multiple sensors, including optical sensors for heart rate monitoring, accelerometers for motion tracking, and additional sensors like gyroscopes for sleep pattern recognition.

Smart Home Automation Systems: Companies like Nest [243], Ecobee [244], and SmartThings [245], provide smart home automation systems based on HPI to enhance comfort, energy efficiency, and convenience. These systems integrate various sensors to make intelligent decisions based on user behavior and preferences. Motion sensors are used for occupancy detection, while RSS is employed to determine the presence of users in different rooms. With this information, smart thermostats are capable of optimizing temperature settings and energy consumption, reducing unnecessary heating or cooling when the room is unoccupied.

I. Communication Efficiency Enhancement

Furthermore, the ability of HPI inference provides valuable data for system control and decision-making, leading to enhanced communication efficiency. With HPI inference, a system can make context-aware decisions that are tailored to the user's specific needs in the following aspects:

- *Optimized Power Consumption:* HPI can help turn on/off the power-saving modes in communication devices. For example, a smart thermostat could adaptively adjust room temperature based on crowd density and human activity [246]. Similarly, smart systems can optimize their functionality to minimize energy consumption and maximize performance during peak usage periods [247].
- *Latency Reduction:* Real-time decisions and data transmissions can be optimized based on the inferred human presence. For example, a study [248] utilizes user position for data pre-buffering and pushing information from the server database to the user's database to ensure minimal latency when the user interacts with the device.
- *Dynamic Bandwidth Allocation:* HPI can enable dynamic bandwidth allocation to improve communication efficiency. For example, when an HPI inference system detects a conference room filling up with people, it can

anticipate increased data usage and allocate additional bandwidth [249] to accommodate the surge in demand.

- *User Experience Enhancement:* Knowing users' location and timing of their presence enables networks to proactively optimize settings. For instance, recent research [250] incorporates multiple decision parameters including RSS, network load (NL), and user service rate requirements for access selection. This approach ensures the provision of high-quality and uninterrupted service, further enhancing the overall user experience.

Additionally, a smart health monitoring system can promptly alert healthcare providers in real time based on detected abnormal vital signs. HPI inference plays a significant role in enhancing communication efficiency.

In conclusion, HPI applications using wireless sensing technologies have shown significant promise. They provide a non-intrusive and ubiquitous approach to monitoring human activities and behaviors, enabling a wide range of applications such as user identification, intrusion detection, indoor user localization and tracking, person counting, gesture recognition, vital signs monitoring, and keystroke recognition. Importantly, these HPI applications are not just theoretical concepts but have practical implementation in our daily lives, thereby further enhancing communication efficiency.

V. FUTURE RESEARCH TRENDS

To employ existing wireless HPI inference techniques, we often need to first deploy the wireless environment, i.e., setting up wireless transceivers, either commercial off-the-shelf WiFi devices or software-defined radio (SDR) systems, around the target user. As IoT with wireless connectivity is more and more pervasive and deployed in critical applications, they may carry sensitive HPI that should be protected. Also, the advance of machine learning techniques complements traditional signal processing methods and brings new opportunities to wireless HPI inference, as machine learning based techniques are able to learn from and adapt to the environment via experience [251]. Besides, mmWave communication technology is increasingly adopted for emerging wireless applications such as virtual reality (VR) [252]. Compared with WiFi signals, mmWave radar signals have shorter wavelengths, which may be exploited to detect subtle motion. With those new changes, this section discusses the future trends for wireless HPI inference, accordingly.

A. Challenges With IoT Devices

IoT is a system of interrelated computing devices connected to a network and/or to one another, exchanging data without necessarily requiring human-to-machine interaction [253]. Due to their low-cost and low-power characteristics, IoT devices have been extensively deployed in various domains, including smart homes, transportation, health care, and manufacturing, as shown in Figure 15. According to a recent study about the global IoT market, it is expected that by 2025, there will be approximately 27 billion connected IoT devices worldwide [254]. However, the proliferation of IoT devices brings several key challenges:

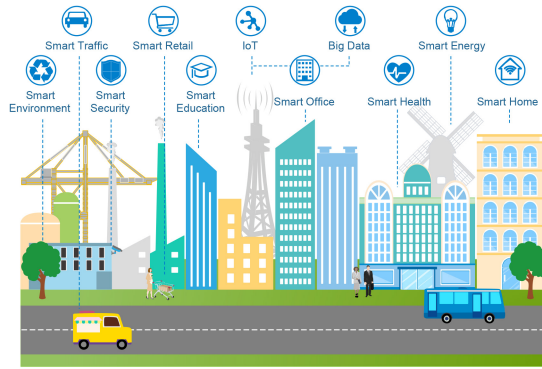


Fig. 15. Extensive applications of IoT.

Challenge 1: Ensuring the security and privacy of sensed data transmitted wirelessly. IoT devices may monitor users' private activities, and the data they collect often carry a great potential for privacy risks regarding the use of the data and its access [255], [256], [257]. Intuitively, the behaviors of various smart devices, such as smart door locks, lighting control systems, and wireless security cameras can be easily affected by human activity. For instance, in a smart home, the lighting condition can be adjusted automatically according to whether the user enters or walks out of the room, and the door will open if the identity is verified for the user who wants to enter the room [258]. As wireless signals may carry important information about these devices' behavior, they can be captured and analyzed to infer HPI.

Challenge 2: Keeping multiple IoT devices to be well-calibrated and synchronized. A vast array of devices with their own set of capabilities are sourced from different manufacturers. Therefore, inter-device data exchange for IoT devices is challenging. First, IoT devices are often portable, and their high mobility may introduce noise and interference, decreasing communication performance. Second, synchronization among diverse IoT devices is crucial when data from multiple devices need to be combined or compared, especially for multi-modal systems that collect data from different sensors, such as capturing audio, video, temperature, motion, and more. For instance, if a surveillance system captures both video and audio [259], a sound should correspond accurately with the visual event causing it; otherwise, it may cause false alarms to make the security system unreliable.

Challenge 3: Addressing the heterogeneity among various IoT devices. On the one hand, different IoT devices may support incompatible communication standards (e.g., WiFi, Bluetooth, and ZigBee) and have varying sensing modalities. For example, a Zigbee smart bulb might struggle to relay its readings to a smart home hub that exclusively supports WiFi or Bluetooth. On the other hand, data quality may differ in different IoT devices. For example, a 4K security camera might capture 4K video while a normal one may only capture 1080p. Such disparities can lead to inconsistencies in data quality, which may deteriorate data aggregation or data fusion.

Challenge 4: Designing new authentication method in emerging IoT devices. Most emerging IoT devices lack a user interface (e.g., a touchscreen or keypad), and traditional

authentication methods using direct text entry become inapplicable. New secure and robust mechanisms are thus required to enable wireless communication among IoT devices [260]. For example, the study [261] develops a robust communicating system for a mobilizable IoT network. It exploits ultrasonic signals at a frequency corresponding to the target receiver, forcing the inertial sensors to resonate, so as to convey information. Also, to authenticate users of IoT devices, [262] presents a virtual sensing technique that allows IoT devices to virtually sense user touches on the devices.

Challenge 5: Achieving real-time sensing while improving energy efficiency. Furthermore, real-time processing and analysis of HPI within IoT devices have become essential requirements for various applications, especially for health monitoring, intruder detection, and fall detection systems. Responding promptly to the new arrival of data and analyzing it without delay is crucial in these contexts. This presents several challenges, including resource constraints of IoT devices, data quality, environmental noise, and the need for robust and adaptable models. Another critical aspect to consider is that devices continuously monitoring and transmitting data rapidly consume energy. To address these issues, future research should focus on developing lightweight machine learning models, efficient data pre-processing techniques, and adaptive learning mechanisms that can operate within IoT device constraints while ensuring accurate and real-time analysis.

B. ML for Wireless HPI Inference

Machine Learning (ML) plays an essential role in wireless HPI inference systems. ML approaches leverage wireless signals to sense our environment, detect and monitor our activities, and localize and track the users. For example, the work [139] proposes a Hidden Markov Model (HMM) for human activity recognition; the study [16] trains k -Nearest Neighbour (k NN) classifiers for recognizing keystrokes; the Support Vector Machine (SVM) model can be built to perfectly classify the gestures [127]. Also, Convolutional Neural Network (CNN) classifiers can be leveraged for sign language recognition [166], and another work [128] exploits a three-layer Deep Neural Network (DNN) for user authentication. These ML approaches are widely used in past studies, providing high-accuracy performance. However, the applications of ML in wireless HPI inference present a lot of challenges.

Challenge 1: Ensuring Reliability of ML Approaches in HPI Inference. Despite the extensive use of ML techniques in wireless HPI inference and their high-accuracy performance, ensuring consistent reliability across diverse scenarios remains a challenge. Different activities and environments introduce significant variations in signal patterns, which may cause a single ML model to underperform in certain situations. Ensemble learning methods offer a solution by combining predictions from multiple models, thereby capturing a wider spectrum of data patterns. For instance, WiARes [263] leverages ensemble learning, fusing predictions from a multiple layer perceptron (MLP), a random forest (RF), and a support vector machine (SVM) to enhance human activity recognition accuracy. Similarly, a recent study [264] presents an ensemble

approach for cross-person activity recognition, demonstrating increased reliability and more robust predictions compared to standalone models.

Challenge 2: Integrating machine learning with multi-modal sensing. Integrating machine learning techniques with wireless sensing technologies offers vast opportunities for a range of applications, with most of these methods focused on a single type of wireless measurement. However, an approach that combines different types of signals (e.g., acoustic, CSI, mmWave, infrared, ultrasound) can provide more comprehensive and diverse information for enhanced HPI inference [265], [266], [267]. For instance, [266] fuses ultrasound (which is immune to ambient noise and provides additional information about the speaker) with acoustic signals (which offer rich auditory data and are less susceptible to airflow) for speech enhancement. If machine learning networks can learn from the combination of these wireless measurements, it can result in more robust HPI inference techniques. Such techniques would be particularly beneficial in ubiquitous deployments, especially in future smart homes with many IoT devices.

Challenge 3: Achieving scalability and adaptability. Existing WiFi sensing techniques based on ML (deep learning) require a labor-intensive and time-consuming process of collecting training data or fingerprints. The training data need to be collected for each target subject or activity across diverse environments. While feasible for a subset of users and typical environments, it becomes impractical when expanding to new users or environments. These constraints limit the applicability of such techniques in larger, more complex settings. Therefore, there is an urgent need for innovative solutions that can reduce the extensive data collection requirement, enabling more scalable and flexible WiFi sensing applications. To address it, [268] applies transfer learning to effectively reuse knowledge across different sites and tasks. In addition, [269] utilizes domain adaptation, allowing the trained model to be applied to untrained domains (e.g., new cars, new drivers) for in-car activity recognition. Despite promising progress in these areas, several challenges, such as ensuring model robustness and reliability in new environments, still need to be addressed. Nonetheless, the potential of these techniques to significantly enhance the scalability and usability of deep learning models makes this an exciting area for future research.

Challenge 4: Mitigating security concerns in wireless ML. The usage of machine learning algorithms in the wireless domain also brings security concerns. Specifically, adversarial machine learning is receiving increasing attention nowadays, which can effectively disrupt wireless communications [270]. It studies vulnerabilities of machine learning approaches in adversarial settings and develops systems to make learning robust to adversarial manipulation [271]. An adversary can carefully design inputs, then feed them to machine learning models in the test or training phase to manipulate the behavior of a legitimate system by launching adversarial attacks [272]. For example, [273] trains a generative adversarial network (GAN) to spoof wireless signals. Hence, adversarial attacks and countermeasures should be considered when applying machine learning tools in achieving wireless HPI inference.

C. HPI Inference With mmWave

Millimeter wave (mmWave) communication has been witnessed as a promising technology for next-generation wireless systems. Millimeter wave frequencies range from 30 GHz to 300 GHz, which are much higher than those used by traditional wireless technologies (e.g., WiFi). As the wavelength of a signal is inversely proportional to its frequency, the wavelength at mmWave frequencies is much shorter than at lower frequencies. Thus, the size of the electronic components designed for transmitting and receiving these signals can be reduced [274], and it is possible to design smaller, more compact, and more portable mmWave-supported devices.

Nowadays, mmWave-supported devices are increasingly popular in everyday life. For example, 5G smartphones are equipped with mmWave technology, which allows them to connect to 5G networks and take advantage of the high speeds and low latency [275]; some wireless routers use mmWave technology to provide high-speed wireless Internet connectivity to devices at home or in an office [276]; autonomous driving systems consisting of mmWave radars provide high-resolution radar images for obstacle detection and avoidance [277]; mmWave frequencies are also used in medical imaging devices, such as CT scanners and MRI machines, to produce detailed images of the human body [278].

Besides, mmWave operates across a wide bandwidth, which results in greater sensing resolution. In detail, the resolution can be computed as $R = \frac{C}{2B}$, where C is the speed of light and B is the sweeping bandwidth. Thus, mmWave technology with a chirp bandwidth of a few GHz will have a range resolution in the order of centimeters (e.g., a chirp bandwidth of 4 GHz translates to a range resolution of 3.75 cm) [274].

The high resolution of mmWave enables it to sense minute human motion. Recent studies show that mmWave systems have improved performance compared with traditional wireless systems in terms of achieving various applications, such as user localization [279], vital signs monitoring [280], [281], activity recognition [116], [282], occupancy detection [283], user identification [284], [285], and speech acquisition [286], [287], [288]. Those studies provide the initial foray into HPI inference using mmWave, and we expect more such schemes will be designed targeting a broader category of HPI with the increased adoption of mmWave techniques.

Challenge 1: Extending the effective range of mmWave sensing. Indeed, millimeter-wave (mmWave) technology demonstrates significant potential for high-precision, non-intrusive HPI inference applications. However, there are inherent challenges that need to be addressed, including occlusion and signal attenuation. mmWave signals are highly susceptible to obstruction by obstacles and suffer from significant signal attenuation over long distances. It makes reliable HPI inference in diverse environments challenging. Because signal attenuation tends to increase with frequency, mmWave radar operating at a higher frequency may have a shorter effective range. To address this problem, an intuitive approach is to simply increase the transmitter power, but

this solution is not energy-efficient and may pose additional issues such as interference with other systems and potential health concerns. Therefore, it is an important direction for the future to investigate intelligent reflecting surfaces (IRS) [289] and reconfigurable intelligent surfaces (RIS) [290], which are employed in the communications domain for signal propagation and beam steering for a larger coverage area.

Challenge 2: Designing advanced integrated circuits and systems. High carrier frequencies and bandwidths introduce design challenges for mmWave communication circuit components and antennas. The high transmit power and large bandwidth can cause nonlinear distortion in power amplifiers. RF integrated circuits also face issues related with phase noise and IQ imbalance. On the other hand, the implementing mmWave technology requires high-frequency and high-speed components, demanding advanced system design and precise manufacturing techniques to produce these energy-efficient, compact, and cost-effective components [274].

Challenge 3: Adopting mmWave techniques in multimodal sensing. Another significant trend in the future is multimodal sensing. Combining mmWave sensing with other sensing modalities (e.g., acoustic, infrared) could enhance the accuracy and robustness of HPI inference. For example, the study [265] integrates mmWave and acoustic signals from a microphone, thereby facilitating a noise-resistant, long-distance speech recognition application. Similarly, the work [291] jointly analyzes mmWave and thermal camera signals, achieving privacy-preserving temperature screening and human tracking. These studies provide exciting opportunities for innovative interaction techniques, applications, and use cases.

In summary, wireless HPI inference is a promising field with significant challenges, including privacy and security concerns, robustness in various environments, and scalability for large-scale deployments. As for future trends, multimodal sensing is expected to gain prominence, as systems integrate different sensing modalities to gather comprehensive data for more accurate and detailed inference. The rise of edge computing and AI, coupled with the growth of the IoT ecosystem, paves the way for real-time data processing and broader integration of wireless HPI sensing. This expansion opens up potential applications in areas such as healthcare, retail, and smart homes. Considering these trends and challenges, wireless HPI sensing is a promising area for future research.

VI. CONCLUSION

In this survey, we review the existing research effort on wireless human profile information (HPI) inference. An overview of the HPI inference system is proposed, including data collection, data preprocessing, feature extraction, algorithm design, and applications. We systematically analyze different wireless sensing modalities, features, and techniques applied in varying phases of HPI inference. Also, we discuss the major applications associated with wireless HPI inference, including user identification, intruder detection, indoor user localization & tracking, person counting, gesture tracking, vital signs monitoring, keystroke recognition, integration into real-world systems, and communication efficiency enhancement.

Furthermore, we highlight future trends of wireless HPI inference from three aspects: challenges with IoT devices, integration with machine learning techniques, and increasing adoption of mmWave applications. Wireless HPI inference is an emerging research area and there is an urgent need for further investigations to detect, quantify, and protect human privacy in ubiquitous wireless environments.

REFERENCES

- [1] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "WiFi-ID: Human identification using WiFi signal," in *Proc. Int. Conf. Distrib. Comput. Sens. Syst. (DCOSS)*, 2016, pp. 75–82.
- [2] D. Vasisht, A. Jain, C.-Y. Hsu, Z. Kabelac, and D. Katabi, "Duet: Estimating user position and identity in smart homes using intermittent and incomplete RF-data," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, pp. 1–21, Jul. 2018.
- [3] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proc. 33rd Annu. ACM Conf. Human Factors Comput. Syst.*, 2015, pp. 837–846.
- [4] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking vital signs during sleep leveraging off-the-shelf WiFi," in *Proc. 16th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2015, pp. 267–276.
- [5] A. S. Abrar, A. Luong, P. Hillyard, and N. Patwari, "Pulse rate monitoring using narrowband received signal strength measurements," in *Proc. 1st ACM Int. Workshop Device-Free Human Sens.*, 2019, pp. 10–13.
- [6] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with Wi-Fi!," in *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 593–604.
- [7] M. Zhao, F. Adib, and D. Katabi, "Emotion recognition using wireless signals," *Commun. ACM*, vol. 61, pp. 91–100, Aug. 2018.
- [8] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proc. 19th Annu. Int. Conf. Mobile Comput. Netw.*, 2013, pp. 27–38.
- [9] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures," in *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 617–628.
- [10] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using WiFi signals," in *Proc. ACM Int. Joint Conf. Perv. Ubiquitous Comput.*, 2016, pp. 363–373.
- [11] H. Abdelnasser, M. Youssef, and K. A. Harras, "WiGest: A ubiquitous WiFi-based gesture recognition system," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, 2015, pp. 1472–1480.
- [12] O. Zhang and K. Srinivasan, "Mudra: User-friendly fine-grained gesture recognition using WiFi signals," in *Proc. 12th Int. Conf. Emerg. Netw. Exp. Technol.*, 2016, pp. 83–96.
- [13] S. Sharma, H. Mohammadmoradi, M. Heydariaan, and O. Gnawali, "Device-free activity recognition using Ultra-Wideband radios," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, 2019, pp. 1029–1033.
- [14] J. Wang, D. Vasisht, and D. Katabi, "RF-IDraw: Virtual touch screen in the air using RF signals," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 44, pp. 235–246, Aug. 2014.
- [15] L. Sun, S. Sen, D. Koutsonikolas, and K.-H. Kim, "WiDraw: Enabling hands-free drawing in the air on commodity WiFi devices," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 77–89.
- [16] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke recognition using WiFi signals," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 90–102.
- [17] S. Fang, I. Markwood, Y. Liu, S. Zhao, Z. Lu, and H. Zhu, "No training hurdles: Fast training-agnostic attacks to infer your typing," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2018, pp. 1747–1760.
- [18] B. Chen, Y. Venamandra, and K. Srinivasan, "Tracking keystrokes using wireless signals," in *Proc. 13th Annu. Int. Conf. Mobile Syst., Appl., Serv.*, 2015, pp. 31–44.
- [19] M. Li et al., "When CSI meets public Wi-Fi: Inferring your mobile phone password via WiFi signals," in *Proc. ACM Conf. Comput. Commun. Secur.*, 2016, pp. 1068–1079.
- [20] E. Yang et al., "Wireless training-free keystroke inference attack and defense," *IEEE/ACM Trans. Netw.*, vol. 30, no. 4, pp. 1733–1748, Aug. 2022.
- [21] E. Yang, Q. He, and S. Fang, "WINK: Wireless inference of numerical keystrokes via zero-training spatiotemporal analysis," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2022, pp. 3033–3047.
- [22] S. Pradhan, W. Sun, G. Baig, and L. Qiu, "Combating replay attacks against voice assistants," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 3, pp. 1–26, Sep. 2019.
- [23] Y. Zhu et al., "Adversarial WiFi sensing," 2018, *arXiv:1810.10109*.
- [24] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single WiFi access point," in *Proc. 13th Usenix Conf. Netw. Syst. Design Implement.*, 2016, pp. 165–178.
- [25] K. Van Loon et al., "Wireless non-invasive continuous respiratory monitoring with FMCW radar: A clinical validation study," *J. Clin. Monit. Comput.*, vol. 30, pp. 797–805, Dec. 2016.
- [26] L. Anitori, A. de Jong, and F. Nennie, "FMCW radar for life-sign detection," in *Proc. IEEE Radar Conf.*, 2009, pp. 1–6.
- [27] C.-Y. Hsu, A. Ahuja, S. Yue, R. Hristov, Z. Kabelac, and D. Katabi, "Zero-effort in-home sleep and insomnia monitoring using radio signals," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 1, pp. 1–18, Sep. 2017.
- [28] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall," *ACM Trans. Graph.*, vol. 34, pp. 1–13, Nov. 2015.
- [29] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3D tracking via body radio reflections," in *Proc. 11th USENIX Conf. Netw. Syst. Design Implement.*, 2014, pp. 317–329.
- [30] C.-Y. Hsu, Y. Liu, Z. Kabelac, R. Hristov, D. Katabi, and C. Liu, "Extracting gait velocity and stride length from surrounding radio signals," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2017, pp. 2116–2126.
- [31] C. Li, J. Ling, J. Li, and J. Lin, "Accurate doppler radar noncontact vital sign detection using the RELAX algorithm," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 3, pp. 687–695, Mar. 2010.
- [32] A. D. Droitcour, O. Boric-Lubecke, and G. T. A. Kovacs, "Signal-to-noise ratio in doppler radar system for heart and respiratory rate measurements," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 10, pp. 2498–2507, Oct. 2009.
- [33] M. Ascione, A. Buonanno, M. D'Urso, L. Angrisani, and R. S. L. Moriello, "A new measurement method based on music algorithm for through-the-wall detection of life signs," *IEEE Trans. Instrument. Meas.*, vol. 62, no. 1, pp. 13–26, Jan. 2013.
- [34] C. Li, V. M. Lubecke, O. Boric-Lubecke, and J. Lin, "A review on recent advances in doppler radar sensors for noncontact healthcare monitoring," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 5, pp. 2046–2060, May 2013.
- [35] T. Rahman et al., "DoppleSleep: A contactless unobtrusive sleep sensing system using short-range doppler radar," in *Proc. ACM Int. Joint Conf. Perv. Ubiquitous Comput.*, 2015, pp. 39–50.
- [36] B. Tan, K. Woodbridge, and K. Chetty, "A real-time high resolution passive WiFi doppler-radar and its applications," in *Proc. Int. Radar Conf.*, 2014, pp. 1–6.
- [37] F. Adib and D. Katabi, "See through walls with WiFi!," in *Proc. ACM SIGCOMM Conf. SIGCOMM*, 2013, pp. 75–86.
- [38] K. Chetty, G. E. Smith, and K. Woodbridge, "Through-the-wall sensing of personnel using passive bistatic WiFi radar at stand-off distances," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1218–1226, Apr. 2012.
- [39] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surv.*, vol. 46, pp. 1–32, Dec. 2013.
- [40] S. Yousefi, H. Narui, S. Dayal, S. Ermon, and S. Valaei, "A survey of human activity recognition using wifi CSI," 2017, *arXiv:1708.07129*.
- [41] Y. Ma, G. Zhou, and S. Wang, "WiFi sensing with channel state information: A survey," *ACM Comput. Surv.*, vol. 52, pp. 1–36, Jun. 2019.
- [42] M. A. A. Al-qaness et al., "Channel state information from pure communication to sense and track human motion: A survey," *Sensors*, vol. 19, p. 3329, Jul. 2019.
- [43] Z. Wang et al., "A survey on human behavior recognition using channel state information," *IEEE Access*, vol. 7, pp. 155986–156024, 2019.
- [44] Z. Wang, Z. Huang, C. Zhang, W. Dou, Y. Guo, and D. Chen, "CSI-based human sensing using model-based approaches: A survey," *J. Comput. Design Eng.*, vol. 8, pp. 510–523, Apr. 2021.
- [45] S. Tan, Y. Ren, J. Yang, and Y. Chen, "Commodity WiFi sensing in ten years: Status, challenges, and opportunities," *IEEE Internet Things J.*, vol. 9, no. 18, pp. 17832–17843, Sep. 2022.

- [46] J. Xiao, Z. Zhou, Y. Yi, and L. M. Ni, "A survey on wireless indoor localization from the device perspective," *ACM Comput. Surv.*, vol. 49, pp. 1–31, Jun. 2016.
- [47] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, "Wireless sensing for human activity: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1629–1645, 3rd Quart., 2020.
- [48] J. Liu, G. Teng, and F. Hong, "Human activity sensing with wireless signals: A survey," *Sensors*, vol. 20, p. 1210, Feb. 2020.
- [49] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [50] S. Fang, I. Markwood, and Y. Liu, "Manipulatable wireless key establishment," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, 2017, pp. 1–9.
- [51] S. Fang, I. Markwood, and Y. Liu, "Wireless-assisted key establishment leveraging channel manipulation," *IEEE Trans. Mobile Comput.*, vol. 20, no. 1, pp. 263–275, Jan. 2021.
- [52] S. Fang, Y. Liu, and P. Ning, "Mimicry attacks against wireless link signature and new defense using time-synched link signature," *IEEE Trans. Inf. Forensics Security*, vol. 11, pp. 1515–1527, 2016.
- [53] J. Xiong and K. Jamieson, "ArrayTrack: A fine-grained indoor location system," in *Proc. 10th USENIX Conf. Netw. Syst. Design Implement.*, 2013, pp. 71–84.
- [54] L. Cheng and J. Wang, "How can i guard my AP? Non-intrusive user identification for mobile devices using WiFi signals," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2016, pp. 91–100.
- [55] C. Xu et al., "SCPL: Indoor device-free multi-subject counting and localization using radio signal strength," in *Proc. 12th Int. Conf. Inf. Process. Sens. Netw. (IPSN)*, 2013, pp. 79–90.
- [56] H. Abdelnasser, K. A. Harras, and M. Youssef, "UbiBreathe: A ubiquitous non-invasive WiFi-based breathing estimator," in *Proc. 16th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2015, pp. 277–286.
- [57] F. Adib, Z. Kabelac, and D. Katabi, "Multi-person localization via RF body reflections," in *Proc. 12th USENIX Conf. Netw. Syst. Design Implement.*, 2015, pp. 279–292.
- [58] J. Lien et al., "Soli: Ubiquitous gesture sensing with millimeter wave radar," *ACM Trans. Graph.*, vol. 35, pp. 1–19, Jul. 2016.
- [59] H. Zhu, F. Xiao, L. Sun, R. Wang, and P. Yang, "R-TTWD: Robust device-free through-the-wall detection of moving human with WiFi," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1090–1103, May 2017.
- [60] C. Li and J. Lin, "Random body movement cancellation in doppler radar vital sign detection," *IEEE Trans. Microwave Theory Techn.*, vol. 56, no. 12, pp. 3143–3152, Dec. 2008.
- [61] M. Kotaru, K. Joshi, D. Bharadia, and S. Katti, "SpotFi: Decimeter level localization using WiFi," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 45, pp. 269–282, Aug. 2015.
- [62] K. Joshi, D. Bharadia, M. Kotaru, and S. Katti, "WiDeo: Fine-grained device-free motion tracing using rf backscatter," in *Proc. USENIX Conf. Netw. Systems Design Implement.*, 2015, pp. 189–204.
- [63] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni, "Pilot: Passive device-free indoor localization using channel state information," in *Proc. IEEE 33rd Int. Conf. Distrib. Comput. Syst.*, 2013, pp. 236–245.
- [64] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni, "FIMD: Fine-grained device-free motion detection," in *Proc. IEEE 18th Int. Conf. Parallel Distrib. Syst.*, 2012, pp. 229–235.
- [65] Y. Wang, K. Wu, and L. M. Ni, "WiFall: Device-free fall detection by wireless networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 581–594, Feb. 2017.
- [66] F. Hong, X. Wang, Y. Yang, Y. Zong, Y. Zhang, and Z. Guo, "WFID: Passive device-free human identification using WiFi signal," in *Proc. 13th Int. Conf. Mobile Ubiquitous Syst. Comput., Netw. Serv.*, 2016, pp. 47–56.
- [67] J. Liu, Y. Chen, Y. Wang, X. Chen, J. Cheng, and J. Yang, "Monitoring vital signs and postures during sleep using WiFi signals," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 2071–2084, Jun. 2018.
- [68] X. Li et al., "IndoTrack: Device-free indoor human tracking with commodity Wi-Fi," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, pp. 1–22, Sep. 2017.
- [69] D. Zhang, Y. Hu, and Y. Chen, "MTrack: Tracking multiperson moving trajectories and vital signs with radio signals," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3904–3914, Mar. 2021.
- [70] J. Wilson and N. Patwari, "See-through walls: Motion tracking using variance-based radio tomography networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 5, pp. 612–621, May 2011.
- [71] S. Y. Seidel and T. S. Rappaport, "914 MHz path loss prediction models for indoor wireless communications in multifloored buildings," *IEEE Trans. Antennas Propag.*, vol. 40, no. 2, pp. 207–217, Feb. 1992.
- [72] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free passive localization for wireless environments," in *Proc. 13th Annu. ACM Int. Conf. Mobile Comput. Netw.*, 2007, pp. 222–229.
- [73] A. S. Paul and E. A. Wan, "RSSI-based indoor localization and tracking using sigma-point Kalman smoothers," *IEEE J. Select. Topics Signal Process.*, vol. 3, no. 5, pp. 860–873, Oct. 2009.
- [74] S. Ikeda, H. Tsuji, and T. Ohtsuki, "Indoor event detection with eigenvector spanning signal subspace for home or office security," in *Proc. IEEE 68th Veh. Technol. Conf.*, 2008, pp. 1–5.
- [75] A. E. Kosba, A. Saeed, and M. Youssef, "RASID: A robust WLAN device-free passive motion detection system," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2012, pp. 180–189.
- [76] M. Moussa and M. Youssef, "Smart devices for smart environments: Device-free passive detection in real environments," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2009, pp. 1–6.
- [77] M. Nakatsuka, H. Iwatani, and J. Katto, "A study on passive crowd density estimation using wireless sensors," in *Proc. 4th Int. Conf. Mobile Comput. Ubiquitous Netw. (ICMU)*, 2008, pp. 1–6.
- [78] Y. Yuan, J. Zhao, C. Qiu, and W. Xi, "Estimating crowd density in an RF-based dynamic environment," *IEEE Sensors J.*, vol. 13, no. 10, pp. 3837–3845, Oct. 2013.
- [79] O. Kaltiokallio, H. Yigitler, R. Jäntti, and N. Patwari, "Non-invasive respiration rate monitoring using a single COTS TX-RX pair," in *Proc. 13th Int. Symp. Inf. Process. Sens. Netw. (IPSN)*, 2014, pp. 59–69.
- [80] N. Patwari, L. Brewer, Q. Tate, O. Kaltiokallio, and M. Bocca, "Breathfinding: A wireless network that monitors and locates breathing in a home," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 1, pp. 30–42, Feb. 2014.
- [81] N. Patwari, J. Wilson, S. Ananthanarayanan, S. K. Kasera, and D. R. Westenskow, "Monitoring breathing via signal strength in wireless networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 8, pp. 1774–1786, Aug. 2013.
- [82] H. Li, W. Yang, J. Wang, Y. Xu, and L. Huang, "WiFinger: Talk to your smart devices with finger-grained gesture," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 250–261.
- [83] M. N. A. Nipu, S. Talukder, M. S. Islam, and A. Chakrabarty, "Human identification using WIFI signal," in *Proc. Joint 7th Int. Conf. Inform., Electron. Vis. (ICIEV) 2nd Int. Conf. Imag., Vis. Pattern Recognit. (icIVPR)*, 2018, pp. 300–304.
- [84] K. Qian, C. Wu, Z. Yang, Y. Liu, and Z. Zhou, "PADS: Passive detection of moving targets with dynamic speed using PHY layer information," in *Proc. 20th IEEE Int. Conf. Parallel Distrib. Syst. (ICPADS)*, 2014, pp. 1–8.
- [85] K. Qian, C. Wu, Z. Yang, Y. Liu, F. He, and T. Xing, "Enabling contactless detection of moving humans with dynamic speeds using CSI," *ACM Trans. Embed. Comput. Syst.*, vol. 17, pp. 1–18, Jan. 2018.
- [86] S. Tan and J. Yang, "WiFinger: Leveraging commodity WiFi for fine-grained finger gesture recognition," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2016, pp. 201–210.
- [87] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool Release: Gathering 802.11n Traces with channel state information," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, p. 53, Jan. 2011.
- [88] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity WiFi," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 53–64.
- [89] A. Virmani and M. Shahzad, "Position and orientation agnostic gesture recognition using WiFi," in *Proc. 15th Annu. Int. Conf. Mobile Syst., Appl., Serv.*, 2017, pp. 252–264.
- [90] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
- [91] E. Soltanaghaei, A. Kalyanaraman, and K. Whitehouse, "Peripheral WiFi vision: Exploiting multipath reflections for more sensitive human sensing," in *Proc. 4th Int. Workshop Phys. Anal.*, 2017, pp. 13–18.
- [92] Y. Xie, J. Xiong, M. Li, and K. Jamieson, "mD-track: Leveraging multi-dimensionality for passive indoor Wi-Fi tracking," in *Proc. 25th Annu. Int. Conf. Mobile Comput. Netw.*, 2019, pp. 1–16.
- [93] W. Dargie and C. Poellabauer, *Fundamentals of Wireless Sensor Networks: Theory and Practice*. Hoboken, NJ, USA: Wiley, 2010.
- [94] B. R. Mahafza, *Radar Systems Analysis and Design Using MATLAB*. Boca Raton, FL, USA: CRC Press, 2002.
- [95] P. Molchanov, S. Gupta, K. Kim, and K. Pulli, "Short-range FMCW monopulse radar for hand-gesture sensing," in *Proc. IEEE Radar Conf. (RadarCon)*, 2015, pp. 1491–1496.
- [96] M. Zhao et al., "Through-wall human pose estimation using radio signals," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7356–7365.

- [97] C.-Y. Hsu, R. Hristov, G.-H. Lee, M. Zhao, and D. Katabi, "Enabling identification and behavioral sensing in homes using radio reflections," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2019, pp. 1–13.
- [98] O. Boric-Lubecke, V. M. Lubecke, A. D. Droitcour, B.-K. Park, and A. Singh, *Doppler Radar Physiological Sensing*. Hoboken, NJ, USA: Wiley, 2016.
- [99] F.-K. Wang, T.-S. Horng, K.-C. Peng, J.-K. Jau, J.-Y. Li, and C.-C. Chen, "Single-antenna doppler radars using self and mutual injection locking for vital sign detection with random body movement cancellation," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 12, pp. 3577–3587, Dec. 2011.
- [100] C. Gu, C. Li, J. Lin, J. Long, J. Huangfu, and L. Ran, "Instrument-based noncontact Doppler radar vital sign detection system using heterodyne digital quadrature demodulation architecture," *IEEE Trans. Instrument. Meas.*, vol. 59, no. 6, pp. 1580–1588, Jun. 2010.
- [101] P. Nguyen, X. Zhang, A. Halbower, and T. Vu, "Continuous and fine-grained breathing volume monitoring from afar using wireless signals," in *Proc. 35th Annu. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, 2016, pp. 1–9.
- [102] H. Zhao, H. Hong, L. Sun, Y. Li, C. Li, and X. Zhu, "Noncontact physiological dynamics detection using low-power digital-IF Doppler radar," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1780–1788, Jul. 2017.
- [103] P. Hillyard et al., "Experience: Cross-technology radio respiratory monitoring performance study," in *Proc. 24th Annu. Int. Conf. Mobile Comput. Netw.*, 2018, pp. 487–496.
- [104] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. IEEE 19th Annu. Joint Conf. IEEE Comput. Commun. Soc. (INFOCOM)*, 2000, pp. 775–784.
- [105] C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu, and J. Cao, "Non-invasive detection of moving and stationary human with WiFi," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2329–2342, Nov. 2015.
- [106] E. Soltanaghaei et al., "Robust and practical WiFi human sensing using on-device learning with a domain adaptive model," in *Proc. 7th ACM Int. Conf. Syst. Energy-Effic. Build., Cities, Transp.*, 2020, pp. 150–159.
- [107] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free your CSI: A channel state information extraction platform for modern Wi-Fi chipsets," in *Proc. 13th Int. Workshop Wireless Netw. Testbeds, Exp. Eval. Charact.*, 2019, pp. 21–28.
- [108] Y. Zhu et al., "Et Tu Alexa? When commodity WiFi devices turn into adversarial motion sensors," in *Proc. Netw. Distrib. Syst. Secur. Symp. (NDSS)*, 2020, pp. 1–15.
- [109] A. T. Mariakakis, S. Sen, J. Lee, and K.-H. Kim, "SAIL: Single access point-based indoor localization," in *Proc. 12th Annu. Int. Conf. Mobile Syst., Appl., Serv.*, 2014, pp. 315–328.
- [110] S. Palipana, D. Rojas, P. Agrawal, and D. Pesch, "FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, pp. 1–25, Jan. 2018.
- [111] L. Davies and U. Gather, "The identification of multiple outliers," *J. Am. Stat. Assoc.*, vol. 88, pp. 782–792, Sep. 1993.
- [112] X. Liu, J. Cao, S. Tang, J. Wen, and P. Guo, "Contactless respiration monitoring via off-the-shelf WiFi devices," *IEEE Trans. Mobile Comput.*, vol. 15, no. 10, pp. 2466–2479, Oct. 2016.
- [113] A. S. Paul et al., "MobileRF: A robust device-free tracking system based on a hybrid neural network HMM classifier," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 159–170.
- [114] S. Sigg, U. Blanke, and G. Tröster, "The telepathic phone: Frictionless activity recognition from Wi-Fi-RSSI," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, 2014, pp. 148–155.
- [115] P. Melgarejo, X. Zhang, P. Ramanathan, and D. Chu, "Leveraging directional antenna capabilities for fine-grained gesture recognition," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 541–551.
- [116] T. Wei and X. Zhang, "mTrack: High-precision passive tracking using millimeter wave radios," in *Proc. ACM Conf. Mobile Comput. Netw. (MobiCom)*, pp. 117–129, 2015.
- [117] W. Xi et al., "Electronic frog eye: Counting crowd using WiFi," in *Proc. IEEE INFOCOM*, 2014, pp. 361–369.
- [118] S. Sigg, M. Scholz, S. Shi, Y. Ji, and M. Beigl, "RF-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals," *IEEE Trans. Mobile Comput.*, vol. 13, no. 4, pp. 907–920, Apr. 2014.
- [119] J. Hong and T. Ohtsuki, "Signal eigenvector-based device-free passive localization using array sensor," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1354–1363, Apr. 2015.
- [120] Y. Chen, W. Dong, Y. Gao, X. Liu, and T. Gu, "Rapid: A multimodal and device-free approach using noise estimation for robust person identification," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, pp. 1–27, Sep. 2017.
- [121] L. Guo, L. Wang, J. Liu, W. Zhou, and B. Lu, "HuAc: Human activity recognition using crowdsourced WiFi signals and skeleton data," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–15, Jan. 2018.
- [122] J. Wang et al., "LiFS: Low human-effort, device-free localization with fine-grained subcarrier information," in *Proc. ACM Conf. Mobile Comput. Netw. (MobiCom)*, 2016, pp. 243–256.
- [123] J. Wang et al., "Low human-effort, device-free localization with fine-grained subcarrier information," *IEEE Trans. Mobile Comput.*, vol. 17, no. 11, pp. 2550–2563, Nov. 2018.
- [124] J. Lv, D. Man, W. Yang, X. Du, and M. Yu, "Robust WLAN-based indoor intrusion detection using PHY layer information," *IEEE Access*, vol. 6, pp. 30117–30127, 2018.
- [125] B. Fang, N. D. Lane, M. Zhang, A. Boran, and F. Kawsar, "BodyScan: Enabling radio-based sensing on wearable devices for contactless activity and vital sign monitoring," in *Proc. ACM Conf. Mobile Syst., Appl., Serv. (MobiSys)*, pp. 97–110, 2016.
- [126] W. He, K. Wu, Y. Zou, and Z. Ming, "WiG: WiFi-based gesture recognition system," in *Proc. 24th Int. Conf. Comput. Commun. Netw. (ICCCN)*, 2015, pp. 1–7.
- [127] J. Shang and J. Wu, "A robust sign language recognition system with multiple Wi-Fi devices," in *Proc. Workshop Mobil. Evolv. Internet Archit.*, 2017, pp. 19–24.
- [128] C. Shi, J. Liu, H. Liu, and Y. Chen, "Smart user authentication through actuation of daily activities leveraging WiFi-enabled IoT," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2017, pp. 1–10.
- [129] D. Wu, D. Zhang, C. Xu, Y. Wang, and H. Wang, "WiDir: Walking direction estimation using wireless signals," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 351–362.
- [130] K. Wu, J. Xiao, Y. Yi, M. Gao, and L. M. Ni, "FILA: Fine-grained indoor localization," in *Proc. IEEE INFOCOM*, 2012, pp. 2210–2218.
- [131] S. Palipana, P. Agrawal, and D. Pesch, "Channel state information based human presence detection using non-linear techniques," in *Proc. 3rd ACM Int. Conf. Syst. Energy-Effic. Built Environ.*, 2016, pp. 177–186.
- [132] R. Zhou, X. Lu, P. Zhao, and J. Chen, "Device-free presence detection and localization with SVM and CSI fingerprinting," *IEEE Sensors J.*, vol. 17, no. 23, pp. 7990–7999, Dec. 2017.
- [133] W. Xi et al., "Instant and robust authentication and key agreement among mobile devices," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2016, pp. 616–627.
- [134] Z. Tian, J. Wang, X. Yang, and M. Zhou, "WiCatch: A Wi-Fi based hand gesture recognition system," *IEEE Access*, vol. 6, pp. 16911–16923, 2018.
- [135] R. Zhou, M. Hao, X. Lu, M. Tang, and Y. Fu, "Device-free localization based on CSI fingerprints and deep neural networks," in *Proc. 15th Annu. IEEE Int. Conf. Sens., Commun., Netw. (SECON)*, 2018, pp. 1–9.
- [136] N. Yu, W. Wang, A. X. Liu, and L. Kong, "QGGesture: Quantifying gesture distance and direction with WiFi signals," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, pp. 1–23, Mar. 2018.
- [137] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "RT-Fall: A real-time and contactless fall detection system with commodity WiFi devices," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 511–526, Feb. 2017.
- [138] W. Wang, Y. Chen, and Q. Zhang, "Privacy-preserving location authentication in Wi-Fi networks using fine-grained physical layer signatures," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1218–1225, Feb. 2016.
- [139] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 65–76.
- [140] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Device-free human activity recognition using commercial WiFi devices," *IEEE Sens. J. Select. Areas Commun.*, vol. 35, no. 5, pp. 1118–1131, May 2017.
- [141] S. D. Regani, Q. Xu, B. Wang, M. Wu, and K. J. R. Liu, "Driver authentication for smart car using wireless sensing," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 2235–2246, Mar. 2020.
- [142] H. Wang et al., "Human respiration detection with commodity Wifi devices: Do user location and body orientation matter?," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 25–36.

- [143] Y. Zeng, P. H. Pathak, and P. Mohapatra, "WiWho: WiFi-based person identification in smart spaces," in *Proc. 15th ACM/IEEE Int. Conf. Inf. Process. Sens. Netw. (IPSN)*, 2016, pp. 1–12.
- [144] Y. Zeng, P. H. Pathak, and P. Mohapatra, "Analyzing shopper's behavior through WiFi signals," in *Proc. 2nd Workshop Workshop Phys. Anal.*, 2015, pp. 13–18.
- [145] Y. Zeng, P. H. Pathak, C. Xu, and P. Mohapatra, "Your AP knows how you move: Fine-grained device motion recognition through WiFi," in *Proc. 1st ACM Workshop Hot Topics Wireless*, 2014, pp. 49–54.
- [146] F. Xiao, J. Chen, X. Xie, L. Gui, L. Sun, and R. Wang, "SEARE: A system for exercise activity recognition and quality evaluation based on green sensing," *IEEE Trans. Emerg. Topics Comput.*, vol. 8, no. 3, pp. 752–761, Jul.–Sep. 2020.
- [147] Q. Gao, J. Wang, X. Ma, X. Feng, and H. Wang, "CSI-based device-free wireless localization and activity recognition using radio image features," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10346–10356, Nov. 2017.
- [148] S. Salvador and P. Chan, "Toward accurate dynamic time warping in linear time and space," *Intell. Data Anal.*, vol. 11, pp. 561–580, Oct. 2007.
- [149] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *Int. J. Comput. Vis.*, vol. 40, pp. 99–121, Nov. 2000.
- [150] K. Chintalapudi, A. P. Iyer, and V. N. Padmanabhan, "Indoor localization without the pain," in *Proc. 16th Annu. Int. Conf. Mobile Comput. Netw.*, 2010, pp. 173–184.
- [151] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice with MATLAB*. Hoboken, NJ, USA: Wiley, 2014.
- [152] M. Bocca, O. Kaltiokallio, N. Patwari, and S. Venkatasubramanian, "Multiple target tracking with RF sensor networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 8, pp. 1787–1800, Aug. 2014.
- [153] S. Nannuru, Y. Li, Y. Zeng, M. Coates, and B. Yang, "Radio-frequency tomography for passive indoor multitarget tracking," *IEEE Trans. Mobile Comput.*, vol. 12, no. 12, pp. 2322–2333, Dec. 2013.
- [154] H. Li et al., "An indoor continuous positioning algorithm on the move by fusing sensors and Wi-Fi on smartphones," *Sensors*, vol. 15, pp. 31244–31267, Dec. 2015.
- [155] S. Duan, T. Yu, and J. He, "WiDriver: Driver activity recognition system based on WiFi CSI," *Int. J. Wireless Inf. Netw.*, vol. 25, pp. 146–156, Jun. 2018.
- [156] F. Zhang et al., "From fresnel diffraction model to fine-grained human respiration sensing with commodity Wi-Fi devices," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, pp. 1–23, Mar. 2018.
- [157] Y. Zeng, D. Wu, R. Gao, T. Gu, and D. Zhang, "FullBreathe: Full human respiration detection exploiting complementarity of CSI phase and amplitude of WiFi signals," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, pp. 1–19, Sep. 2018.
- [158] J. Zhang, W. Xu, W. Hu, and S. S. Kanhere, "WiCare: Towards in-situ breath monitoring," in *Proc. 14th EAI Int. Conf. Mobile Ubiquitous Syst. Comput., Netw. Serv.*, 2017, pp. 126–135.
- [159] X. Wang, C. Yang, and S. Mao, "PhaseBeat: Exploiting CSI phase data for vital sign monitoring with commodity WiFi devices," in *Proc. IEEE 37th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, 2017, pp. 1230–1239.
- [160] S. Sigg, S. Shi, F. Buesching, Y. Ji, and L. Wolf, "Leveraging RF-channel fluctuation for activity recognition: Active and passive systems, continuous and RSSI-based signal features," in *Proc. Int. Conf. Adv. Mobile Comput. Multimedia*, 2013, pp. 43–52.
- [161] H. Liu, Y. Wang, J. Liu, J. Yang, and Y. Chen, "Practical user authentication leveraging channel state information (CSI)," in *Proc. 9th ACM Symp. Inf. Comput. Commun. Secur.*, 2014, pp. 389–400.
- [162] X. Guo, B. Liu, C. Shi, H. Liu, Y. Chen, and M. C. Chuah, "WiFi-enabled smart human dynamics monitoring," in *Proc. 15th ACM Conf. Embed. Netw. Sens. Syst.*, 2017, pp. 1–13.
- [163] K. Ohara, T. Maekawa, and Y. Matsushita, "Detecting state changes of indoor everyday objects using Wi-Fi channel state information," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, pp. 1–28, Sep. 2017.
- [164] J. Zhang et al., "Data augmentation and dense-LSTM for human activity recognition using WiFi signal," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4628–4641, Mar. 2021.
- [165] H. Chen, Y. Zhang, W. Li, X. Tao, and P. Zhang, "ConFi: Convolutional neural networks based indoor Wi-Fi localization using channel state information," *IEEE Access*, vol. 5, pp. 18066–18074, 2017.
- [166] Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "SignFi: Sign language recognition using WiFi," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 2, pp. 1–21, Mar. 2018.
- [167] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Recognizing keystrokes using WiFi devices," *IEEE J. Select. Areas Commun.*, vol. 35, no. 5, pp. 1175–1190, Mar. 2017.
- [168] X. Zhu and A. B. Goldberg, "Introduction to semi-supervised learning," in *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 3, Springer, Cham, Switzerland: 2009, pp. 1–130.
- [169] H. Liu, Y. Wang, J. Liu, J. Yang, Y. Chen, and H. V. Poor, "Authenticating users through fine-grained channel information," *IEEE Trans. Mobile Comput.*, vol. 17, no. 2, pp. 251–264, Feb. 2018.
- [170] C. Li, M. Liu, and Z. Cao, "WiHF: Enable user identified gesture recognition with WiFi," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, 2020, pp. 586–595.
- [171] D. Wang, J. Yang, W. Cui, L. Xie, and S. Sun, "CAUTION: A robust WiFi-based Human authentication system via few-shot open-set recognition," *IEEE Internet Things J.*, vol. 9, no. 18, pp. 17323–17333, Sep. 2022.
- [172] C. Lin et al., "A contactless authentication system based on WiFi CSI," *ACM Trans. Sens. Netw.*, vol. 19, pp. 1–20, Mar. 2023.
- [173] M. S. Nixon, T. Tan, and R. Chellappa, *Human Identification Based on Gait*, vol. 4, Berlin, Germany: Springer, 2010.
- [174] Z. Zhou, Z. Yang, C. Wu, L. Shangguan, and Y. Liu, "Towards omnidirectional passive human detection," in *Proc. IEEE INFOCOM*, 2013, pp. 3057–3065.
- [175] J. Wang, Z. Tian, M. Zhou, J. Wang, X. Yang, and X. Liu, "Leveraging hypothesis testing for CSI based passive human intrusion direction detection," *IEEE Trans. Veh. Technol.*, vol. 70, no. 8, pp. 7749–7763, Aug. 2021.
- [176] Z. Ni and B. Huang, "Gait-based person identification and intruder detection using mm-wave sensing in multi-person scenario," *IEEE Sens. J.*, vol. 22, no. 10, pp. 9713–9723, May. 2022.
- [177] S. Kang, J.-K. Paik, A. Koschan, B. R. Abidi, and M. A. Abidi, "Real-time video tracking using PTZ cameras," in *Proc. 6th Int. Conf. Qual. Control Artif. Vis.*, 2003, pp. 103–111.
- [178] M. Moghavvemi and L. C. Seng, "Pyroelectric infrared sensor for intruder detection," in *Proc. IEEE Region 10 Conf. (TENCON)*, 2004, pp. 656–659.
- [179] R. Ismail, Z. Omar, and S. Suaibun, "Obstacle-avoiding robot with IR and PIR motion sensors," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 152, Oct. 2016, Art. no. 012064.
- [180] Q. He and S. Fang, "Phantom-CSI attacks against wireless liveness detection," in *Proc. 26th Int. Symp. Res. Attacks, Intrus. Def.*, 2023, pp. 440–454.
- [181] M. Abbas, M. Elhamshary, H. Rizk, M. Torki, and M. Youssef, "WiDeep: WiFi-based accurate and robust indoor localization system using deep learning," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, 2019, pp. 1–10.
- [182] K. Qian, C. Wu, Z. Yang, Y. Liu, and K. Jamieson, "Widar: Decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2017, pp. 1–10.
- [183] X. Tong, Y. Wan, Q. Li, X. Tian, and X. Wang, "CSI fingerprinting localization with low human efforts," *IEEE/ACM Trans. Netw.*, vol. 29, no. 1, pp. 372–385, Feb. 2021.
- [184] J. Ding, Y. Wang, H. Si, S. Gao, and J. Xing, "Three-dimensional indoor localization and tracking for mobile target based on WiFi sensing," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21687–21701, Nov. 2022.
- [185] S. Fang, Y. Liu, W. Shen, and H. Zhu, "Where are you from? Confusing location distinction using virtual multipath camouflage," in *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 225–236.
- [186] S. Fang, Y. Liu, W. Shen, H. Zhu, and T. Wang, "Virtual multipath attack and defense for location distinction in wireless networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 566–580, Feb. 2017.
- [187] H. Choi, M. Fujimoto, T. Matsui, S. Misaki, and K. Yasumoto, "Wi-CaL: WiFi sensing and machine learning based device-free crowd counting and localization," *IEEE Access*, vol. 10, pp. 24395–24410, 2022.
- [188] Z. Guo, F. Xiao, B. Sheng, L. Sun, and S. Yu, "TWCC: A robust through-the-wall crowd counting system using ambient WiFi signals," *IEEE Trans. Veh. Technol.*, vol. 71, no. 4, pp. 4198–4211, Apr. 2022.
- [189] D. Khan and I. W.-H. Ho, "CrossCount: Efficient device-free crowd counting by leveraging transfer learning," *IEEE Internet Things J.*, vol. 10, no. 5, pp. 4049–4058, Mar. 2023.
- [190] L. Ren, A. Yarovoy, and F. Fioranelli, "Grouped people counting using mm-wave FMCW MIMO radar," *IEEE Internet Things J.*, vol. 10, no. 22, pp. 20107–20119, Nov. 2023.

- [191] B. Wei, W. Hu, M. Yang, and C. T. Chou, "Radio-based device-free activity recognition with radio frequency interference," in *Proc. 14th Int. Conf. Inf. Process. Sens. Netw.*, 2015, pp. 154–165.
- [192] W. Meng, X. Chen, W. Cui, and J. Guo, "WiHGR: A robust WiFi-based human gesture recognition system via sparse recovery and modified attention-based BGRU," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 10272–10282, Jun. 2022.
- [193] Z. Wang, Z. Yu, X. Lou, B. Guo, and L. Chen, "Gesture-radar: A dual doppler radar based system for robust recognition and quantitative profiling of human gestures," *IEEE Trans. Human-Mach. Syst.*, vol. 51, no. 1, pp. 32–43, Feb. 2021.
- [194] S. Palipana, D. Salami, L. A. Leiva, and S. Sigg, "Pantomime: Mid-air gesture recognition with sparse millimeter-wave radar point clouds," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 5, pp. 1–27, Mar. 2021.
- [195] D. Zhang, H. Wang, Y. Wang, and J. Ma, "Anti-fall: A non-intrusive and real-time fall detector leveraging CSI from commodity WiFi devices," in *Proc. 13th Int. Conf. Smart Homes Health Telematics*, 2015, pp. 181–193.
- [196] Z. Yang, Y. Zhang, and Q. Zhang, "Rethinking fall detection with Wi-Fi," *IEEE Trans. Mobile Comput.*, vol. 22, no. 10, pp. 6126–6143, Oct. 2023.
- [197] S. Chen, W. Yang, Y. Xu, Y. Geng, B. Xin, and L. Huang, "AFall: Wi-Fi-based device-free fall detection system using spatial angle of arrival," *IEEE Trans. Mobile Comput.*, vol. 22, no. 8, pp. 4471–4484, Aug. 2023.
- [198] W. Li et al., "Real-time fall detection using Mmwave radar," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2022, pp. 16–20.
- [199] R. Dillier et al., "Continuous respiratory monitoring for sleep apnea screening by ambulatory hemodynamic monitor," *World J. Cardiol.*, vol. 4, pp. 121–127, Apr. 2012.
- [200] S.-Y. Lee, C. Guilleminault, H.-Y. Chiu, and S. S. Sullivan, "Mouth breathing," "nasal disuse," and pediatric sleep-disordered breathing," *Sleep Breath.*, vol. 19, pp. 1257–1264, Dec. 2015.
- [201] B. Peters, "What is an overnight sleep study (Polysomnogram)." 2022. [Online]. Available: <https://www.verywellhealth.com/what-to-expect-in-a-sleep-study-3015121>
- [202] Masimo. "MightySat fingertip pulse oximeter with bluetooth LE, RRP, & PVI." 2022. [Online]. Available: <https://www.masimopersonalhealth.com/products/mightysat-fingertip-pulse-oximeter-with-bluetooth-le-rrp-pvi>
- [203] C. G. Scully et al., "Physiological parameter monitoring from optical recordings with a mobile phone," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 2, pp. 303–306, Feb. 2012.
- [204] X. Wang, C. Yang, and S. Mao, "TensorBeat: Tensor decomposition for monitoring multiperson breathing beats with commodity WiFi," *ACM Trans. Intell. Syst. Technol. (TIST)*, vol. 9, pp. 1–27, Sep. 2017.
- [205] X. Zhang et al., "Wital: A COTS WiFi devices based vital signs monitoring system using NLOS sensing model," *IEEE Trans. Human-Mach. Syst.*, vol. 53, no. 3, pp. 629–641, Jun. 2023.
- [206] Z. Guo, W. Yuan, L. Gui, B. Sheng, and F. Xiao, "BreatheBand: A fine-grained and robust respiration monitor system using WiFi signals," *ACM Trans. Sens. Netw.*, vol. 19, pp. 1–18, May 2023.
- [207] Y. Li, C. Gu, T. Nikoubin, and C. Li, "Wireless radar devices for smart human-computer interaction," in *Proc. IEEE 57th Int. Midwest Symp. Circuits Syst. (MWSCAS)*, 2014, pp. 65–68.
- [208] P.-H. Juan, C.-Y. Chueh, and F.-K. Wang, "Distributed MIMO CW radar for locating multiple people and detecting their vital signs," *IEEE Trans. Microw. Theory Techn.*, vol. 71, no. 3, pp. 1312–1325, Mar. 2023.
- [209] G. Li, Y. Ge, Y. Wang, Q. Chen, and G. Wang, "Detection of human breathing in non-line-of-sight region by using mmWave FMCW radar," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, Sep. 2022.
- [210] X. Liu, J. Cao, S. Tang, and J. Wen, "Wi-sleep: Contactless sleep monitoring via WiFi signals," in *Proc. IEEE Real-Time Syst. Symp.*, 2014, pp. 346–355.
- [211] Q. He, E. Yang, S. Fang, and S. Zhao, "HoneyBreath: An ambush tactic against wireless breath inference," in *Proc. 19th Int. Conf. Mobile Ubiquitous Syst. Comput., Netw. Serv.*, 2023, pp. 203–226.
- [212] Y. Xiao, J. Lin, O. Boric-Lubecke, and V. M. Lubecke, "Frequency-tuning technique for remote detection of heartbeat and respiration using low-power double-sideband transmission in the ka-band," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 5, pp. 2023–2032, May. 2006.
- [213] C. Gu, "Short-range noncontact sensors for healthcare and other emerging applications: A review," *Sensors*, vol. 16, p. 1169, Jul. 2016.
- [214] G. Oliveri, S. Sciancalepore, S. Raponi, and R. Di Pietro, "BrokenStrokes: On the (in)security of wireless keyboards," in *Proc. 13th ACM Conf. Secur. Privacy Wireless Mobile Netw.*, 2020, pp. 231–241.
- [215] F. Li, X. Wang, H. Chen, K. Sharif, and Y. Wang, "ClickLeak: Keystroke leaks through multimodal sensors in cyber-physical social networks," *IEEE Access*, vol. 5, pp. 27311–27321, 2017.
- [216] K. Ling, Y. Liu, K. Sun, W. Wang, L. Xie, and Q. Gu, "SpiderMon: Towards using cell towers as illuminating sources for keystroke monitoring," in *Proc. INFOCOM*, 2020, pp. 666–675.
- [217] D. Shukla, R. Kumar, A. Serwadda, and V. V. Phoha, "Beware, your hands reveal your secrets!," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2014, pp. 904–917.
- [218] Y. Chen, T. Li, R. Zhang, Y. Zhang, and T. Hedgpeth, "EyeTell: Video-assisted touchscreen keystroke inference from eye movements," in *Proc. IEEE Symp. Secur. Privacy (SP)*, 2018, pp. 144–160.
- [219] P. Marquardt, A. Verma, H. Carter, and P. Traynor, "(Sp)iPhone: Decoding vibrations from nearby keyboards using mobile phone accelerometers," in *Proc. 18th ACM Conf. Comput. Commun. Secur.*, 2011, pp. 551–562.
- [220] H. Wang, T. T.-T. Lai, and R. Roy Choudhury, "MoLe: Motion leaks through smartwatch sensors," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 155–166.
- [221] Z. Zhang et al., "WiPOS: A POS terminal password inference system based on wireless signals," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 7506–7516, Aug. 2020.
- [222] (Microsoft Corp., Redmond, WA, USA). *Learn about Windows Hello and set it up.* (2023). [Online]. Available: <https://learn.microsoft.com/en-us/windows-hardware/design/device-experiences/windows-hello-fingerprint-authentication>
- [223] (Samsung, Suwon-si, South Korea). *What is the iris Scanning and How to use it on My Galaxy S9/S9+.* (2023). [Online]. Available: <https://www.samsung.com/ie/support/mobile-devices/what-is-the-iris-scanning-and-how-to-use-it-on-my-galaxy-s9s9/>
- [224] "Apple face ID." 2023. [Online]. Available: <https://support.apple.com/en-us/HT208108>
- [225] "Amazon Alexa." 2023. [Online]. Available: <https://developer.amazon.com/en-US/alexa>
- [226] "Google assistant." 2023. [Online]. Available: <https://assistant.google.com/>
- [227] "Arlo." 2023. [Online]. Available: <https://www.arlo.com/en-us/security-system/arlo-security-system.html>
- [228] "Ring." 2023. [Online]. Available: <https://ring.com/security-system>
- [229] "SimpliSafe." 2023. [Online]. Available: <https://simplisafe.com/home-security-shop-packages>
- [230] Y. He, Q. He, S. Fang, and Y. Liu, "MotionCompass: Pinpointing wireless camera via motion-activated traffic," in *Proc. 19th Annu. Int. Conf. Mobile Syst., Appl., Services*, 2021, pp. 215–227.
- [231] Y. He, Q. He, S. Fang, and Y. Liu, "Precise wireless camera localization leveraging traffic-aided spatial analysis," *IEEE Trans. Mobile Comput.*, early access, Nov. 15, 2023, doi: [10.1109/TMC.2023.3333272](https://doi.org/10.1109/TMC.2023.3333272).
- [232] Y. He, Q. He, S. Fang, and Y. Liu, "When free tier becomes free to enter: A non-intrusive way to identify security cameras with no cloud subscription," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2023, pp. 651–665.
- [233] "IndoorAtlas." 2023. [Online]. Available: <https://www.indooratlas.com/>
- [234] "estimote." 2023. [Online]. Available: <https://estimote.com/#This-is-awesome>
- [235] "AXIS people counter." 2023. [Online]. Available: <https://www.axis.com/products/axis-people-counter>
- [236] "Hikvision people counting." 2023. [Online]. Available: <https://www.hikvision.com/en/solutions/solutions-by-function/people-counting/>
- [237] "Azure Kinect DK." 2023. [Online]. Available: <https://www.microsoft.com/en-us/d/azure-kinect-dk/8pp5vxdm9nhq?activetab=pivot:overviewtab>
- [238] "Ultraleap 3Di." 2023. [Online]. Available: <https://www.ultraleap.com/product/ultraleap-3di/>
- [239] "Intel RealSense." 2023. [Online]. Available: <https://www.intelrealsense.com/>
- [240] "Fitbit." 2023. [Online]. Available: https://www.fitbit.com/global/us/home?utm_medium=sem&utm_source=google&utm_campaign=US_PF_ROAS&gad=1&gclid=CjwKCAjw44mlBhAQEiwAqP3eVurkCv1fgaucF389rqF_tveT9T3eRJUFNqV5UEoD0rCeP1yDo2cQEhCoeUQAvD_BwE&gclid=aw.ds
- [241] "Apple watch." 2023. [Online]. Available: <https://www.apple.com/shop/buy-watch?afid=p238>

- [242] "Garmin." 2023. [Online]. Available: <https://www.garmin.com/en-US/c/wearables-smartwatches/?series=BRAND18900&sorter=featuredProducts-desc>
- [243] "Nest." 2023. [Online]. Available: https://store.google.com/category/connected_home?hl=en-US
- [244] "Ecobee." 2023. [Online]. Available: <https://www.ecobee.com/en-us/SmartThings>
- [245] "SmartThings." 2023. <https://www.smartthings.com/>
- [246] L. Özgür, V. K. Akram, M. Challenger, and O. Dağdeviren, "An IoT based smart thermostat," in *Proc. Int. Conf. Electr. Electron. Eng. (ICEEE)*, 2018, pp. 252–256.
- [247] B. Zhou et al., "Smart home energy management systems: Concept, configurations, and scheduling strategies," *Renew. Sustain. Energy Rev.*, vol. 61, pp. 30–40, Aug. 2016.
- [248] O. Krejcar, "User localization for large artifacts prebuffering and safety possibilities in mobile embedded systems," in *Proc. Int. Conf. Adv. Comput. Control*, 2009, pp. 41–45.
- [249] T.-Y. Tung and D. Gündüz, "DeepWiVe: Deep-learning-aided wireless video transmission," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 9, pp. 2570–2583, Sep. 2022.
- [250] G. Liang, H. Yu, X. Guo, and Y. Qin, "Joint access selection and bandwidth allocation algorithm supporting user requirements and preferences in heterogeneous wireless networks," *IEEE Access*, vol. 7, pp. 23914–23929, 2019.
- [251] M. Chen, U. Challita, W. Saad, C. Yin, and M. Debbah, "Artificial neural networks-based machine learning for wireless networks: A tutorial," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3039–3071, 4th Quart., 2019.
- [252] M. S. Elbamby, C. Perfecto, M. Bennis, and K. Doppler, "Edge computing meets millimeter-wave enabled VR: Paving the way to cutting the cord," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2018, pp. 1–6.
- [253] (Congressional Research Service, Washington, DC, USA). *The Internet of Things (IoT): An Overview*. 2020. [Online]. Available: <https://crsreports.congress.gov/product/pdf/IF/IF11239>
- [254] M. Hasan, "State of IoT 2022: Number of connected IoT devices growing % to 14.4 billion globally." 2022. [Online]. Available: <https://iot-analytics.com/number-connected-iot-devices/#:text=The>
- [255] T. J. Pierson, T. Peters, R. Peterson, and D. Kotz, "Proximity detection with single-antenna IoT devices," in *Proc. 25th Annu. Int. Conf. Mobile Comput. Netw.*, 2019, pp. 1–15.
- [256] E. Yang and S. Fang, "GPSKey: GPS-based secret key establishment for intra-vehicle environment," in *Proc. Workshop Automot. Auton. Veh. Secur.*, 2022, pp. 1–8.
- [257] E. Yang, S. Fang, and D. Shen, "DASK: Driving-assisted secret key establishment," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, 2022, pp. 73–81.
- [258] A. C. Jose and R. Malekian, "Improving smart home security: Integrating logical sensing into smart home," *IEEE Sensors J.*, vol. 17, no. 13, pp. 4269–4286, Jul. 2017.
- [259] N. Lakshmanan, I. Bang, M. S. Kang, J. Han, and J. T. Lee, "SurFI: Detecting surveillance camera looping attacks with Wi-Fi channel state information," in *Proc. 12th Conf. Secur. Priv. Wireless Mobile Netw.*, 2019, pp. 239–244.
- [260] J. Han et al., "Do you feel what i hear? Enabling autonomous IoT device pairing using different sensor types," in *Proc. IEEE Symp. Security Privacy (SP)*, 2018, pp. 836–852.
- [261] M. Gao et al., "Deaf-Aid: Mobile IoT communication exploiting stealthy speaker-to-gyroscope channel," in *Proc. 26th Annu. Int. Conf. Mobile Comput. Netw.*, 2020, pp. 1–13.
- [262] X. Li, F. Yan, F. Zuo, Q. Zeng, and L. Luo, "Touch well before use: Intuitive and secure authentication for IoT devices," in *Proc. 25th Annu. Int. Conf. Mobile Comput. Netw.*, 2019, pp. 1–17.
- [263] W. Cui, B. Li, L. Zhang, and Z. Chen, "Device-free single-user activity recognition using diversified deep ensemble learning," *Appl. Soft Comput.*, vol. 102, 2021, Art. no. 107066.
- [264] S. Xu, Z. He, W. Shi, Y. Wang, T. Ohtsuki, and G. Guai, "Cross-person activity recognition method using snapshot ensemble learning," in *Proc. IEEE 96th Veh. Technol. Conf.*, 2022, pp. 1–5.
- [265] T. Liu et al., "Wavevoice: A noise-resistant multi-modal speech recognition system fusing mmWave and audio signals," in *Proc. 19th ACM Conf. Embed. Netw. Sensor Syst.*, 2021, pp. 97–110.
- [266] H. Ding et al., "UltraSpeech: Speech enhancement by interaction between ultrasound and speech," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 6, no. 3, pp. 1–25, Sep. 2022.
- [267] X. Lu et al., "AutoDLAR: A semi-supervised cross-modal contact-free human activity recognition system," *ACM Trans. Sen. Netw.*, to be published.
- [268] J. Zhang, Z. Tang, M. Li, D. Fang, P. Nurmi, and Z. Wang, "CrossSense: Towards cross-site and large-scale WiFi sensing," in *Proc. 24th Annu. Int. Conf. Mobile Comput. Netw.*, 2018, pp. 305–320.
- [269] F. Wang, J. Liu, and W. Gong, "WiCAR: WiFi-based in-car activity recognition with multi-adversarial domain adaptation," in *Proc. Int. Symp. Qual. Service*, 2019, pp. 1–10.
- [270] D. Adesina, C.-C. Hsieh, Y. E. Sagduyu, and L. Qian, "Adversarial machine learning in wireless communications using RF data: A review," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 1, pp. 77–100, 1st Quart., 2022.
- [271] Y. Vorobeychik and M. Kantarcioglu, *Adversarial Machine Learning*. San Rafael, CA, USA: Morgan & Claypool, 2018.
- [272] A. Kurakin, I. J. Goodfellow, and S. Bengio, "Adversarial machine learning at scale," in *Proc. Int. Conf. Learn. Represent.*, 2017, pp. 1–17.
- [273] Y. Shi, K. Davaslioglu, and Y. E. Sagduyu, "Generative adversarial network for wireless signal spoofing," in *Proc. ACM Workshop Wireless Secur. Mach. Learn.*, 2019, pp. 55–60.
- [274] C. Iovescu and S. Rao, "The fundamentals of millimeter wave radar sensors." 2020. [Online]. Available: https://www.ti.com/lit/wp/spyy005a/spyy005a.pdf?ts=1673037835044&ref_url=https
- [275] W. Hong, K.-H. Baek, and S. Ko, "Millimeter-wave 5G antennas for smartphones: Overview and experimental demonstration," *IEEE Trans. Antennas Propag.*, vol. 65, no. 2, pp. 6250–6261, Dec. 2017.
- [276] "5G routers, gateways and antennas (updated 2021)." Novotech.com. 2022. [Online]. Available: <https://novotech.com/5g-routers-gateways-antennas/>
- [277] Z. Wei, F. Zhang, S. Chang, Y. Liu, H. Wu, and Z. Feng, "MmWave radar and vision fusion for object detection in autonomous driving: A review," *Sensors*, vol. 22, p. 2542, Apr. 2022.
- [278] S. S. Ahmed, A. Schiessl, F. Gumbmann, M. Tiebout, S. Methfessel, and L.-P. Schmidt, "Advanced microwave imaging," *IEEE Microw. Mag.*, vol. 13, no. 6, pp. 26–43, Sep./Oct. 2012.
- [279] J. Palacios, G. Bielsa, P. Casari, and J. Widmer, "Communication-driven localization and mapping for millimeter wave networks," in *Proc. IEEE Conf. Comput. Commun.*, 2018, pp. 2402–2410.
- [280] Z. Yang, P. H. Pathak, Y. Zeng, X. Liran, and P. Mohapatra, "Monitoring vital signs using millimeter wave," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2016, pp. 211–220.
- [281] F. Wang, X. Zeng, C. Wu, B. Wang, and K. J. R. Liu, "Driver vital signs monitoring using millimeter wave radio," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 11283–11298, Jul. 2022.
- [282] W. Jiang et al., "Towards environment independent device free human activity recognition," in *Proc. 24th Annu. Int. Conf. Mobile Comput. Netw.*, 2018, pp. 289–304.
- [283] A. Santra, R. V. Ulaganathan, and T. Finke, "Short-range millimetric-wave radar system for occupancy sensing application," *IEEE Sensors Lett.*, vol. 2, no. 3, pp. 1–4, Sep. 2018.
- [284] P. Zhao et al., "mID: Tracking and identifying people with millimeter wave radar," in *Proc. 15th Int. Conf. Distrib. Comput. Sensor Syst. (DCOSS)*, 2019, pp. 33–40.
- [285] P. Janakaraaj, K. Jakkala, A. Bhuyan, Z. Sun, P. Wang, and M. Lee, "STAR: Simultaneous tracking and recognition through millimeter waves and deep learning," in *Proc. 12th IFIP Wireless Mobile Netw. Conf. (WMNC)*, 2019, pp. 211–218.
- [286] S. Basak and M. Gowda, "mmSpy: Spying phone calls using mmWave radars," in *Proc. IEEE Symp. Security Privacy (SP)*, 2022, pp. 1211–1228.
- [287] C. Wang, F. Lin, Z. Ba, F. Zhang, W. Xu, and K. Ren, "Wavesdropper: Through-wall word detection of human speech via commercial mmWave devices," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 6, no. 2, pp. 1–26, Jul. 2022.
- [288] P. Hu, W. Li, R. Spolaor, and X. Cheng, "mmEcho: A mmWave-based acoustic eavesdropping method," in *Proc. IEEE Symp. Security Privacy (SP)*, pp. 836–852, 2022, pp. 1840–1856.
- [289] Y. Cao, T. Lv, and W. Ni, "Intelligent reflecting surface aided multi-user mmWave communications for coverage enhancement," in *Proc. IEEE 31st Annu. Int. Symp. Pers., Indoor Mobile Radio Commun.*, 2020, pp. 1–6.
- [290] M. Nemati, J. Park, and J. Choi, "RIS-assisted coverage enhancement in millimeter-wave cellular networks," *IEEE Access*, vol. 8, pp. 188171–188185, 2020.
- [291] M. Canil, J. Pegoraro, and M. Rossi, "milliTRACE-IR: Contact tracing and temperature screening via mmWave and infrared sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 2, pp. 208–223, Dec. 2022.



Qiuye He received the M.S. degree from Xidian University, Xi'an, China, in 2019. She is pursuing the Ph.D. degree in computer science with the University of Oklahoma. Her research interests are in the area of wireless security and mobile computing.



Song Fang received the Ph.D. degree in computer science from the University of South Florida in 2018. He is currently an Assistant Professor with the School of Computer Science, University of Oklahoma. His research interests include wireless and mobile system security, cyber-physical systems and IoT security, and mobile computing. He is also interested in applying machine learning in cybersecurity.



Edwin Yang received the M.S. degree from Yonsei University, Seoul, South Korea, in 2017. He is currently pursuing the Ph.D. degree in computer science with the University of Oklahoma. His research interests are in the area of mobile system security and Internet of Things security.