Assisting Living by Wireless Sensing: The Role of Integrated Sensing and Communications in 6G Era

Xinyu Li, Student Member, IEEE, Yuanhao Cui, Member, IEEE, J. Andrew Zhang, Senior Member, IEEE, Fan Liu, Member, IEEE, Xiaojun Jing, Member, IEEE, Octavia A. Dobre, Fellow, IEEE

Abstract—Advance in wireless communication and signal processing facilitates integrated sensing and communication (ISAC) - a technology that combines sensing and communication functionalities to efficiently utilize congested wireless/hardware resources, and to pursue mutual benefits. Consequently, the future communications network will be perceptive. In this article, we provide a review of human-related sensing in the context of ISAC. We first present a general ISAC receiver signal processing framework, with a focus on human activity recognition (HAR). Based on geographical deployments, we then categorize current ISAC HAR into three classical configurations, namely monostatic, bistatic and distributed deployments, and discuss their properties, critical research problems and solutions. In order to facilitate system realization and improve the recognition performance, we further explore inherent connections between physical-layer system parameters and HAR performance metrics. Finally, we review major technical challenges and identify key open research problems.

Index Terms—Joint radar and communications, Integrated sensing and communications, Signal processing, Wireless sensing, Human activity recognition.

I. INTRODUCTION

BOOSTED by innovations in artificial intelligence and sensing technologies, human-centric decision-making systems are continuously explored to bridge a broad range of compelling applications, such as safety protection, smart cities, and remote health-caring. Particularly, with the merits of contactless, non-intrusive and all-weather day-and-night availability, various standardized wireless signals (WiFi, LoRa, etc.) have emerged as a new medium to capture ambient human motions, relying on predefined channel estimation outputs such as the received signal strength indicator (RSSI), and channel state information (CSI)\(^1\). However, such sensing functionality has been primarily implemented passively using the received wireless signals. The quality of above sensory data is fundamentally determined by the prefixed pilot structure, the standardized waveform, and the spatial relationships of deployed commodity wireless products.

Integrated sensing and communication (ISAC) can complement existing wireless communication commodity by building-in sensing functionality \(^2\). As such, wireless devices are able to jointly meet the requirements of sensing and communications, by exploring and balancing all dimensional (spatial/time/frequency) wireless resources, rather than only the standardized pilot structure \(^3\). Therefore, the future networks and user equipment will not only be capable of wireless communications, but also be perceptive \(^4\), i.e., intelligently sensing targets from reflected and scattered echoes for, e.g., human activity recognition (HAR).

In general, human activities impact the wireless signal propagation properties such as reflection, diffraction and scattering, which provides recognition opportunities through analyzing and mapping the variations of the received signals with a specific human activity \(^5\). However, on the one hand, when ISAC is employed in the cellular network to sense the surrounding environment, HAR may be performed in complicated scenarios such as multiple-person and outdoor. This recognition problem is more challenging comparing to its indoor counterpart, due to the difficulties to distinguish the human target from the objects of no interest. On the other hand, various network deployments also open up new opportunities to HAR. Therefore, it is essential to evaluate both wireless communication and sensing from a systemic viewpoint, from raw data processing to recognition algorithms to provide a bird’s eye view for new researchers in this area.

In this article, we provide a review of human-related sensing in the context of ISAC. We first elaborate on the overlap and divergence between wireless communication and sensing processing pipelines, by providing a systemic overview of the ISAC signal processing framework. After that, to explore the impact of wireless devices’ geographical and spatial relationships on sensing performance, we identify three typical ISAC configurations (e.g., monostatic, bistatic and distributed configurations), and discuss their respective key challenges in deployment and implementation, respectively. Furthermore, we analyze the impact of several ISAC physical system parameters on HAR performance, and discuss the relevant optimization principle by jointly considering human sensing and communications. Finally, we put forward research challenges and open research opportunities before concluding this work.
II. A Systemic View of ISAC Signal Processing

A communication- and sensing-capable radio emission can simultaneously extract the environmental information, while conveying communication data from the transmitter to the intended receiver(s). Even though a number of signaling strategies are able to achieve a unified sensing and communication waveform, the most straightforward implementation is to reuse the communication infrastructures for wireless sensing, with a low-cost and fast-deployment footprint [6]. To provide a systematic view for the evolution from the communication-only devices to the ISAC infrastructure, in this section, we introduce a general ISAC receiver signal processing framework by examining the similarities and differences of current communication and sensing signal processing procedures, with reference to the sensing application of HAR. For each procedure, various state-of-the-art technologies and corresponding challenges are detailed.

A. The Shared Procedures

This subsection presents the shared receiver procedures between the communication and sensing signal processing pipelines, as shown in the blue boxes of Fig. 1.

1) Time/Frequency Synchronization: Time and frequency synchronization are important for both communication and sensing, and existing techniques for communication can still be applied. However, residual time and frequency offsets have different impacts on the functions. For communication, the residual can be absorbed into channel estimation and further compensated using embedded pilots; for sensing, however, it is hard to compensate for such residuals, which may cause estimation error in sensing parameters, particularly when the transmitter and receiver are not synchronized. On the other hand, synchronization and reference signals with high auto-correlation characteristics are better candidates for sensing usage, because they promise a higher matched filter gain to resist background noise.

2) Noise/Interference Reduction: Interference, noise along with other signal distortions constitute unintended but ubiquitous aspects of any radio system. It is well known that low signal-to-interference-plus-noise ratio (SINR) severely degrades both communication quality and sensing performance, however, the communication and sensing functionalities show several divergences when dealing with interference. For instance, the desired communication and sensing signals may travel to the receiver via the same or different paths. All paths contain effective signals for communication, whereas non-desired paths act as interference for sensing and shall be removed. In most cases, this module may be implemented before and/or after the next signal transformation module.

3) Signal Transformation: After preliminary time-domain processing, signal transformation is employed for space-time-frequency analysis of ISAC signal measurements, and a 3D cube is constructed from complex-valued baseband samples, as shown in the right of Fig. 1. Such transformation may be shared by or exclusive to communication and sensing. In a nutshell, different signal transforms hold different properties, e.g., the fast Fourier transform (FFT) transforms the time series of signals to frequency domain for communication, and can also be used to obtain delay-Doppler presentation for sensing.

4) Signal Separation: Depending on the adopted signaling strategy, the sensing and communication signals may be tightly integrated in a unified ISAC waveform, or be loosely combined in the time/frequency/space domains. A coordination center is commonly employed to balance the communication and sensing performance. When the waveforms of communication and sensing clearly differs in time, frequency, space or code domains, existing signal classification techniques can be straightforwardly applied. However, separation becomes challenging when communication and sensing signals overlap in one or more domains. For instance, in a monastic ISAC base station (BS), the aliasing of the uplink communication signal and the target echo would probably lead to artifacts in the sensing pipeline.

B. Separate Procedures for Sensing

The communication processing steps are well known as shown in Fig. 1. We focus on the differences between communication and HAR pipelines in this subsection. Additionally, in Table 1, we summarize the sensing characteristics of different wireless networks from signal structure, deployment, and data processing perspectives.

1) Subcarriers/Antennas Selection: Due to frequency-selective fading and antenna deployments, the delay-Doppler patterns of some human motions may be different, depending on spectrum and antennas, and are susceptible to external factors such as the moving direction of human and the locations of antennas. As a result, the received sensing signals from some antenna links may not show significant variations with the human target moving, and cannot improve the HAR performance of the ISAC system, or may even degrade the performance. In this case, subcarriers/antennas selection is indispensable to retain the signals that show significant fluctuations with the human movement [7], by using, e.g., convex optimization technologies. Additionally, clustering algorithms such as K-means can also be applied to the subcarrier selection task.

2) Signal Compression: The goal of signal compression is to remove the redundancy in the discrete 3D range-Doppler-angle sensing signals, such as static background clutter and outdoor environmental noise. There are mainly two signal compression strategies. The first strategy is based on statistical dimension reduction approaches, such as independent component analysis (ICA) and PCA, while the second is the clustering approach based on the range-Doppler-angle estimates. Specifically, the first strategy generally removes the noise in the data by reducing the data dimension and retaining main components. However, since each dimension in the 3D data cube has a clear physical meaning, dimension reduction techniques may destroy the structure and the physical explanatory nature of the data cube. In contrast, the second strategy employs clustering-based methods, which can retain the data structure of the range-Doppler-angle data cube [8]. Specifically, CFAR may first be employed on the 3D range-Doppler-angle sensing data to detect the reflection points whose intensities exceed a reflection amplitude threshold. In
this way, the discrete data points that may belong to the human target are retained, while noise can be filtered out. However, with CFAR, the reflection points corresponding to objects of no interest often remain in the point cloud. To deal with this issue, clustering methods such as DBSCAN is applied to group the points of interest and remove the false alarm clusters. With the second strategy, the sensing data of the human target can be represented as a 3D range-Doppler-angle point cloud, and used for the subsequent feature extraction and activity recognition.

3) Feature Extraction: Radio features could be extracted by manual feature engineering, or by automatic DL algorithms. For manual feature extraction, the amplitude and phase of the received sensing signals are the most common features, because they can characterize the impact of human activity on signal propagation. In addition, the time-varying Doppler/micro-Doppler frequency shifts in signal phase, which correspond to the velocities of different components of the human body, are also effective for single-person activity recognition. Particularly, when using cloud points data for feature extraction, the intensity of the reflected points and the shape of the point cloud can be adopted to describe diverse human activities.

Several ML models [9] can also be employed to automatically extract HAR related features. Early researchers have exploited the combination of convolutions neural network (CNN) with radio-based HAR feature extraction. However, the temporal information of human motions gradually vanishes during CNN training, which may seriously degrade the classification performance in the next stage. Hence, even with higher computational complexity, the memory-enabled recurrent NNs, such as long short-term memory (LSTM) and gated recurrent unit (GRU), are commonly adopted to extract time-varying HAR-related features, resulting in significantly improved performances [9].

4) Activity Recognition: The algorithms for HAR can be divided into two categories: model-based algorithms and learning-based algorithms [5]. Model-based approaches are based on physical theories (e.g., the Fresnel Zone model) or statistical models (e.g., the Rician fading model), characterizing the relationship between human activities and the resultant signal variations. On the other hand, learning-based
approaches aim to learn the mapping between sensing measurements and the corresponding ground truth human activities.

In model-based algorithms, the amplitude and phase of the received sensing signals are generally employed as input features, and the signal dynamics concerning human movement are quantitatively calculated. However, device placement often has significant impact on the performance of model-based approaches [5], and it is challenging to infer human activities under complex scenarios, e.g., multi-person scenario.

In learning-based algorithms, the classifiers can be divided into two categories: conventional ML classifier and DL classifier. Hidden Markov model (HMM), K nearest neighbor (KNN) and support vector machine (SVM) are commonly used ML classifiers, while DL classifiers are often fully-connected NNs, CNNs, RNNs, etc. Both manually extracted features (e.g., range, angle, Doppler and micro-Doppler) in the receiving signals or the intensity and the shape of the cloud points) and automatically extracted features can be input to learning-based classifiers. However, in most cases, DL-based feature extractor and classifier are combined to process input data and then, classify human activity in an end-to-end manner without any human intervention.

III. HAR SIGNAL PROCESSING WITH VARIOUS DEPLOYMENTS

There are mainly three types of ISAC systems, i.e., monostatic deployment, bistatic deployment and distributed deployment (see Fig. 3), according to the locations of the transmitters and receivers. In this section, we discuss the properties of the three deployments for HAR, and present the key problems and techniques in these deployments.

A. Monostatic Deployment

A monostatic system transmits wireless signals to sense the environment, and captures the target echoes via the co-located sensing receiver. One example is a 5G New Radio (NR) BS that senses the environment using the echoes of its transmitted downlink communication signals. Self-interference (SI) and far-field HAR are two major issues of concern in such a deployment.

1) Self-Interference (SI): In a monostatic system, the leaked transmitted signals can interfere with the desired received echo signals, also known as SI. The strong SI can saturate the receiver and overwhelm the target echoes. Since most modern communication systems transmit continuous waveform, it is infeasible to use a dumb period for receiving echoes after an ultra short transmission period, like in a pulse radar. It is also impossible to remove the SI by using the transmission signal as the local oscillator input, like in a frequency modulated

| TABLE I | SENSING CHARACTERISTICS OF DIVERSE WIRELESS NETWORKS. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | 802.11n         | 802.11ac        | 802.11ad        | 802.11ax        | LTE             | NR              | LoRa            |
| Signal Structure | Range Resolution | Reference Signals | Pseudo Random Sequence | Available Frequency | Coverage | Deployment | Data Processing | Computing Hardware | Measurement for HAR | HAR Algorithm Challenge |
|                 | ~2 m            | STF LTF         | Golay           | 2.4 GHz 5 GHz   | ~8 m          | Cooperative Sensing | Protocol supported | Access point      | CSI, RSSI Round trip time (RTT) | HAR with small-size data Generalization to new persons Location/orientation-dependent HAR |
|                 | ~5 m            | LTF             | Zadoff Chu M Gold | 5 GHz 60 GHz | ~5 m        | Protocol supported | Protocol supported | Phone             | CSI RSSI Amplitude Phase | Location/orientation-dependent HAR Location/orientation-dependent HAR Slow fading |
|                 | ~1 m            | STF CEF         | Zadoff Chu M Gold | 2.4 GHz 5 GHz | ~8 m        | Protocol supported | Protocol supported | BS Phone          | Amplitude Phase | Recognition of remote micro-activities Slow fading |
|                 | ~2 m            | LTF             | 2 GHz 60 GHz 1 GHz 1 GHz | 800 MHz 1.8 GHz 2.6 GHz | ~8 m        | Protocol supported | Protocol supported | Edge device       | 169 MHz 433 MHz 868 MHz 915 MHz | |
|                 | ~0.05 m         | STF LTF         | SC-FDMA OFDMA   | 800 MHz 1.8 GHz 2.6 GHz | ~5 m        | Protocol supported | Protocol supported | 5 m              | 10 m 5 m 1 m | |
|                 | ~1 m            | LTF             | DFT-S-OFDM CP-OFDM | 7.125 GHz 52.6 GHz | ~5 m        | Protocol supported | Protocol supported | 2 m              | 2 m 1 m | |
|                 | ~0.5 m          | STF LTF         | Chirps          | 8 GHz 3 GHz 1 GHz | ~0.5 m      | Protocol supported | Protocol supported | 1 m              | 1 m 0.5 m | |
|                 | ~10 m           | DMRS SRS, CRS CSI, PRS | / | 8 GHz 3 GHz 1 GHz | ~15 m      | Protocol supported | Protocol supported | 5 m              | 5 m 1 m | |
|                 | ~50 m           | DMRS SRS, CRS | / | 8 GHz 3 GHz 1 GHz | ~100 m     | Protocol supported | Protocol supported | 10 m             | 10 m 5 m | |
|                 | ~30 m           | PRS SRS, CSI | / | 8 GHz 3 GHz 1 GHz | ~300 m     | Protocol supported | Protocol supported | 100 m            | 100 m 50 m | |
|                 | ~500 m          | Upchirps Sync Word Downchirps | / | 8 GHz 3 GHz 1 GHz | ~1000 m     | Protocol supported | Protocol supported | 1000 m           | 1000 m 500 m | |

1 SFD: Start Frame Delimiter; STF: Short Training Field; LTF: Long Training Field; CEF: Channel Estimation Field; DMRS: Demodulation Reference Signal; SRS: Sounding Reference Signal; CRS: Cell Reference Signal; CSI: Channel State Information; PRS: Positioning Reference Signals; PTRS: Phase Tracking Reference Signal. 2 OFDMA: Orthogonal Frequency Division Multiple Access; SC-FDMA: Single Carrier-FDMA; DFT-S-OFDM: Discrete Fourier Transform-Spread OFDM; CP-OFDM: Cyclic Prefix-OFDM.
continuous radar. Full duplex is a long-term solution to this issue, as described in [4], together with suboptimal near-term solutions, such as deploying a receiving antenna dedicated to sensing, widely separated from other antennas.

2) Far-Field HAR: According to the geographical relationship between transmitter and targets, the sensing area can be divided into two distinct regions, i.e., near-field and far-field. The electromagnetic characteristics in near-field and far-field regions result in diverse patterns of receiving signals, further affecting the sensing performance. In far-field, due to the low received signal power, low signal-to-noise ratio (SNR) and the multipath propagation, the reflected signals from the human target may be overwhelmed by the background clutter [10]. Especially, the micro-Doppler frequencies, which are efficient features for HAR, may be contained in signals too weak to be captured. In this case, other features such as the time-varying range information of different human body segments can also be employed for recognition [11]. In addition, in far-field HAR, when the beam is focused on the human body, most energy of reflected sensing signals is likely to be scattered away from the direction of the collocated receiver. As a result, the received echoes may be quite weak and overwhelmed by background noise. Receiver beamforming, beam alignment and initial scanning are therefore necessary to enhance the received SINR.

B. Bistatic Deployment

Benefiting from its easy realization, bistatic deployment draws much attention from both academia and industry. One typical example is Wi-Fi sensing [12]. Compared with monostatic systems, the bistatic deployment is more compatible with existing communication networks such as WiFi and cellular networks. Furthermore, the SI issue is naturally avoided with the spatially isolated transmitter and receiver. Nevertheless, phase offset removal and unknown data payload are two key issues in bistatic deployments for performing accurate HAR.

1) Phase Offset Removal: Due to the oscillator instability, phase offset usually exists in the received signals of bistatic systems, leading to measurement ambiguity and accuracy degradation. For instance, sampling frequency offset (SFO) generally introduces high variations in the phase of sensing signals, and can even drown out the small phase changes caused by human movement. Carrier frequency offset (CFO) causes Doppler estimation ambiguity and then speed ambiguity. There exist two types of strategies to deal with the phase offset. The first strategy aims to compensate the phase offset and recover the information loss to improve the performance of multi-class classification applications. Existing techniques typically exploit the fact that the phase offsets between different receive antennas are the same, based on which, cross-antenna correlation and cross-antenna ratio techniques are applied [4]. The second strategy discards the phase information and only
uses the signal magnitude, which results in degraded sensing performance [5].

2) Exploiting Data Payload for Sensing: A typical communication frame consists of standardized pilot signals and data payload signals to convey user data from the transmitter to the receiver. By treating the entire frame as sensing signals, we can potentially achieve higher SNRs and larger Doppler frequency resolution range, compared to using pilot only. However, in most existing bistatic ISAC systems, e.g., Wi-Fi sensing applications, the data payload is unknown at the receiver side. To this end, a possible strategy is to first decode the unknown data payload according to the channel estimation results, and then employing the entire frame for sensing. However, the actual benefit of achieving improved SNRs with the sensing-after-decoding scheme is yet to be verified, and may only be prominent over a limited range of SNRs. In the low SNR region, the estimation error would spread from channel estimation to data payload recovery, such that sensing performance may be suppressed when the estimation error is high. In the high SNR region, the SNR improvement from higher matched filter gain may contribute only marginally to the overall detection and estimation performance.

C. Distributed Deployment

Distributed deployment can provide spatial diversity of the illuminated target and obtain human activity information from various spatial perspectives to deal with target fluctuation, greatly improving the sensing performance [4]. However, with more transmitting and receiving devices, a systematic design and arrangement, such as receiving data fusion, infrastructure deployment and scheduling issues, are indispensable.

1) Data Fusion: Based on the relations of data sources, the data to be fused for HAR can be divided into three categories: complementary data, redundant data and cooperative data.
   - Complementary data: When the receiving data from the device represents different parts of the scene, fusing these data can lead to more complete global information.
   - Redundant data: When several input sources provide similar information, such as having a close spatial location relative to the target, these data can be fused to increase the confidence.
   - Cooperative data: The provided information can be combined to generate new features, such as location information about the target used in cooperative localization, which is typically more complex than the original representations.

To fuse the three types of data, three corresponding data fusion strategies can be employed, i.e., data-level fusion, feature-level fusion and decision-level fusion.

   - In data-level fusion, each node first sends its raw data to the fusion center (FC), the FC then aggregates the data and extracts the HAR-related information from the fused data for classification.
   - In feature-level fusion, each node can do some preprocessing of their data, e.g., by estimating sensing parameters from their individual data, and then send the outputs of the preprocessing to the FC. Such a strategy can not only extract unique information from various devices but also allow flexible algorithm designs since the feature extraction structures of different branches can be diverse. Additionally, federated learning can also be employed to fuse the parameters from diverse nodes while addressing privacy concerns.
     - In decision-level fusion, each node processes their observations locally and sends the respective decision results to the FC. With the recognition results transmitted by each node, the fusion strategies, e.g., weighted decision fusion and generalized likelihood ratio tests (GLRT), can be implemented for a decision-level fusion at FC. This strategy requires the least data exchange between FC and each node.

By fusing all the observations for a final decision, centralized data-level fusion can achieve excellent sensing performance. However, sending a huge amount of sensing data to the FC causes large communication burden and high hardware costs. Therefore, such a deployment has higher requirements for computing power, energy consumption and hardware performance. In contrast, although there is no significant performance advantage, decentralized feature-level fusion and decision-level fusion has less data transmission, resulting in saving on energy and computing at the FC.

2) Deployment of Host and Slave Devices: The deployment in a communication-centric ISAC system can be considered for two types of devices: the host devices such as BSs and WiFi access point, and the slave devices such as terminal equipment. Slave devices get communication access by linking to host devices, but both types of devices can also act as sensing transceivers. Host device deployment requires to consider interference (e.g. interference between BSs in cellular network). The principle in communication systems is that there is generally little signal overlap between cells, which is not suitable for ISAC systems, since interference also contains useful sensing information. Therefore, in ISAC systems, the deployment of host devices may face a trade-off between communication interference and sensing performance.

On the other hand, though slave device deployment is independent of the communication performance, the placement of terminals can affect sensing factors such as coverage, orientation and angles, and needs to be optimized for better sensing performance. For instance, though balanced deployment can increase coverage overlapping and reduce blind zones to a certain extent, it may suffer from the mismatching between performance and demand, and lead to a waste of resources. Therefore, slave device deployment needs to consider a trade-off between sensing performance and resource efficiency.

3) Scheduling Issue with Target Echoes: In addition to data fusion and device deployment, sensing human activities also impose challenges in resource scheduling on distributed ISAC systems. Since the human echoes could randomly appear in time, frequency, and spatial domains, novel ISAC scheduling algorithms are required to predict the appearance of random echo signals and schedule the echoes in an orderly manner [3]. To this end, some model-based or learning-based algorithms can be applied to echo prediction and detection. Furthermore, intelligent resource allocation algorithms can be designed to
generate scheduling strategies for all device nodes, based on context information like quality-of-service (QoS) requirements and battery consumption.

In light of the discussions above, we summarize the properties and challenges of the three types of deployments in Table I. It can be inferred that the distributed deployment can yield the best HAR performance due to spatial diversity and wide coverage. However, more signal processing and computational resources are required. In real-world applications, one can select the ISAC deployment by considering the site and resource constraints.

IV. DESIGN FACTORS

To design an ISAC HAR system, accuracy and coverage are two important performance metrics to be considered.

- **Accuracy**: Accuracy is a statistical measure that reflects how well the HAR can be performed. The higher the accuracy, the better the HAR performance of a system can be achieved.
- **Coverage**: The coverage of the HAR system is defined as the range within which the activities of human targets can be identified.

The parameters and properties of current communication signals and systems can have significant influences on sensing accuracy and coverage. The impact of several design factors on HAR performance are detailed in Table II.

V. CHALLENGES AND FUTURE OPPORTUNITIES

There exist a number of challenges in the research and development of HAR with ISAC signals. In this section, we briefly discuss several major research challenges and future opportunities.

A. Multi-person Sensing

When there are multiple moving persons in the sensing area, identifying the target of interest or recognizing the activities of multiple persons is a challenging problem in ISAC systems. The general idea is to extract the reflected signals of each person and then recognize the corresponding human activity with the separated signals. This can be typically realized in two strategies: separation via physical location and moving speed [8], or separation via signal statistics [14].

In the first strategy, the signals for different targets may be separated from the spatial dimension by using the range, angle, and/or moving speed information, which requires high resolutions in these domains. Furthermore, the reflected signals at different time need to be associated. However, because the human body is an extended target, classical data association algorithms such as joint probabilistic data association cannot be straightforwardly applied, which makes it imperative to adapt the ISAC signal association algorithm to human targets.

In the second strategy, separation via signal statistics, signals from multiple humans may be modelled as a linear sum of statistical independent signals, and the separation can be cast as a blind source separation problem. Techniques such as ICA can then be applied to separate the mixed signal [14]. The main challenges of applying this approach are that signals from multiple users may not always be statistically independent, and the signals may not always be linearly combined due to complicated propagation phenomena such as multi-hop and blockage.

B. Joint Design and Optimization

Communication and HAR generally have conflicting requirements for antenna placement and grouping. When using an array, sensing aims to increase antenna aperture and resolution by optimizing antenna placements and virtual subarrays, while communication focuses on beamforming gain for spatial diversity and spatial multiplexing with a low signal correlation among antennas. Such different requirements demand a trade-off for joint design. For instance, considering the benefits of antenna grouping in communication and sensing, using hybrid antenna arrays can be a low-cost, balanced option.

Moreover, to achieve higher accuracy of HAR, the transmitted signals can be optimized by jointly considering performance metrics for communication and HAR. An overview of such joint optimization for general ISAC is available [4]. For HAR, signal optimization can be conducted by referring to the impacts presented in Table III. In addition, the performance of HAR may depend on both the estimation accuracy of sensing parameters, such as delay, Doppler and angles, and the channel

| Deployment       | Properties                                      | Challenges                              | Applicable Scenarios              | Supported Wireless Systems |
|------------------|-------------------------------------------------|-----------------------------------------|-----------------------------------|----------------------------|
| Monostatic       | Known sensing signals                           | Self-interference                       | Single moving human target        | LTE, NR, LoRa              |
|                  | Synchronized transceiver                        | Far-field HAR                           | Short-range HAR                   |                            |
| Bistatic         | Tx/Rx spatially isolated                        | Phase offset removal                    | Through-the-wall HAR              | WiFi, LTE, NR, LoRa        |
|                  | Compatible with existing networks               | Unknown data payload                    | Multiple moving human targets     |                            |
| Distributed      | Wide coverage                                   | Data fusion                             | Through-the-wall HAR              | WiFi, LTE, NR, LoRa        |
| Deployment       | Multidirectional sensing                        | Device deployment                       | Compound activity recognition     |                            |
|                  | Multi-node collaboration                        | Scheduling issue                        | Multiple moving human targets     |                            |
|                  |                                                 |                                         | Persons moving with free locations/orientations |                            |

TABLE II
PROPERTIES AND CHALLENGES OF THE THREE ISAC DEPLOYMENTS.
**Impact on Sensing Pipeline**

- **Total Signal Bandwidth** $B$
  - Larger $B$ leads to better range resolution

- **Carrier Frequency** $f_c$
  - Greater $f_c$ leads to finer velocity resolution but smaller unambiguous velocity \(^2\)

- **Symbol Duration** $T_s$
  - Larger $T_s$ leads to longer unambiguous range but lower range resolution

- **Subcarrier Interval** $T_s^1$
  - Smaller $T$ leads to larger unambiguous velocity but lower velocity resolution

- **Antenna Array Design**
  - Better antenna aperture and higher angular resolution with more antennas

- **Transmission Power**
  - Positively correlated with the coverage of the sensing system

- **Antenna Gain**
  - Greater coverage in a certain direction with higher antenna gain

**Impact on HAR performance**

- Better multi-targets separation ability along range domain with larger $B$

- Different moving components of the human target can be recorded with finer granularity in velocity domain

- Longer detectable range with $T_s$ increasing, promising for far-field HAR

- Wider coverage in velocity domain to record the activities with higher velocity components

- Locating human target more precisely at the angular direction

- Distinguishing multiple targets at close angular directions

- Wider HAR coverage with higher transmission power

- Stronger echo signals and more robust to interference

- Wider HAR coverage

- Improving intensity of the reflected echoes with higher antenna gain

---

1 For single-subcarrier signals, $T$ is $T_s$, and is equal to $1/B$.
2 For OFDM signals, $T$ includes $T_s$ and cyclic prefix, and is equal to $N/B$, where $N$ is the number of subcarriers.

---

### TABLE III
**IMPACT OF ISAC PHYSICAL PARAMETERS ON HAR PERFORMANCE.**

| Physical Parameters             | Impact on Sensing Pipeline                                                                 | Impact on HAR performance                                                                 |
|---------------------------------|-------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|
| **Total Signal Bandwidth $B$**  | Larger $B$ leads to better range resolution                                               | Better multi-targets separation ability along range domain with larger $B$                |
| **Carrier Frequency $f_c$**     | Greater $f_c$ leads to finer velocity resolution but smaller unambiguous velocity \(^2\) | Different moving components of the human target can be recorded with finer granularity in velocity domain |
| **Symbol Duration $T_s$**       | Larger $T_s$ leads to longer unambiguous range but lower range resolution                 | Longer detectable range with $T_s$ increasing, promising for far-field HAR                |
| **Subcarrier Interval $T_s^1$** | Smaller $T$ leads to larger unambiguous velocity but lower velocity resolution             | Wider coverage in velocity domain to record the activities with higher velocity components |
| **Antenna Array Design**        | Better antenna aperture and higher angular resolution with more antennas                 | Locating human target more precisely at the angular direction                             |
| **Transmission Power**          | Positively correlated with the coverage of the sensing system                            | Distinguishing multiple targets at close angular directions                                |
| **Antenna Gain**                | Greater coverage in a certain direction with higher antenna gain                          | Wider HAR coverage with higher transmission power                                         |
|                                 |                                                                                            | Stronger echo signals and more robust to interference                                     |

---

information contained in the received signals. Hence, both Cramer-Rao lower bound (CRLB) and mutual information (MI) can be used as performance metric for HAR.

### C. Through-the-wall HAR

Sensing through walls is also a challenging but essential task in HAR. The first problem is how to model the correlation between signal attenuation and walls \([15]\). In existing wall-loss models, power-loss as the signal passes through the wall can generally be well estimated. However, RF signals generally experience unpredictable reflection and absorption as they pass through walls, greatly weakening the receiving signals and hence reducing the HAR information. With the increased environmental noise, the characteristics of human activities contained in the receiving signals are also reduced, which seriously affects the subsequent feature extraction, especially for similar activities with fine differences. DL-based approaches require less signal processing and human intervention to classify human activities through the walls. On the other hand, although being less studied, physical or statistical model-based algorithms can characterize the mathematical relationship between the received sensing signals and the human activities in the non-line-of-sight through-the-wall environment.

### D. HAR Robustness and Generalization

HAR with ISAC signals is sensitive to many factors such as the sensing environment, network settings, relative location of human target, geometry and mobility situations. For instance, different moving directions and orientations of the person with respect to the transceivers can result in various Doppler/micro-Doppler frequencies. How to improve the system generalization on recognizing human activities from diverse directions is a challenging issue. Furthermore, due to the unique behaviour of each individual, when different persons, such as a young person and an old person, perform the same activity, their movements are discrepant. It is essential to generalize the trained HAR algorithms when new persons or new environments emerge.

### E. Distributed Sensing

With the potentially significant improvement on coverage and HAR accuracy, sensing based on a distributed topology is the general trend. However, research on sensing under a distributed topology is still very limited. As presented in Section III-C, the challenges for distributed HAR mainly lie in deployment and cooperation between transceivers, fusion strategies on data from diverse receivers and scheduling issues with target return. In addition, there is almost no discussion on the HAR performance bound for distributed sensing networks yet, which is also a promising research direction.

### VI. Conclusion

In this article, a review of human activity sensing in the context of ISAC is provided. We illustrated the general pipeline of ISAC signal processing for HAR and analyzed the sensing characteristics of several typical wireless networks. According to the spatial locations of transceivers, we categorized ISAC systems into three typical deployments and then, elaborated characteristics and problems for HAR in these three deployments. Furthermore, the impact of several physical indicators
in ISAC system on the performance of HAR was discussed. To achieve higher accuracy of HAR, the parameters of signals can be optimized from time, space and frequency domains. Finally, we presented five key challenges and open research problems to facilitate a transition of the techniques into real-world assisted living applications.

REFERENCES

[1] B. Tan, Q. Chen, K. Chetty, K. Woodbridge, W. Li, and R. Piechocki, “Exploiting WiFi channel state information for residential healthcare informatics,” *IEEE Communications Magazine*, vol. 56, no. 5, pp. 130–137, May 2018.

[2] Y. Cui, F. Liu, X. Jing, and J. Mu, “Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges,” *IEEE Network*, vol. 35, no. 5, pp. 158–167, Sep. 2021.

[3] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, “Integrated sensing and communications: Towards dual-functional wireless networks for 6G and beyond,” 2021. [Online]. Available: https://arxiv.org/abs/2108.07165

[4] J. A. Zhang, M. L. Rahman, K. Wu, X. Huang, Y. J. Guo, S. Chen, and J. Yuan, “Enabling joint communication and radar sensing in mobile networks - a survey,” *IEEE Communications Surveys & Tutorials*, early access, 2021.

[5] Y. Ma, G. Zhou, and S. Wang, “WiFi sensing with channel state information: A survey,” *ACM Comput. Surv.*, vol. 52, no. 3, May 2020.

[6] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu, “An overview of signal processing techniques for joint communication and radar sensing,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1295–1315, Nov. 2021.

[7] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, “Device-free human activity recognition using commercial WiFi devices,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 5, pp. 1118–1131, May 2017.

[8] J. Pegoraro, F. Meneghello, and M. Rossi, “Multiperson continuous tracking and identification from mm-wave micro-doppler signatures,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 2994–3009, April 2021.

[9] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu, and Y. Liu, “Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities,” *ACM Computing Surveys (CSUR)*, vol. 54, no. 4, pp. 1–40, May 2021.

[10] E. Piuzzi, P. D’Atanasio, S. Pisa, E. Pittella, and A. Zambotti, “Complex radar cross section measurements of the human body for breath-activity monitoring applications,” *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 8, pp. 2247–2258, Aug. 2015.

[11] B. Jokanović and M. Amin, “Fall detection using deep learning in range-doppler radars,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 1, pp. 180–189, Feb. 2018.

[12] Z. Shi, J. A. Zhang, R. Y. Xu, and Q. Cheng, “Environment-robust device-free human activity recognition with channel-state-information enhancement and one-shot learning,” *IEEE Transactions on Mobile Computing*, vol. 21, no. 2, pp. 540–554, Feb. 2022.

[13] J. F. Chamberland and V. Veeravalli, “How dense should a sensor network be for detection with correlated observations?” *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5099–5106, Nov. 2006.

[14] Y. Zeng, D. Wu, J. Xiong, J. Liu, Z. Liu, and D. Zhang, “Multisense: Enabling multi-person respiration sensing with commodity wifi,” *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 4, no. 3, Sep. 2020.

[15] Z. Wang, K. Jiang, Y. Hou, Z. Huang, W. Dou, C. Zhang, and Y. Guo, “A survey on CSI-based human behavior recognition in through-the-wall scenario,” *IEEE Access*, vol. 7, pp. 78 772–78 793, 2019.