Integrated Sensing and Communication

From Theory to *Standardization*

HUAWEI
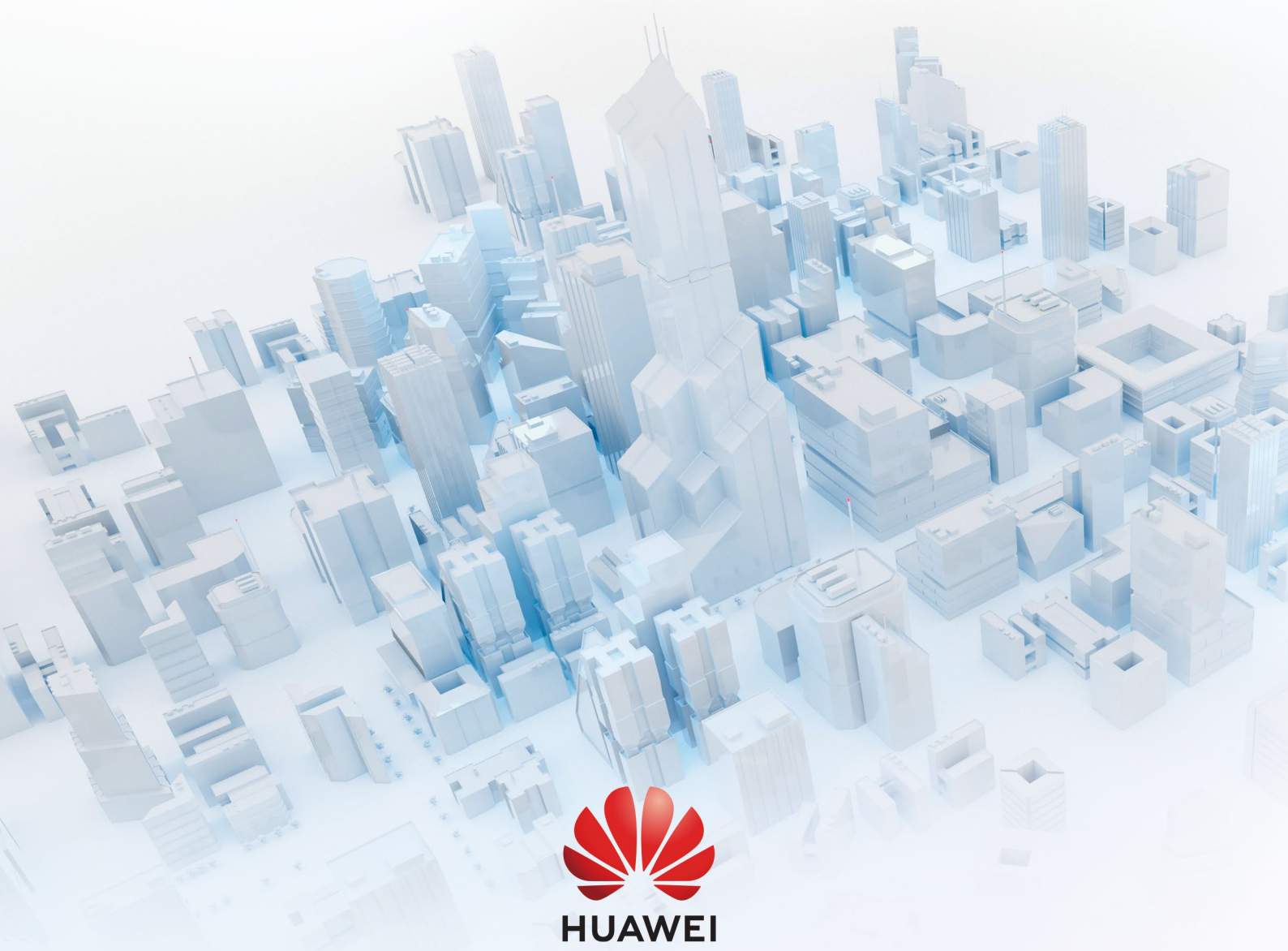
Editorial Note

Integrated sensing and communication (ISAC) is a technological innovation for 6G. In ISAC, wireless sensing serves as a foundational technology based on radio transmission sciences. By detecting radio wave propagation in the physical-world surroundings using digital receivers and advanced signal processing algorithms, we can achieve higher accuracy positioning of mobile devices and reconstruct the physical-world environment in real time. This sensing data can then be used to build a digital replica of the real world.

For 6G, the use of smaller radio wavelengths, such as centimeter wave, millimeter wave, and sub-terahertz wave, together with the advantages of higher frequency and wider spectrum bandwidth, allows us to create a digital twin of the physical world with high accuracy and resolution. Providing the capability of even detecting material properties, these new features will enable 6G to enter the era of beyond connectivity.

Two new enablers make wireless sensing more powerful in 6G. The first enabler is the native integration of sensing and communication. This integration enables every mobile device and every base station to perform 6G sensing without requiring additional spectrum or network investments. 6G ISAC can perform joint collaborative sensing of mobile devices and base stations in both monostatic and bistatic modes. Real-time reconstruction of the real world enables intelligent communications and drastically enhances wireless transmission reliability and capacity while also reducing power consumption. The second enabler is the integration of sensing and machine learning, which allows us to utilize 6G sensing data as a training dataset. In this case, machine learning, especially large model-based AI, can be constantly tuned to the real world to deliver customized AI services.

6G sensing is the engine for data-driven intelligence services and applications. Even if only one percent of the capacity of 6G networks were allocated for sensing, the amount of sensing-generated big data will reach the quettabyte level per day for all mobile devices and the zettabyte level per day for all base stations. From a machine learning point of view, 6G sensing data will be a major data source for AI training and will drive generative and interactive AI services.

In light of this, 6G sensing is the crucial interface bridging the physical and digital worlds. 6G sensing data is the primary source for building and delivering advanced AI services. The pervasiveness of 6G sensing will extend to every device and every base station, in conjunction with AI services and applications.

This special issue focuses on the research progress of 6G ISAC. The Huawei research team and our academic friends and industrial partners contribute their latest results. These heavy-lifting efforts will definitely lay the foundation for 6G ISAC technology. We want to express our heartfelt gratitude to everyone involved for their breakthrough efforts and contributions.

Dr. Wen Tong
Huawei Fellow

Dr. Peiying Zhu
Huawei Fellow

CONTENTS

ISAC Standardization

ISAC for Environment Sensing

ISAC Theories and Algorithms

ISAC Applications and Demos

Other

Technical Performance Requirements and Evaluation Methodologies for 6G Integrated Sensing and Communication

— The Next Milestone of ITU-R 6G Standardization

Yunhao Zhang [1], Min Zhang [2], Yan Chen [3], Peiying Zhu [3]

[1] Wireless Technology Lab

[2] Munich Advanced Wireless Technology Lab

[3] Ottawa Advanced Wireless Technology Lab

Abstract

The concept of integrated sensing and communication (ISAC), one of six usage scenarios of 6G to have been formally recognized by ITU-R, is gaining increasing attention from academia, industry, and standardization bodies. The next milestone in the ITU-R 6G standardization process centers on the technical performance requirements (TPRs) and evaluation methodologies of the six usage scenarios. In this paper, we will first introduce the timeline of TPR and evaluation studies in ITU-R and 3GPP. Then we will explore the potential ISAC use cases for 6G and discuss the service requirements that those cases give rise to, focusing on functionality and key performance metrics that motivate the use case-oriented test environment proposal for sensing-related TPR evaluation. Finally, we discuss the evaluation methodology of each test environment, and then introduce the progress of system- and link-level simulation.

Keywords

6G, sensing, integrated sensing and communication (ISAC), technical performance requirements (TPRs), evaluation methodology

1 Introduction

1.1 ITU-R Framework for IMT-2030

The use of higher frequency bands (from centimeter wave [cmWave] to millimeter wave [mmWave] and up to terahertz [THz]), wider bandwidth, and denser distribution of massive antenna arrays in future communication systems will enable sensing and communication to mutually enhance each other [1]. On the one hand, the communication system as a whole can be served as a sensor. It can explore wireless signal transmission, reflection, and scattering to "understand" the physical world, and potentially offer a broad range of new user experiences, which are commonly known as "network as a sensor." On the other hand, the capabilities of high-accuracy localization, imaging, and environment reconstruction achieved by passive sensing could potentially enhance communication performance, thereby realizing more accurate beamforming, faster beam failure recovery, and less signal overhead to track channel state information, which are commonly known as "sensing-assisted communication."

By sharing spectrum, hardware, signal processing, protocol stacks, and even cross-module or cross-layer information, the integrated sensing and communication (ISAC) functionality can be supported at different system levels, ranging from loosely coupled all the way to fully integrated. Compared with traditional sensing functionality, the ISAC in future wireless systems has potential benefits to reuse communication infrastructure and thus significantly reduce the implementation cost of supporting additional sensing functionality. A further

benefit is that it can improve sensing performance by leveraging the large-scale availability of deployed base stations and user equipment (UE).

With the completion of ITU-R Framework Recommendation for IMT-2030 in June 2023 [2], ISAC has been formally considered as one of the six usage scenarios of 6G (the sixth generation of mobile communications), as shown in Figure 1a, the so-called "wheel diagram." (ITU-R is the radio communication division of the International Telecommunication Union, and IMT-2030 is the International Mobile Telecommunications system for 2030 and beyond.) The typical use cases of this usage scenario include network-assisted navigation, gesture/posture detection and movement tracking, environment monitoring, and provision of sensing data/information on surroundings for artificial intelligence (AI), extended reality (XR, which includes both augmented reality and virtual reality), and digital twin applications. Along with communication capabilities, ISAC usage scenarios in 6G require support of high-precision positioning and sensing-related capabilities, as shown in Figure 1b, the so-called "palette diagram." These scenarios may also require capability metrics such as range/velocity/angle estimation, object and presence detection, localization, imaging, and mapping.

1.2 Timeline of Technical Performance Requirement and Evaluation Methodology Standardization

The next steps of ISAC standardization for 6G include the following aspects, which will interwork with and enhance each other despite being discussed by different groups:

(a) IMT-2030 "wheel diagram"

(b) IMT-2030 "palette diagram"
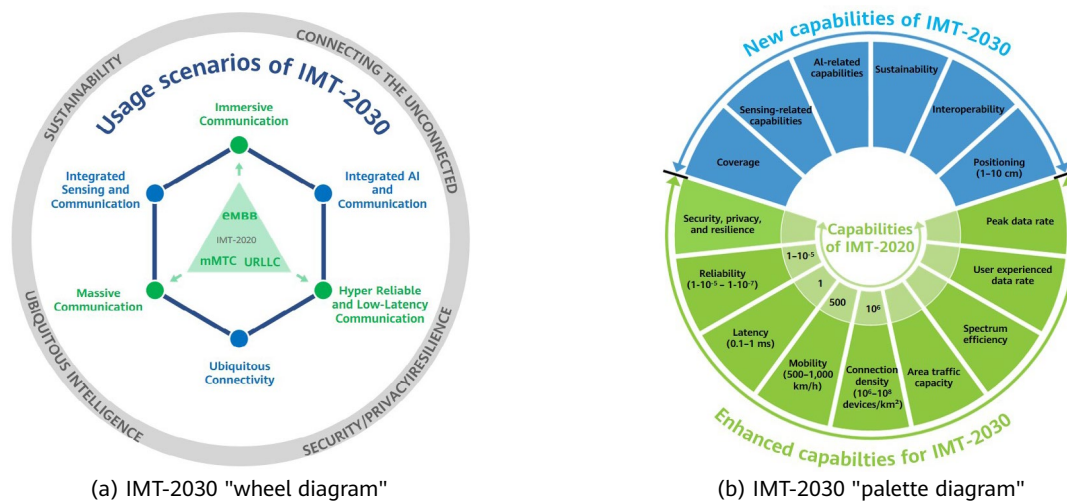
Figure 1 (a) Usage scenarios and overarching aspects of IMT-2030 (6G) in the "wheel diagram" and (b) Capabilities of IMT-2030 in the "palette diagram" [2]
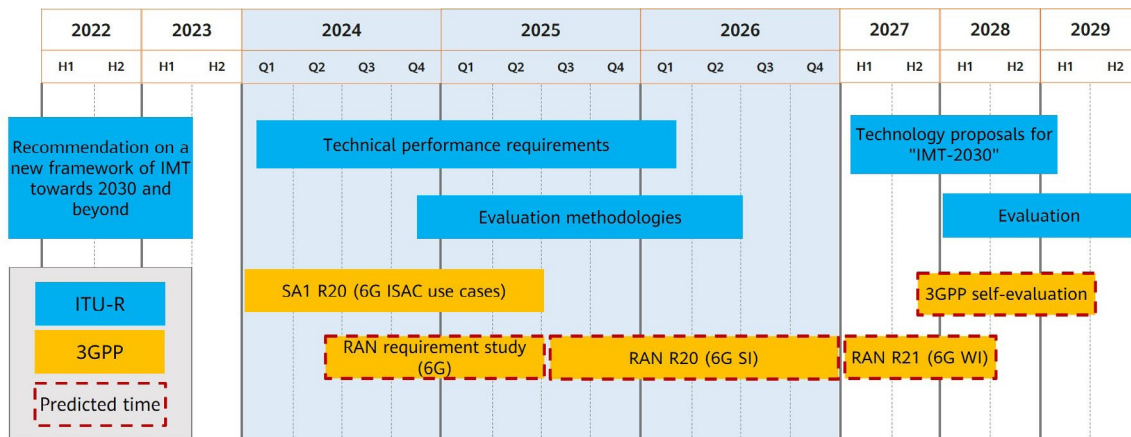
Figure 2 Timeline of 6G TPR and evaluation studies in ITU-R and 3GPP

- **Technical performance requirements (TPRs) and evaluation methodology study in ITU-R:** This step forms part of the second phase of ITU-R's IMT-2030 study planned for the period of 2024–2026. During this phase, performance metrics for future radio interface technologies (RITs) will be established, and the minimum values with respect to a set of typical test environments will be defined. These minimum requirements will serve as the prerequisites that RITs must meet in order for them to be considered 6G technologies. In addition, the evaluation methodology corresponding to each TPR will be specified to ensure the comparability of the evaluation campaigns conducted by various institutions, companies, and administrations during the period of 2028-2030.

- **ISAC use case and service requirement study in 3GPP:** "Feasibility Study on Integrated Sensing and Communication" in 3GPP System Architecture sub-Group 1 (SA1) Release 19, kicked off in early 2022. The latest study report TR22.837 [3] indicates that 32 use cases for ISAC have been identified, each including details about service flows, functionality requirements, and key performance indicators (KPIs). This study is progressing toward consolidation and further specification in SA1. Discussions are already underway for potential new ISAC use cases in the context of SA1 Release 20.

- **Technical study, specification, and self-evaluation in 3GPP:** Radio access network (RAN) requirements, including channel modelling and evaluation methodology, will first be discussed at the 3GPP RAN level, followed by detailed standardization of RAN-related solutions at the working group level. The latter may include aspects such as the design and selection of sensing waveforms and the collaboration procedures among multiple nodes for joint sensing and feedback. 3GPP RAN will then perform self-evaluation of sensing-

related TPRs under certain test environments defined by ITU-R, submitting the results to ITU-R to check whether IMT-2030 requirements are met.

2 Study of Sensing-related Technical Performance Requirements

2.1 Potential Use Cases for 6G ISAC Evaluation

Future wireless communication systems that integrate sensing capabilities are expected to offer wide-area multidimensional sensing. They will achieve this by providing temporal-spatial information not only about connected devices (such as cell phones) but also about objects in the environment not sending or receiving any wireless signals (such as cars, pedestrians, or buildings). These advanced wireless systems with ISAC capabilities will support various innovative applications [4] such as high-precision localization, high-resolution imaging, gesture/posture recognition, and real-time 3D-mapping — such applications are expected to accelerate growth in numerous industries, including autonomous driving, smart city, industrial automation, digital health, and immersive XR.

Although ISAC was first proposed as a part of 6G [1], 5G evolution, formally known as 5G Advanced, has also foreseen its value. However, by potentially adopting new waveforms, frame structures, spectrum, and network architecture designs, 6G can offer greater design and implementation freedom to support all existing and advanced ISAC use cases.

Sensing Use Cases by Deployment Scenario	Common Functionalities for Sensing Use Cases	Technical Performance Metrics

Smart Transportation
- Intruder detection at highways/railways
- Sensing-assisted automotive maneuvering and navigation
- Sensing at crossroads with obstacles
- Tourist spot traffic management
- Vehicle sensing
- Blind spot detection
- UAV flight trajectory tracing
- UAV collision avoidance
- UAV intrusion detection
- Digital twin

Smart Industry
- AGV detection/tracking
- AMR collision avoidance
- Integrated sensing and positioning
- Collaborative robots
- THz imaging
- Digital twin

Smart Home
- Intruder detection inside rooms
- Intruder detection in surroundings
- Sleep monitoring
- Immersive experience in rooms
- Gesture recognition for application navigation and immersive interaction

Smart City
- Parking space determination
- UAV/vehicle/pedestrian detection near smart grid equipment
- Public safety search and rescue

Common Functionalities for Sensing Use Cases:
- Detection, localization, and tracking
- Environment reconstruction and imaging
- Gesture/Posture recognition
- Health/Safety/Traffic monitoring

Technical Performance Metrics:
- Accuracy of localization/velocity estimation
- Accuracy of reconstruction/imaging
- Missed detection/false alarm probability
- Resolution (range, angle, velocity)
- Sensing service latency

Figure 3 Illustration and analysis of potential 6G ISAC use cases

2.1.1 ISAC Use Cases Studied in 5G-Advanced

Starting in early 2022, 3GPP SA1 Release 19 undertook a feasibility study on ISAC services. The study aims to identify ISAC use cases and requirements for the 3GPP system, enhance the utilization of network infrastructure, offer novel user experience, and support vertical market implementation of ISAC. To date, this study has identified 32 use cases [3], as shown in Figure 3, mainly in the following five deployment scenarios:

- **Smart Transportation:** By leveraging the extensive coverage offered by network infrastructure, ISAC can provide a fresh perspective on the transportation environment. One potential ISAC use case is for intruder detection along highways or railways or at street corners under NLOS, where ISAC can identify potential intruders that may cause fatal incidents. Additionally, thanks to its passive sensing capabilities and wide network coverage, ISAC can be applied to tourist spots and other locations where traditional traffic information may be limited — in this case, it can provide real-time traffic flow information across the country. And through network-assisted sensing, 3GPP-based sensing information obtained via ISAC can help achieve high-quality sensing results for advanced driver assistant system (ADAS) data

fusion by providing supplementary sensing information. This will empower vehicles to navigate challenging driving conditions such as visual obstructions and adverse weather conditions.

- **Smart Industry:** With ISAC functionality, automated guided vehicles (AGVs) can enhance worker interaction and safety. ISAC will play a vital role in proactively detecting and avoiding both static and dynamic areas or obstacles, including identifying and maintaining safe zones around workers or AGVs. Comprehensive sensing coverage, for continuous detection and tracking of targets within the designated sensing service area, is imperative to ensure collision-free movement of AGVs throughout smart factories.

- **Smart Home:** Potential use cases include intruder detection, immersive experiences, contactless sleep monitoring, and gesture recognition. For intruder detection, the presence of an intruder (human or animal) induces variations in the sensing signals relative to certain RF reference characteristics, and these light-of-sight (LOS) and non-light-of-sight (NLOS) variations are utilized to determine the intrusion status. For immersive experiences, ISAC facilitates passive but coarse indoor tracking with respect to a reference point, enhancing the overall user experience of home theaters. And for both contactless sleep monitoring and gesture recognition, finer motion or micro-object tracking is required to enable cost-effective

and less intrusive mechanisms for pattern recognition of human respiration, gestures, and falls.

- **Smart City:** ISAC opens up new possibilities for supporting public services, such as rainfall monitoring, city flood monitoring, parking space detection and tracking, intrusion detection for smart grids or other public facilities, public search and rescue operations, and more. The desired sensing results in smart city applications can vary significantly, ranging from rainfall rates and coarse location of city flooding to the precise positioning of individuals during search and rescue operations.

- **Smart UAV:** ISAC provides passive sensing capabilities to enhance unmanned aerial vehicle (UAV) implementations in a cost-effective manner. In addition to checking adherence to designated UAV flight routes, ISAC functionality can detect and track unauthorized UAVs flying over restricted areas, addressing concerns related to public privacy. Furthermore, ISAC can offer network assistance for UAV collision avoidance and trajectory tracking. And by delivering high-quality sensing results, ISAC can assist UAVs in navigating challenging environments.

2.1.2 Potential New Use Cases in 6G

The advancements in 6G RAN and architecture design will bring significant benefits to the use cases already studied in 5G-Advanced , enhancing the support of ISAC services. In the future, 6G is expected to introduce additional ISAC use cases that feature more stringent yet exciting service requirements and KPIs. Some of these new use cases are described below and shown in Figure 4.

- **Collaborative robots:** The standalone sensors in a robot could be significantly reduced and replaced by the sensing capabilities provided by onboard ISAC modules and network assisted ISAC functionalities in order to facilitate collaboration among robots. The whitepaper in [5] has identified many use cases of collaborative robots for 6G, including collaborative carrying, drone-enabled inspection, interactive education, and healthcare at home.

- **Environment reconstruction for digital twin services:** With the same beamforming structure used for communication and sensing, an ISAC system could reconstruct the environment of a target area and provide real-time digital twining. The study in [6] shows the feasibility of environment reconstruction using the proposed scatter polygon method. It also suggests that the reconstruction accuracy can be improved through the collaboration of multiple sensing transmit/receive points.

- **THz imaging on portable devices:** Recent developments in semiconductor technology have bridged the "THz band gap" and are expected to give rise to various THz sensing applications. The study in [7] demonstrates the feasibility of THz imaging with millimeter-level resolution. Due to its compact form-factor and non-ionizing safety, THz sensing is expected to be integrated into future robotics, mobile devices, and even wearables.

2.2 Sensing-related Technical Performance Metrics

The 32 identified use cases introduced in Section 2.1 are categorized by various deployment scenarios, and common functionalities among use cases can be generalized and illustrated in Figure 2. These commonalities can further drive
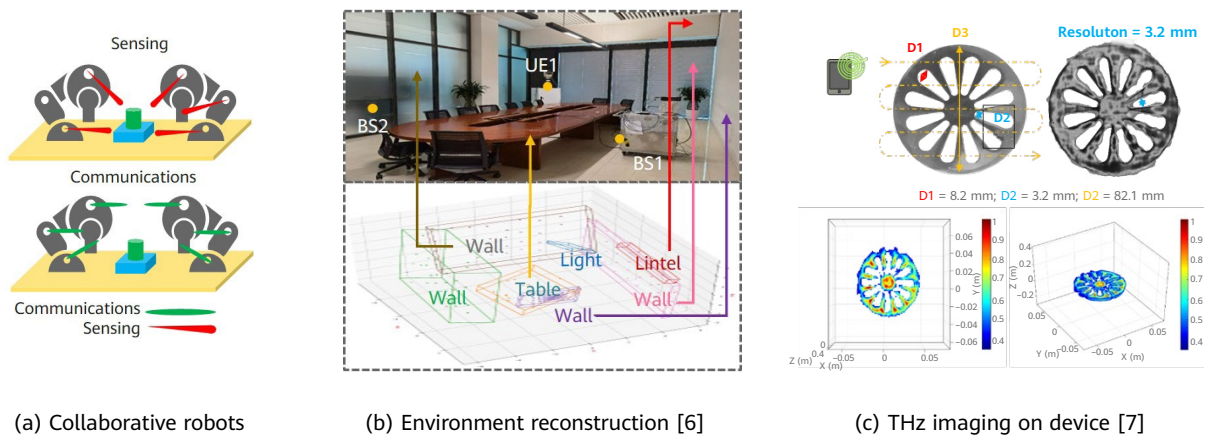

(a) Collaborative robots   (b) Environment reconstruction [6]   (c) THz imaging on device [7]

Figure 4 Illustration of potential new ISAC use cases in 6G

fundamental measurements and performance metrics for each use case. In this section, we analyze the key performance metrics of each use case for further discussion of test environments, with the definition of metrics listed in Table 1.

For object detection, localization, and tracking, it is required to detect whether unwanted or unexpected objects are present in specific areas, and to locate and track these objects based on service requirements. Furthermore, it is essential to detect intruders as early and as accurately as possible. It is also essential to estimate object locations and velocity accurately. As such, for detection, localization, and tracking, metrics like detection probability/false alarm probability, localization/velocity accuracy, and service latency are preferred to be evaluated.

For environment reconstruction and imaging, the main purpose is to reconstruct objects — ranging from large outdoor buildings to small indoor objects — for digital twinning. In this regard, the focus is placed on the detection and localization accuracy of reconstructed scatterers and the similarity of reconstructed and real object shapes. Additionally, the RF images of surrounding environments and objects are captured so that a 3D map can be reconstructed for unknown environments. In smart transportation, self-driving cars can use imaging to "see" around corners. And in smart factories, robots and AGVs can move more freely according to real-time imaging of surrounding environments. Therefore, the main metrics here include the accuracy of object reconstruction and imaging, detection/false alarm probability, and accuracy of localization.

For gesture/posture recognition, a key technology enabling new human-machine interfaces (HMIs), humans can interact with devices by using predefined gestures and actions. As such, the main metrics here include detection, localization, and recognition latency.

2.3 Use Case-oriented Test Environment for 6G ISAC Evaluation

Because mobile communication networks are deployed in many environments (e.g., roads, factories, and homes), the service requirements and network parameters vary in different scenarios. For example, the inter-site distance, number of antennas, frequency bands, and user distribution may differ from one scenario to another. As a result, the TPRs in different scenarios also need to differ. To evaluate communication performance in the 4G and 5G eras, the test environments are therefore defined so that performance metrics are evaluated in different test environments. According to the ITU-R IMT-2020 (5G) evaluation guide [8], each test environment reflects the combination of a geographic environment and a usage scenario and corresponds to a set of evaluation configurations and simulation parameters. From the possible combinations of

Table 1 Definition of sensing-related technical performance metrics

Sensing-related Technical Performance Metrics	Definition
Accuracy of localization	The closeness of the target object's measured position to its true position. This metric can be further derived into horizontal sensing accuracy (i.e., the sensing result error in a 2D reference or horizontal plane) and vertical sensing accuracy (i.e., the sensing result error on the vertical axis or altitude).
Accuracy of velocity estimation	The closeness of the target object's measured velocity to its true velocity.
Accuracy of reconstruction/ imaging	The closeness of the reconstructed scattering points set of the target object to its true set. Potential evaluation mechanisms include Chamfer distance and F-score.
Missed detection probability	The probability of not detecting the presence of an existing target object/environment/event. It applies only to binary sensing results.
False alarm probability	The probability of falsely detecting the presence of a non-existent target object/environment/event. It applies only to binary sensing results.
Sensing resolution	The minimum difference between measurements in different magnitude (e.g., range, angle, and velocity).
Sensing service latency	The time elapsed between an event occurring or sensing triggered by the sensing system and the availability of sensing results on the sensing system.

Mapping of Test Environments and TPRs from 5G to 6G

5G Usage Scenarios	eMBB			mMTC	URLLC
Test Environments	Indoor Hotspot – eMBB	Dense Urban – eMBB	Rural – eMBB	Urban Macro – mMTC	Urban Macro – URLLC
Peak data rate	√	√	√	/	/
User experienced data rate	/	√	/	/	/
Latency	√	√	√	/	√
Connection density	/	/	/	√	/
Reliability	/	/	/	/	√
…	…	…	…	…	…

① Detection, localization, and tracking
② Environment reconstruction and imaging
③ Gesture/Posture recognition

6G Usage Scenarios		ISAC					Immersive Communication	…	
Test Environments	Geographic Environment – Usage Scenario	Dense Urban – ISAC		Indoor Factory – ISAC			Indoor Hotspot – Immersive Communication	…	…
	Use case	①	②	①	②	③			
Sensing-related TPRs	Accuracy of localization	√	√	√	√	√	Communication/positioning/AI-related TPRs		
	Accuracy of velocity estimation	√	√	√	√	√			
	Accuracy of reconstruction/imaging	/	√	/	√	√			
	Missed detection/false alarm probability	√	√	√	√	√			
	Sensing resolution	√	√	√	√	√			
	Sensing service latency	√	√	√	√	√			

Figure 5 Illustration of the test environment-TPR mapping from 5G to 6G

four geographic environments and three usage scenarios, the following five test environments (shown in the upper table in Figure 5) have been chosen for 5G assessment to cover critical aspects in system design and performance:

- **Indoor Hotspot – eMBB:** An indoor isolated environment at offices and/or in shopping malls based on stationary and pedestrian users with very high user density.

- **Dense Urban – eMBB:** An urban environment with high user density and traffic loads focusing on pedestrian and vehicular users.

- **Rural – eMBB:** A rural environment with larger and continuous wide area coverage, supporting pedestrian, vehicular, and high-speed vehicular users.

- **Urban Macro – mMTC:** An urban macro environment targeting continuous coverage and focusing on a high number of connected machine-type devices.

- **Urban Macro – URLLC:** An urban macro environment targeting ultra-reliable and low-latency communications.

According to the study from 3GPP SA1 summarized in Figure 2, sensing use cases span multiple geographic environments and they require significantly different functionalities in the same geographic environment. Taking smart transportation as an example of a geographic environment, detection, localization, and tracking focus on metrics such as detection probability and the accuracy of location and velocity estimation, whereas environment reconstruction and imaging focus on the resolution and accuracy of RF imaging. As a result, the values of the metrics required by each use case might vary significantly.

Therefore, based on the combinations of geographic environments and usage scenarios for defining test environments in IMT-2020 [1], the following format of test environments for IMT-2030 sensing-related TPRs is proposed: "Geographic Environment – Usage Scenario (i.e., ISAC) – Use Case." Examples that adhere to this format include "Indoor Factory – ISAC – Environment Reconstruction" and "Dense Urban – ISAC – Detection, Localization, and Tracking." For each test environment, the evaluation parameters, methodologies, and numerical requirements of metrics shall be specified independently, as shown in Figure 5. And in terms of the geographic environment, it is advisable to select one typical outdoor and one indoor scenario — specifically, "Dense Urban" for outdoor environments like smart transportation, and "Indoor Factory" for indoor environments like smart factory. The main characteristics of each test environment, in terms of propagation conditions, deployment, sensed targets, etc., are summarized in Table 2.

Table 2 Recommended evaluation methodologies for different ISAC test environments

Test Environment (Geographic Environment – ISAC – Use Case)	Definition	Recommended Evaluation Methodologies	Key Elements
Dense Urban – ISAC – Detection, Localization, and Tracking	An urban environment of ISAC focusing on the detection, localization, and tracking of objects (e.g., animals, UAVs, and vehicles).	Full-/Semi-SLS	An object can be represented by one point and dropped into the coverage of networks. Objects' appearance, vanishing, and movement are modeled in simulation.
Dense Urban – ISAC – Environment Reconstruction and Imaging	An urban environment of ISAC focusing on the reconstruction and imaging of objects (e.g., buildings, roads, and vehicles).	Semi-/Full-SLS	An object can be represented by a set of multi-scattering points. A digital map can be imported into simulation to deploy the set of scattering points in the environment if necessary. In simulation, the detection, localization, reconstruction, and/or imaging algorithms can be used to obtain the characteristics of scattering points, e.g., locations, and the characteristics of the object, such as its shape and orientation.
Indoor Factory – ISAC – Detection, Localization, and Tracking	An indoor isolated factory environment of ISAC focusing on the detection, localization, and tracking of objects (e.g., robots, AGVs, and staff).	Full-/Semi-SLS	Same as "Dense Urban – ISAC – Detection, Localization, and Tracking."
Indoor Factory – ISAC – Environment Reconstruction and Imaging	An indoor isolated factory environment of ISAC focusing on the reconstruction and imaging of objects (e.g., walls, machines, robots, and products).	Semi-/Full-SLS	Same as "Dense Urban – ISAC – Environment Reconstruction and Imaging."
Indoor Factory – ISAC – Gesture/Posture Recognition	An indoor isolated factory environment of ISAC focusing on the gesture/posture recognition of staff/visitors.	LLS/Semi-SLS	The gesture/posture recognition algorithm is executed to obtain recognition results.

3 Evaluation Methodologies of Sensing-related Metrics

3.1 Discussion on the Evaluation Methodology of Each Test Environment

ISAC system evaluation may include link-level simulation (LLS) and system-level simulation (SLS). LLS usually involves a small-scale network setup (such as a single pair/group of transmitters/receivers) and mainly evaluates the performance of signal processing algorithms (such as detection and sensing parameter estimation) under given configurations. Conversely, SLS usually involves a larger-scale network setup (such as 21 or 57 cells) and models the impact of traffic models, user scheduling, and dynamics of inter-cell interference. The SLS performance of signal processing is usually abstracted from LLS and imported into the SLS system through a process known as link-to-system interface or physical layer abstraction. SLS can be further classified into Full-SLS and Semi-SLS, both of which adopt a larger-scale network setup but Semi-SLS adopts simplified

multi-drop channel generation and multi-slot duration. An example application of Semi-SLS is the positioning evaluation of UEs in 5G. Table 2 presents the ISAC test environments and suggests the corresponding simulation methodologies and key issues.

For detection, localization, and tracking, the appearance and vanishing of intruders or other objects on the road or in the factory need to be simulated so that system-level network deployment with objects' appearance and vanishing can be modeled in simulation. Because it requires relatively long time slots to evaluate the sensing service latency (e.g., the time between an object's appearance and the sensed time slot) during simulation, Semi-SLS may not meet the evaluation requirements of the service latency metric — in this case, Full-SLS is more suitable. Semi-SLS can be an alternative methodology if objects' dynamic changes are ignored and service latency is assessed through analysis.

For environment reconstruction and imaging, although system-level network deployment is required, the objects to be reconstructed/imaged generally remain unchanged during the simulation. As such, the sensing latency can be obtained through analysis in Semi-SLS if this metric needs to be evaluated. Full-SLS is an alternative methodology but involves a high degree of complexity. It can be still open about whether and how to evaluate the impact of resource sharing or interference between sensing and communication by using Full-SLS.

For gesture/posture recognition, if multiple devices (e.g., base stations or terminals) are involved, system-level network deployment can be considered. In this case, evaluation can be performed by using imaging and/or gesture/posture recognition algorithms. For single-device cases, LLS also supports the evaluation of gesture/posture recognition.

The exemplary procedures of Full-/Semi-SLS and LLS are further discussed in the following sections, each procedure with a typical sensing use case.

3.2 Exemplary Full- and Semi-System-Level Simulation Procedures

For both Full-SLS and Semi-SLS, system-level network and object deployment are important features that distinguish SLS from LLS. As such, the system-level deployment of base stations and objects in the prioritized geographic environments is described first.

For smart transportation (i.e., dense urban environment),

several base stations (e.g., 7 or 19) with three sectors on each are deployed to form a hexagonal network and cover full horizontal directions — this is commonly used in 5G evaluation. Target objects can be vehicles, pedestrians/animals, or flying objects. Certain target areas like "road" and "shoulder" with certain width can be deployed in the environment, and the proportions of objects dropped in different areas can be preconfigured according to the real-world scenario. For example, all vehicles are deployed in road areas; and 80% of pedestrians/animals are deployed in shoulder areas, 1% are deployed in road areas as important objects, and 19% are deployed in other areas. Figure 6 illustrates an example deployment.
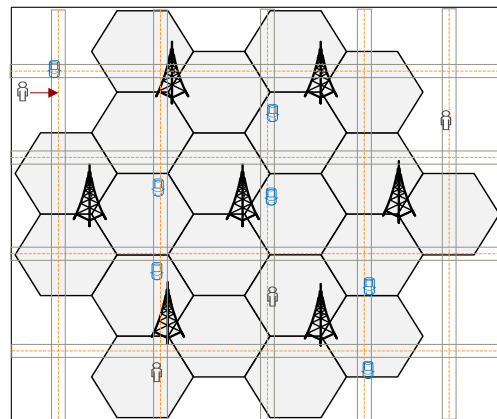
Figure 6 Illustration of network and object deployment in a dense urban environment

For smart industry (i.e., indoor factory environment), the deployment scenario can refer to the indoor factory deployment in 5G evaluation. For example, multiple base stations are deployed on the roof of the factory, as shown in Figure 7. The dropped objects may include AGVs, machines, staff, and products. When deploying different types of objects, it is noted not to deploy the objects at duplicate locations.

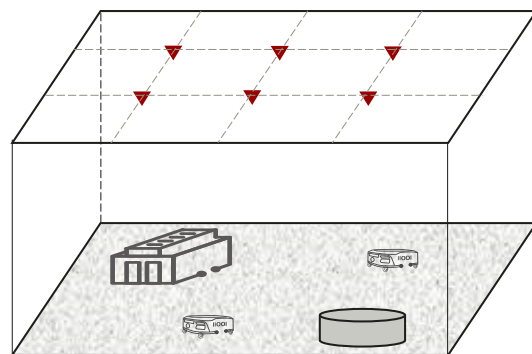**Figure 7** Illustration of network and object deployment in an indoor factory environment

Full-SLS procedure

Configure parameters and network layout.

Initialize the sensing channel.

Schedule ISAC resources.

Drop objects and update the sensing channel.

Receive sensing signals.

Estimate sensing parameters (e.g., using physical abstraction).

Detect and localize objects.

Collect overall statistics on metrics.

Subsequent slots

(a)

Semi-SLS procedure

Configure parameters and network layout.

Drop objects and generate the sensing channel.

Schedule ISAC resources.

Receive sensing signals.

Estimate sensing parameters.

Reconstruct and image objects.

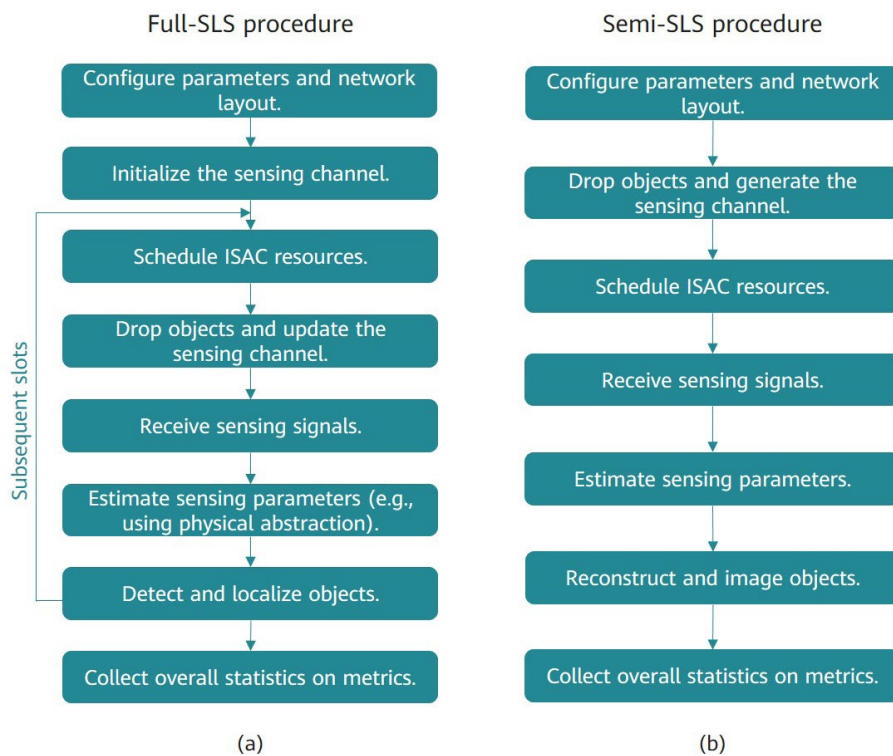Collect overall statistics on metrics.

(b)

Figure 8 Illustration of the typical (a) Full-SLS and (b) Semi-SLS procedures for ISAC evaluation

The typical Full-SLS and Semi-SLS procedures (which share most key steps) for ISAC evaluation are illustrated in Figure 8. In the following, the steps of Full-SLS for detection, localization, and tracking use cases are introduced as an example, with some specific remarks on Semi-SLS for environment reconstruction and imaging.

- **Step 1 Configure parameters and network layout.** For the application of smart transportation, which primarily operates in outdoor environments, a cellular layout resembling 3GPP UMa or UMi (shown in Figure 6) can be considered. ISAC parameters, such as sensing antenna array size, frequency band, and sensing mode (monostatic, bistatic, or multi-static mode, or a combination thereof) need to be thoroughly discussed and agreed upon by 3GPP. It is important to note that each sensing mode may have its own preferred parameter sets due to implementation restrictions.

- **Step 2 Initialize the sensing channel.** For the use case of detection, localization, and tracking, base station-based monostatic, bistatic, and multi-static sensing modes can be considered. Consequently, the sensing channel between each pair of base stations and the monostatic sensing channel for each base station need to be generated in this step.

- **Step 3 Schedule ISAC resources.** At the beginning of the ith slot, the ISAC scheduler updates the state of the

sensing buffer. This buffer includes information about detected targets/intruders from previous slots and hypothetical targets to be sensed in the ith slot. The determination of hypothetical targets is subject to the ISAC scheduling design. Each set of coordinated ISAC transmit/receive groups can schedule time, frequency, power, and antenna resources based on the status of the sensing buffer.

- **Step 4 Drop objects and update the sensing channel.** In the ith slot, the position and velocity of the detected objects are updated and new objects (e.g., intruders) may be dropped according to a predefined generation model. The sensing channel will be updated based on all objects (regardless of whether they are detected).

  Δ In Semi-SLS, simplified dropping of dynamic objects with limited channel updates is expected, and repeated simulation for subsequent slots may not be needed. Therefore, to implement object dropping and sensing channel generation only once, after the network deployment, Step 4 can be merged into Step 2.

- **Step 5 Receive sensing signals.** Taking into account both Step 3 and Step 4 in the ith slot, the sensing signals at each receiver are generated. Potential interference from data communication, self-interference, intra-cell interference, and inter-cell interference needs to be considered in the received signals.

- **Step 6 Estimate sensing parameters.** In this step, the sensing channel parameters — for example, power, delay, angle of arrival (AoA), and Doppler — for the dominant paths are estimated by using algorithms such as MUSIC [9]. As mentioned earlier, certain forms of link-to-system mappings are expected to reduce the simulation consumption. One possible form is to pre-derive and store the relationships between estimation accuracy and signal to interference plus noise ratio (SINR). The methodologies for physical abstraction should be further discussed within the framework of 3GPP.

  Δ Alternatively, dedicated sensing algorithms can directly give rise to targeted sensing results without link-to-system mapping in this step. This is especially true in Semi-SLS. For example, detection/localization of intruders, or object imaging by detecting and locating multiple scatterers, can be implemented but potentially at the cost of increasing the simulation time.

- **Step 7 Detect and localize objects.** Using the derived sensing channel parameters, a dedicated detection algorithm is employed to detect individual intruders by each cell or jointly across cells. This process may involve setting dynamic thresholds for power and Doppler to isolate the intruders' profile from the surrounding environment. For each identified intruder, a follow-up localization and tracking algorithm estimates its position and velocity. Monostatic, bistatic, and multi-static localization can be applied separately or together — each one has both advantages and limitations in terms of performance benefits and implementation constraints.

  Δ Alternatively, for object reconstruction and imaging, the reconstruction algorithm may be combined with the estimation of sensing parameters in Step 6.

- **Step 8 Repeat Steps 3–7 for each subsequent slot.** For each subsequent slot, Steps 3–7 are repeated until the end of the preconfigured simulation time.

  Δ In Semi-SLS, this step is not required because usually only limited slots are simulated.

3.3 Exemplary Link-Level Simulation Procedure

In ISAC evaluation, LLS typically serves two purposes:

- To achieve a link-to-system interface (referred to as physical layer abstraction)

- To evaluate certain use cases such as gesture/posture recognition

These two purposes can share a similar LLS procedure, as follows:

- **Step 1 Configure parameters.** Unlike SLS, LLS involves only one or a few sites, meaning that network deployment is generally not required.

- **Step 2 Generate the sensing channel based on object features.** For example, in the gesture/posture recognition use case, a sensing channel is generated based on the channel features predefined for each gesture/posture pattern.

- **Step 3 Receive sensing signals (and preconfigure SNRs).** LLS typically requires observing performance under different signal-to-noise ratios (SNRs). Therefore, in this step, noises with a certain power level are added to the received signals to achieve the preconfigured SNRs.

- **Step 4 Estimate sensing parameters.** LLS runs the sensing algorithm to estimate the sensing parameters directly. If the purpose of LLS is to obtain the physical layer abstraction of parameter estimation, statistics can be collected after this step.

- **Step 5 Recognize the gesture/posture of objects.** Based on the results of sensing parameter estimation obtained in Step 4, the gesture/posture recognition algorithm, e.g., based on AI, is executed to obtain recognition results. These results provide supplementary information for the overall statistics of TPRs.
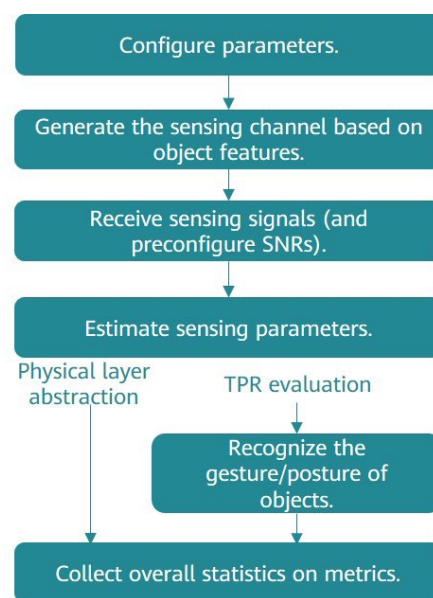

**Figure 9** Typical LLS procedure for ISAC evaluation

- **Step 6 Collect overall statistics on metrics.** For physical layer abstraction, the impact of SNRs on the estimation accuracy of sensing parameters such as latency, angle, and Doppler can be derived from statistics. For gesture/posture recognition, this step also involves collecting statistics on the TPRs listed in Figure 5.

4 Conclusion

Recognized as one of the six usage scenarios of 6G by ITU-R, ISAC presents a broad range of novel applications across various domains and elevates safety requirements, convenience, and efficiency in our daily life. The next milestone of ITU-R 6G standardization centers on TPRs and evaluation methodologies of the six usage scenarios including ISAC, which is our aspiration to contribute to. In this paper, we explore the TPRs motivated by the potential ISAC use cases for 6G, and propose the evaluation methodology for each proposed use case-oriented test environment. We also share some initial thoughts on SLS and LLS — these are among the fundamental tasks to be addressed by 3GPP RAN standardization studies in the future.

References

[1] Wen Tong and Peiying Zhu, "6G: The next horizon – From connected people and things to connected intelligence," Cambridge University Press, May 2021.

[2] Draft new Recommendation ITU-R M. [IMT. FRAMEWORK FOR 2030 AND BEYOND], "Framework and overall objectives of the future development of IMT for 2030 and beyond," Source: ITU-R internal document, June 2023. [Online]. Available: https://www.itu.int/md/meetingdoc. asp?lang=en&parent=R19-SG05-C-0131

[3] 3GPP Technical Report, TR 22.837 V2.0.0, "Feasibility study on integrated sensing and communication," 15 June 2023.

[4] D. K. P. Tan, J. He, Y. Li, A. Bayesteh, Y. Chen, P. Zhu and W. Tong, "Integrated sensing and communication in 6G: Motivations, use cases, requirements, challenges and future directions," in 1st IEEE International Online Symposium on JC&S, 23–24 February, 2021.

[5] one6G, whitepaper, "6G & robotics, use cases and potential service requirement," June 2023. Available online: https://one6g.org/resources/publications/

[6] Z. Zhou *et al*., "6G integrated sensing and communication – Sensing-assisted environmental reconstruction and communication," ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023.

[7] O. Li, *et al*., "Integrated sensing and communication in 6G: A prototype of high-resolution THz sensing on portable device," in European Conference on Networks and Communications (EuCNC), 8–11 June, 2021.

[8] ITU, "Guidelines for evaluation of radio interface technologies for IMT-2020: M. 2412-0." (2017).

[9] Bazzi, Ahmad, Dirk TM Slock, and Lisa Meilhac, "On joint angle and delay estimation in the presence of local scattering," IEEE International Conference on Communications Workshops (ICC), 2016.
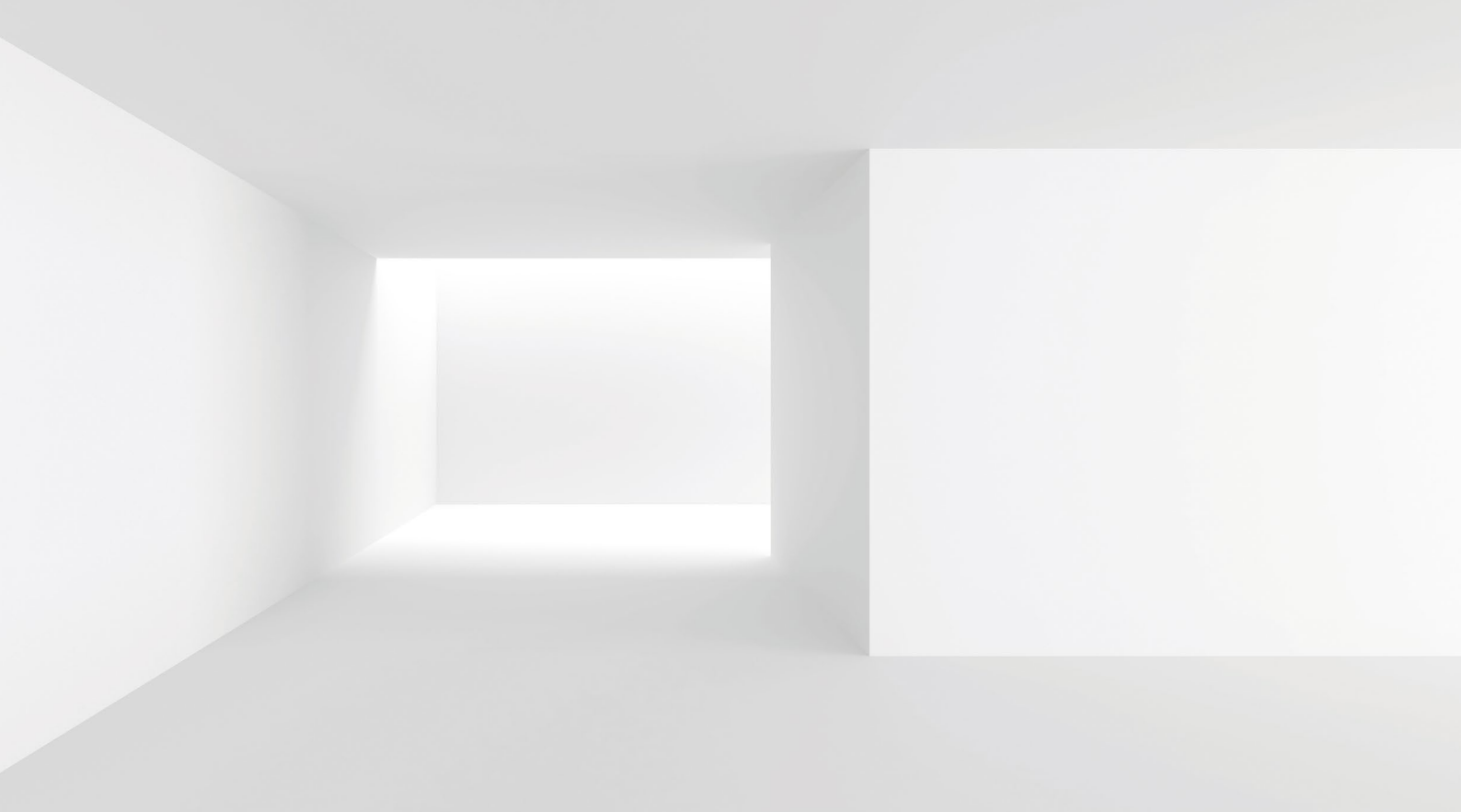
Novel Hybrid Channel Modeling on 6G Integrated Sensing and Communication

Yi Chen, Wenfei Yang, Ziming Yu
Wireless Technology Lab

Abstract

Integrated sensing and communication (ISAC), which can be used for localization, identification, environment reconstruction, and communication assistance purposes, is one of the six usage scenarios specified in ITU-R's IMT-2030 (6G) framework. ISAC performance evaluation depends critically on channel modeling, but because traditional models are oriented only to communication, such models cannot meet the requirements imposed by ISAC. Consequently, we need to explore a new channel model for ISAC. In this paper, we first introduce the use cases of ISAC and the corresponding requirements for channel modeling. Then, we propose a novel channel modeling method that incorporates statistical and deterministic modeling. This modeling method independently considers sensing targets and background channels, and adopts the multi-scattering center theory to model the target's electromagnetic features. We verify the accuracy and effectiveness of the proposed model through simulation and measurement.

Keywords

ISAC, multi-scattering center, channel model

1 Introduction

Integrated sensing and communication (ISAC) is a key technology of the 6G communications network [1, 2]. It enables sensing and communication to be performed at the same time using the same hardware rather than being conducted separately. Two key benefits of this approach are improved spectral efficiency and reduced power consumption. Because we need to model the ISAC channel in order to evaluate the performance of ISAC in different cases, research on the ISAC channel is extremely important.

Given that a channel model should support the evaluation of different use cases, the ISAC channel model needs to take into account the specific use cases and corresponding evaluation indicators in order to abstract supported features. ISAC use cases can be roughly divided into three categories: localization, detection and tracking, and environment reconstruction and pattern recognition. In terms of the channel, performance evaluation of these use cases poses new requirements compared with traditional communications use cases. This means that we need to explore a new architecture and features of the ISAC channel model and new modeling methods. From the architecture perspective, the traditional communications channel only considers the scatterers that generate communications multipaths in the environment — this is insufficient for ISAC evaluation, in which the ISAC channel should also consider sensing targets (which vary with use cases) in the environment. For example, in environment reconstruction, the sensing target can be a building, the ground, or even both. And in localization and tracking use cases, the preceding objects may be modeled as the background. The architecture of the ISAC channel should therefore be more flexible given that there are more components in the ISAC channel than in the traditional communications channel.

Similar to the methods for modeling the communications channel, those for modeling the ISAC channel can be classified into deterministic, statistical, and hybrid methods [1]. Deterministic modeling determines the ISAC channel based on the locations of preset transmitters and receivers. Examples of such methods include full-wave simulation, ray tracing, and channel measurement. Statistical modeling models the parameters of the ISAC channel as random distribution, which usually consists of large-scale parameters (such as path loss, K factor, delay spread, and angular spread) and small-scale parameters (such as multipath delay, angle, and power). Hybrid modeling [3], as the name suggests, is a combination of the other two models,

leveraging the advantages of both. In the propagation channel, the deterministic method is used to model the major multipath component, and the statistical method is used to model the other multipath component. The major channel component in an ISAC channel model is related to the sensing target — we believe that the deterministic method can be used to model this component. Given the variety of ISAC use cases, there are differences in the major component in the ISAC channel. For instance, in localization and identification, humans and vehicles are sensing targets, whereas in environment reconstruction and sensing-assisted communication, buildings and the ground are multipath components in the propagation environment (consequently, they should also be regarded as sensing targets).

Because the ISAC channel focuses more on the scattering features of the target, it is important to model the target in the ISAC model. In traditional ISAC model research [4], a common practice is to model the radar cross section (RCS) of a sensing target as a scattering point in order to represent the entire target. This is referred to as the single-scattering point model, in which there are three methods for describing the RCS of a sensing target: (1) fixed value; (2) statistical modeling; and (3) deterministic modeling. The first method regards the sensing target's RCS as a fixed value that is irrelevant to the angle. The second considers that the RCS follows a random distribution. And the third calculates the RCS based on simulation or measurement, taking into account the impact of shape, frequency, and incident angle. Modeling a shaped object as a single scattering point ignores the spatial and geometric features of the object. Therefore, the single-scattering point model only supports sensing cases such as localization and identification. To address this issue, we propose the multi-scattering center model. In this model, a sensing target is represented by multiple scattering centers. These scattering centers are obtained by measuring the object's scattering field and simulating its electromagnetic features. In contrast with its single counterpart, the multi-scattering center model can accurately restore the RCS and geometric features of a sensing target, making it suitable for use cases of imaging, environment reconstruction, and pattern recognition.

Section 2 briefly discusses ISAC use cases and their categories, and identifies the features that the ISAC channel model needs to support in order to meet the requirements of the use cases. Section 3 proposes a hybrid ISAC channel model and introduces the components and modeling methods of the proposed model. In this model, the components are divided into two parts: target-related and background components.

Target-related components are described by multiple scattering centers, and the background component is modeled using a statistical method. Section 4 describes the measurement of ISAC channels.

2 ISAC Channel Model Requirements and Challenges

Channel modeling must consider the requirements for evaluating the performance of different use cases. Integrated Sensing and Communication Technical Report, compiled by IMT-2030 (6G) ISAC Task Group [4], and 3GPP SA1 TR22.837 [5] discussed the use cases of ISAC. The use cases identified include intrusion detection [5], vehicle detection and tracking [5], meteorological monitoring [5], health monitoring [5], high-accuracy localization and tracking [5, 6], environment reconstruction [5, 6], augmented human sense [6], and gesture and activity recognition [6]. These cases can be classified into three categories:

- Localization, detection, and tracing. This category includes smart transportation, intrusion detection in smart manufacturing, and vehicle and robot localization and tracking. Such services focus on the probability of detection (PoD), probability of false alarm (PoFA), localization accuracy, speed estimation error, target density, refresh rate, and sensing delay.

- Environment reconstruction and imaging. This category is commonly used in smart transportation and manufacturing. The evaluation of these cases focuses on the target reconstruction rate, reconstruction distance error, and direction error.

- Pattern recognition. This category mainly involves posture and gesture recognition, such as breathing, heartbeat, and fall recognition in smart home or health care scenarios.

Currently, the statistical channel model proposed in 3GPP TR38.901 is used for simulation of localization applications defined by 3GPP [8]. 6G ISAC raises requirements for channel modeling from the following perspectives:

- Passive targets. The sensing targets defined in 3GPP R18 only involve terminals that access the wireless network — these objects are active targets. In the future 6G network, base stations should be able to detect passive objects such as animals that stray into roads and unmanned aerial vehicles (UAVs) that fly into restricted airspace.

- Different sensing mode. In active localization, the wireless channel is established based on the wireless transmission link between a base station and a terminal, and is referred to as a bistatic channel. Conversely, in passive sensing, the base station needs to receive the signals it sends so that it can obtain information about a target and localize and track the target — this is performed via a monostatic channel.

- Richer target features. The change of sensing mode gives rise to sensing requirements for modeling the scattering of the target's physical features. For example, in vehicle intrusion sensing simulation by a base station, we expect that the channel model should accurately depict the typical RCS model of vehicles. The traditional cluster-based channel model fails to model the electromagnetic features of sensing targets, something that the ISAC channel model must be able to do in order to restore the target's shape and identify its material.

- Scatterer mapping. To improve the accuracy of localization and tracking, the transceiver must extract information about the channel's spatial geometric propagation path and reflection points. Specifically, it is necessary to extract the spatial consistency of the reflection path, physical position, and electromagnetic scattering features of the moving target. The conventional cluster-based channel model uses mathematical techniques to randomly depict features such as the channel angle, delay, and power in a statistical manner in order to form the cluster and scale parameter distribution. However, this distribution has no geometric or physical meaning.

- Predefined information. Before the evaluation of ISAC applications can be performed, certain information about the target (e.g., the target type, vehicles, pedestrians, and UAVs) and environment (e.g., crossroads) must be defined.

- Micro-Doppler signature. Posture recognition relies on the modeling of channel changes based on micromotions of the target. For example, the rise and fall of a human's chest during breathing can lead to micro-Doppler signature changes.

- Model architecture. Because the statistical channel model cannot precisely reflect the environment, we need novel channel modeling methods and model architectures. A hybrid channel modeling method integrating both statistical and deterministic techniques and a deterministic channel modeling method based on ray tracing and electromagnetism are candidate solutions.

3 Channel Modeling Methodology

Based on the stochastic communications channel model defined in 3GPP TR38.901 [8], we propose a novel channel model that consists of target-related components and a background component. Figure 1a and Figure 1b show the models of a monostatic and a bistatic sensing channel, respectively. The target-related components include all paths involving the target. These components are cascaded by two segments: the transmitter-target segment, where the target receives the paths from the transmitter; and the target-receiver segment, where the target sends the paths to the receiver. The background component contains paths from the background environment — these paths do not change with the the number or location of targets. If there are multiple sensing targets, the target-related component is generated for each sensing target, respectively, and then

Monostatic sensing channel

(a) Monostatic sensing channel
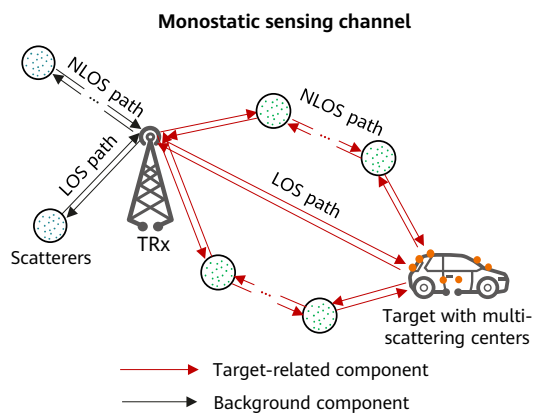
Bistatic sensing channel
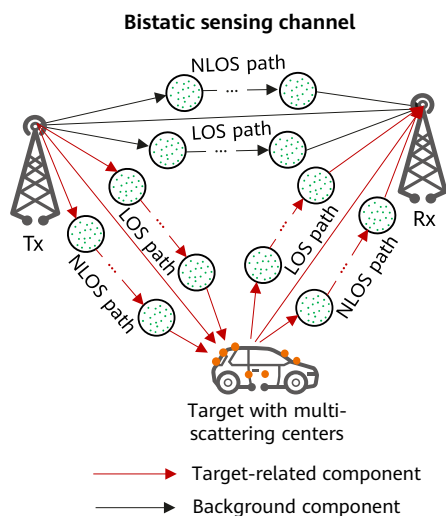
(b) Bistatic sensing channel

Figure 1 ISAC channel models

the background component of the sensing environment is generated. We can then combine the target-related and the background components to obtain the sensing channel. The method for generating the background component and each segment of target-related components of the sensing channel is similar to the communications channel generation procedure specified in 3GPP TR38.901. The sensing channel and communications channel are associated by sharing some or all of the statistical parameters. If a user in the communications channel is a sensing target in the sensing channel, a segment of the target-related components in the sensing channel may reuse the statistical parameters of the communications channel. And if the communications user is a sensing transmitter or receiver in the bistatic sensing channel, the communications channel also functions as the sensing channel.

3.1 Background Component Generation

In a monostatic sensing scenario, the transmitter and receiver are at the same location. This differs from conventional communications links, in which data is sent and received by different devices. The background component of the monostatic sensing channel depends on the location of the transceiver in the propagation environment, meaning that statistical models (e.g., the path loss model and the line-of-sight probability model) and parameters used for the communications channel are no longer applicable. So that we can describe the loss in the environment in the monostatic sensing channel, a total power gain coefficient irrelevant to the Tx-Rx distance needs to be defined during model parameterization. For other reusable channel parameters (e.g., the delay spread and angular spread), model parameterization should be completed for the monostatic background component according to channel measurement results.

In a bistatic sensing scenario, the background component consists of the interference power — transmitted through a bistatic link — directly received by a transmitter without passing through a target. The generation process is the same as that used to generate the communications channel in 3GPP TR38.901. In addition to transmitters and receivers with height deviation (as is the case in traditional communications links), paths between base stations exist in a bistatic sensing scenario — this is proof of the diversity of sensing modes and scenarios. Figure 1b shows an example of two base stations jointly sensing a vehicle. In that example, the background

component contains the paths between the base stations. Because the transmitter and receiver are at a similar height, we can refer to [8] to define statistical models for certain channel parameters, including large-scale parameters such as the delay spread, angular spread, and shadow fading, and small-scale parameters such as the path-cluster power, delay, and phase.

3.2 Target-related Component Generation

As previously mentioned, the target-related components consist of two segments. When we cascade these two segments, it is necessary to consider what impact the sensing target's features will have on the target-related component. The generation of the target-related component can be described as follows: The large- and small-scale channel parameters of the transmitter-target and the target-receiver channel segments are generated first. Then, the target is represented by a multi-scattering center model (which has an essential impact on the path loss model, and details will be provided in the next subsection). After the target's scattering field is calculated based on the incident and scattering angles of the two segments, the two segments are cascaded to form the target-related component of the sensing channel. In a monostatic sensing scenario, the large- and small-scale parameters of the two segments are the same. [9] and [10] provide details about channel parameters in scenarios involving a UAV and vehicle, respectively. The next subsection uses a car as an example to illustrate the multi-scattering center model.

3.3 Multi-Scattering Center-based Modeling for Sensing Targets

The multi-scattering center model regards the electromagnetic scattering response of a sensing target to be equivalent to the sum of scattering responses of multiple independent scattering centers. The locations of these scattering centers and electromagnetic parameters can be combined to represent the target's scattering features. And based on the scattering centers, the target's geometric structure can be obtained. Applying the multi-scattering center model to the ISAC channel model will simplify the modeling of sensing targets by representing the complex geometric structure and electromagnetic features of sensing targets with a small number of scattering centers.

The scattering field $E(f, \varphi, \theta)$ of the sensing target can be expressed as a sum of scattering fields of multiple scattering centers as:

$$E(f, \varphi, \theta) = \sum_{i=1}^{I} E_i(f, \varphi, \theta)$$

where f, φ, θ are frequency, azimuth, and pitch angle, respectively. I denotes the number of scattering centers, and $E_i(f, \varphi, \theta)$ denotes the scattering field of the i-th scattering center. This equation only considers the scattering field of echoes in a single direction in a monostatic sensing scenario (i.e., the scattering field when the incident and scattering directions are the same). In a bistatic channel, the electromagnetic waves travel to the sensing target in one direction and are scattered in another direction. According to research [11] by R.E. Kell, a bistatic scattering field can be approximated by the aforementioned monostatic scattering field. This means that we can use the multi-scattering center model to model both monostatic and bistatic channels during ISAC channel modeling.

In order to verify the multi-scattering center model, an electromagnetic simulator is used to simulate the rear part of a car. We obtained the monostatic scattering field, and extracted scattering centers from it (Figure 2). The simulation frequency range is 10 GHz to 11 GHz, at a step of 2 MHz. The scanning angle range is from 0° to 180°, at a step of 2.5°. The scanning begins from the X-axis direction (0°) in Figure 2. The simulation adopts the configuration of a co-located transmitter and receiver located on a circle with a radius of 20 meters and the car as the center. In the figure, each solid dot represents a scattering center — there are 100 such centers in total. The color of a dot indicates its field strength. The larger the dot, the higher the strength. It
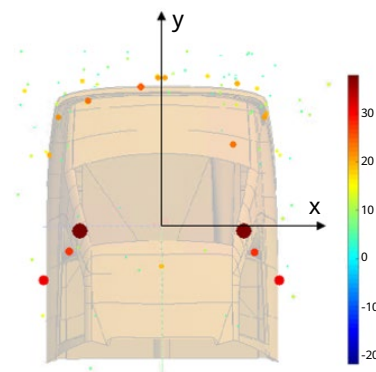

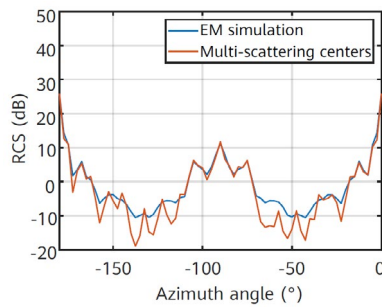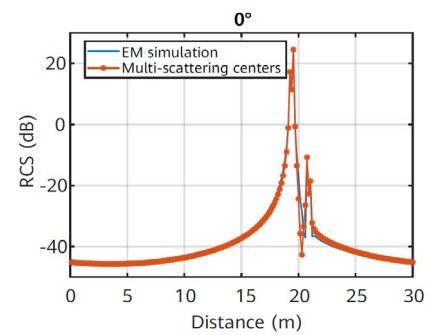**Figure 2** Extracted scattering centers of a car

Figure 3 Scattering field of a car versus the angle and the scattering field reconstructed from five scattering centers

can be observed that most scattering centers are near the car's fenders and two rear windows, and all the scattering centers roughly envelop the car.

Figure 3 shows the function of the car's scattering field with respect to the angle and the scattering field reconstructed from the extracted scattering centers. For ease of understanding, the strength of the scattering field is converted to the RCS of the car. In the proposed model, the strength of each scattering center changes with the angle. Therefore, when reconstructing the scattering field at each angle, five multi-scattering centers with the highest field strength are selected from the 100 centers, and the remaining ones are considered invisible. By comparing the simulation results with the reconstruction results, it can be seen that the scattering field reconstructed based on five scattering centers is consistent with that obtained by electromagnetic simulation in different incident directions. We calculated that the average energy reconstruction rate using five scattering centers is 93.41%. (The energy construction rate at each incident angle is calculated as: Summed energy of the five scattering centers/Summed energy of all scattering centers.) We can therefore conclude that, with the proposed multi-scattering center model, a small number of scattering centers are enough to accurately describe the scattering field features of the sensing target.

Figure 4 shows the simulation results of the car's scattering field with respect to the distance and the scattering fields reconstructed from five centers at different incident angles (0°, 45°, and 90°). In the simulation, the distance between the car and the observation point is set to 20 meters. As shown in the figure, the RCS impulses are concentrated around 20 meters. The scattering field of the car contains multiple impulse responses at each observation angle, meaning that the car's scattering field is composed of different scattering field structures and cannot be considered equivalent to a single scattering point.

(a) 0°
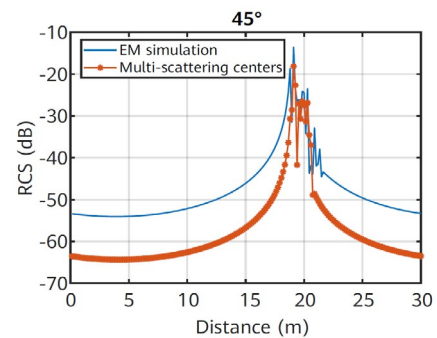
(b) 45°
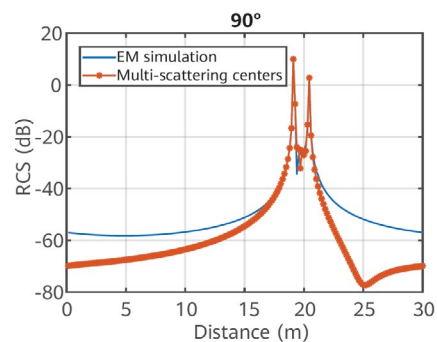
(c) 90°

Figure 4 Scattering field of a car versus distance and the scattering field reconstructed from five scattering centers

Figure 4 proves that the multi-scattering center model can restore the car's scattering field response in the distance domain at different incident angles. Whereas the single-scattering center model can only describe the relationship between the angle and scattering field of the target, the proposed multi-scattering center model can restore more details of the target's scattering field. Our simulation results show that the multi-scattering center model can describe the geometric and scattering features of the sensing target in terms of angle and distance. This demonstrates that the multi-scattering center model is suitable for localization, identification, and environment reconstruction use cases of ISAC.

4 Measurements

By measuring the ISAC channel, we can obtain background environment-related statistical parameters required for establishing the ISAC channel model. The measurement results can be used to analyze ISAC channel features, extract channel model parameters, and verify the channel model. Monostatic and bistatic ISAC channels of 10 GHz and 140 GHz are measured. Table 1 lists the measurement parameters.

4.1 Indoor ISAC Channel Measurement

The indoor ISAC channels are measured in a conference room. In the conference room, we set up a transmitter (Tx) 2 meters above the floor and two transceivers (TRx) 1.4 meters above the floor. The Tx only transmits signals, whereas the TRxs can both transmit and receive signals, as shown in Figure 5. In this scenario, we measured the Tx-TRx bistatic channel and the TRx monostatic channel. For the bistatic channel, the Tx transmits signals towards TRx1 and TRx2. Each TRx, as a receiver, can receive incoming signals in the horizontal and elevation directions after mechanical adjustment. Figure 6 shows the angle-delay-power spectrum of both the monostatic channel and the bistatic channel in the conference room scenario. Given that only a small
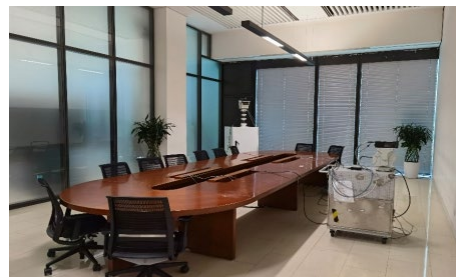
(a) Conference room
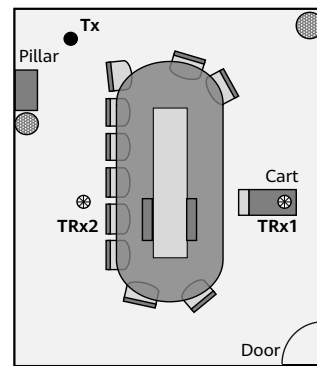
(b) Conference room measurement configuration

Figure 5 Indoor measurement in a conference room

number of paths exist in each cluster of the bistatic channel, the channel is sparse. The paths in the monostatic channel are mostly echoes of the propagation environment and are more abundant.

Table 1 ISAC channel measurement parameters

Frequency Band	10 GHz		140 GHz	
Mode	Monostatic	Bistatic	Monostatic	Bistatic
Frequency	10–12 GHz	10–12 GHz	130–133 GHz	130–133 GHz
Scan Step	1 MHz	1 MHz	1 MHz	1 MHz
Transmit Antenna Gain	20 dBi	20 dBi	25 dBi	15 dBi
Receive Antenna Gain	20 dBi	20 dBi	25 dBi	25 dBi
Horizontal Angle	0°:10°:360°			
Pitch Angle	−30°:10°:30°			
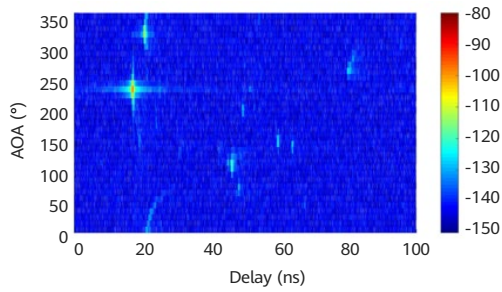
(a) Tx-TRx1 bistatic channel
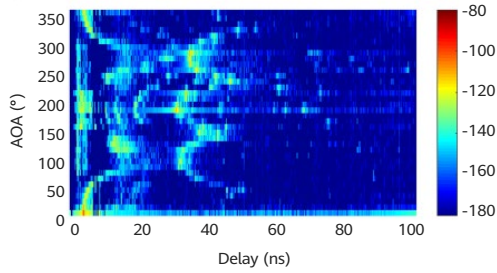
(b) TRx1 monostatic channel

Figure 6 Indoor measurement results in a conference room
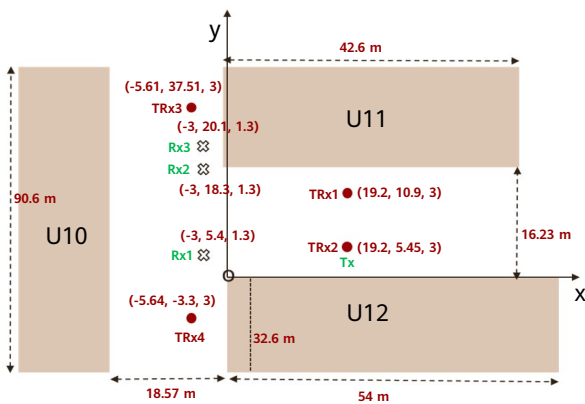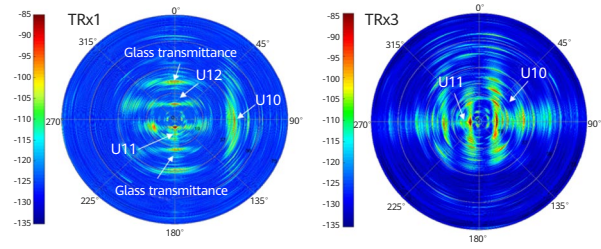
Figure 8 Measurement results of monostatic channels at an outdoor UMi site

4.2 Outdoor ISAC Channel Measurement

Outdoor ISAC channels are measured in an urban micro (UMi) scenario, which contains both monostatic and bistatic channels. Figure 8 shows the angle-delay-power spectrum in polar coordinates of the monostatic channels of TRx1 and TRx3 in the UMi scenario. The multipath reflections of buildings around the TRxs can be observed in Figure 8, proving that the geometric features of the propagation environment can be restored in the monostatic channel. These features can provide reference for communications channel prediction.

5 Conclusion

In this paper, we identified the requirements for channel modeling oriented to the application scenarios of three ISAC use case categories (i.e., localization and identification, environment reconstruction and imaging, and pattern recognition). Based on the identified requirements, we proposed a novel method that integrates statistical and deterministic techniques to model ISAC channels. And leveraging the communication model defined in 3GPP TR38.901, we proposed an ISAC channel model composed of the target-related and background components. Drawing on the multi-scattering center theory, we modeled the electromagnetic features of sensing targets. Furthermore, we measured monostatic and bistatic channels, obtaining results that prove the feasibility of parameterizing the proposed model. We also verified the effectiveness and accuracy of the multi-scattering center model using electromagnetic simulation. The simulation results show that the proposed ISAC channel model is compatible with the widely used 3GPP TR38.901 communications channel model, and can meet the requirements of ISAC use cases for channel modeling. Therefore, the proposed model can be utilized in performance evaluation and optimization for ISAC systems.

(a) UMi site

(b) UMi measurement configuration

Figure 7 Measurement at an outdoor site

References

[1] A. Bayesteh, J. He, Y. Chen, P. Zhu, J. Ma, A. W. Shaban, Z. Yu, Y. Zhang, Z. Zhou, and G. Wang, "Integrated sensing and communication (ISAC) — From concept to practice," Huawei Research, no. 2, 2022.

[2] G. Wang, H. Gu, X. Li, Z. Yu, O. Li, Q. Liu, K. Zeng, J. He, Y. Chen, J. Lu, W. Tong, and D. Wessel, "Terahertz sensing and communication towards future intelligence connected networks," Huawei Research, no. 2, 2022.

[3] Y. Chen, Z. Yu, J. He, *et al.* "A scatterer-based hybrid channel model for integrated sensing and communications (ISAC)," in *Proc. of IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2023.

[4] IMT-2030 (6G) Promotion Group, "Integrated sensing and communication technical report," Technical Report 1, 2021. [Online]. Available: https://www.imt2030.org.cn/html/default/zhongwen/chengguofabu/yanjiubaogao/list-2.html?index=2

[5] 3GPP, "Study on integrated sensing and communication," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 22.837, 2022. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4044

[6] Huawei Technologies, "6G: The Next Horizon," https://www-file.huawei.com/-/media/corp2020/pdf/tech-insights/1/6g-white-paper-en.pdf, 2021

[7] IMT-2020 (5G) Task Group, "Report on 5G-Advanced ISAC simulation," Technical Report, 2023.

[8] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.901, 2017. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3173

[9] 3GPP, "Enhanced LTE support for aerial vehicles," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 36.777, 2018. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3231

[10] 3GPP, "Study on evaluation methodology of new vehicle-to-everything (V2X) use cases for LTE and NR," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 37.885, 2019. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3209

[11] R. Kell, "On the derivation of bistatic RCS from monostatic measurements," in *Proceedings of the IEEE*, vol. 53, no. 8, pp. 983–988. 1965.

Candidate Technologies of Future Air Interfaces for Integrated Sensing and Communication

Xia Shen, Jiamo Jiang, Xiaoyan Xu[*]

Mobile Communications Innovation Center, China Academy of Information and Communications Technology

Abstract

The revolutionary concept of integrated sensing and communication (ISAC) integrates communication and sensing functionalities through a shared communications infrastructure, offering new avenues for advancement in 5G-Advanced (5G-A) and 6G communication systems. This paper investigates emerging technologies that could define future air interfaces for ISAC systems, covering key aspects like frequency band and bandwidth configuration, networking resource allocation, beam management, and measurement procedure design. This paper proposes diverse deployment strategies for ISAC systems across various frequency bands as well as coordination approaches to optimize the allocation of sensing and communication resources at base stations (BSs) and in neighboring cells. Additionally, the paper explores the viability of reusing communication beams as sensing beams in different sensing modes and outlines the configuration and reporting mechanisms for sensing measurements corresponding these modes. The research findings presented in this paper shed light on the potential deployment of ISAC systems in the near term.

Keywords

ISAC, 5G-A, 6G

* Corresponding author

1 Introduction

Integrated sensing and communication (ISAC) is one of key potential technologies in the near 5G-Advanced (5G-A) and future 6G eras [1, 2]. With ISAC, communication networks can achieve combined sensing and communication functions, i.e., detecting targets in their surrounding environments while fulfilling their primary communication purposes.

The significance of ISAC lies in its ability to utilize spectrum resources efficiently. Traditionally, frequency bands reserved for communication and radar sensing do not overlap in order to avoid interference, even though they may be close in frequency. For example, L-band (1–2 GHz) is used for cellular communications, remote monitoring, and air traffic control; S-band (2–4 GHz) is used for cellular communications, Wi-Fi (802.11b/g/n/ax), air traffic, and airborne early warning; C-band (4–8 GHz) is used for 5G and Wi-Fi (802.11a/n/ac/ax) communications, weather and ground monitoring. The millimeter wave (mmWave) band is used for environment sensing, high-resolution imaging radar detection [3], and Wi-Fi (802.11ad/ay). This band is also a promising candidate for cellular communications. An ISAC system integrates communication and sensing functions through methods such as spectrum sharing, device sharing, and waveform sharing on the same frequency band. This allows for more efficient use of spectrum resources and enables a range of novel physical sensing applications, providing foundational capabilities for the upcoming 6G digital intelligent world.

5G is designed to serve three primary application scenarios: enhanced mobile broadband (eMBB), massive machine-type communications (mMTC), and ultra-reliable low-latency communications (URLLC). It aims to enhance communication KPIs, including spectral efficiency, latency, device connectivity, and communication rate. In addition to enhancing communication KPIs, ISAC involves research and design of sensing KPIs and ISAC KPIs, which will drive more innovative applications based on communication devices and physical world sensing. By incorporating environment sensing, communication services can be enhanced, and more environment-sensing services can be developed based on communications infrastructure, which will ultimately facilitate the creation of a digital intelligent world.

To realize ISAC, key technical challenges include integrated waveform design, frame structure design, resource allocation, measurement procedure, and beam management on the air interface side. Technical challenges on the network side include converged architecture design, data management and transmission, and security mechanisms [3, 4]. In the 5G-A phase, wireless channel features will be used to acquire sufficient environmental information and implement fundamental sensing applications based on the 5G infrastructure and enhanced air interface design. Currently, dedicated sensing reference signals are designed based on OFDM waveforms for two typical sensing modes, i.e., monostatic sensing (where A transmits and A receives) or bistatic sensing (where A transmits and B receives) applied at the BS side. The network architecture design considers to introduce a new network element for the management of sensing applications and sensing measurement procedures.

To facilitate new sensing services in the upcoming 6G era, researchers are exploring new sensing waveforms, including OFDM, FMCW, and OTFS [5], new sensing frequency bands such as the Terahertz (THz) band [6, 7], and new technologies like AI. In the future, sensing and communication will be integrated on a more advanced level to enable coexistence, coordination, and reciprocity.

This paper explores candidate technologies for future air interface design in ISAC systems. It covers aspects such as frequency band and bandwidth configuration, networking resource allocation, beam management, and measurement procedures design. Different sensing frequency bands and bandwidths can support different sensing service requirements and schemes for reusing resources with communication functionality, which can pave the way for ISAC device deployment. Networking resource allocation and coordinating sensing and communication interference between BSs is a practical challenge for deploying BS-based sensing at scale. ISAC beam management and measurement are necessary mechanisms for realizing the sensing function on the air interface side. ISAC beam management involves managing sensing beams and allowing the reuse of sensing beams and communication beams. ISAC measurement procedure design involves resource configuration of sensing measurement signals in different sensing modes and the sensing measurement reporting mechanism between the user equipment (UE) and BS.

2 Research on Key ISAC Technologies

2.1 Frequency Band and Bandwidth Configuration

The implementation of ISAC technology varies depending on the frequency band, waveform, and application scenario. The frequency band, in particular, significantly impacts the sensing waveform and application scenario. Therefore, the deployment of ISAC must be designed based on the characteristics of the frequency band.

In medium- and low-frequency bands (< 10 GHz), the sensing bandwidth can reach 200 MHz or 400 MHz. This enables positioning, intrusion, and health detection sensing applications with meter-level or decimeter-level distance precision. Communication services mainly operate in medium and low-frequency bands, which require high coverage, high rates, massive connections, and high reliability. Therefore, the principle on how to meet the requirements of both sensing and communication applications should be determined before applying ISAC technology in these bands. It could be serving the sensing application in a best-effort manner, or fulfilling the requirements of high-demand sensing services. For the best-effort mode, the deployment of ISAC integrates sensing without affecting communication performance. The sensing signal can reuse communication signals as much as possible, and sensing and communication share the same coverage . This eliminates the need for separate deployment of additional communications infrastructure to meet sensing requirements.

The mmWave band ranges from 10 to 100 GHz, being able to offer a sensing bandwidth of 800 MHz to several GHz. This makes it ideal for various sensing applications such as positioning, tracking, imaging, and posture recognition. This band allows for precise sensing with a distance accuracy of decimeters or even centimeters. While the mmWave band is primarily used indoors in homes and factories to meet high-speed and low-latency communication needs, it also provides abundant frequency resources and is the main band for target positioning and sensing in surrounding environments. To deploy ISAC effectively in the mmWave band, it is crucial to integrate sensing into the communications infrastructure and achieve cost-effective integration benefits. Additionally, it is necessary to explore suitable temporal, frequency, and spatial multiplexing schemes to minimize interference between the sensing and communication functionalities.

THz frequencies, specifically in the sub-THz band ranging from 100 GHz to 300 GHz, offer a wide sensing bandwidth that spans several dozen GHz. This capability enables various sensing applications, including high-resolution imaging and security checks. The precision of THz sensing can reach the millimeter level or higher. THz communication is primarily applied in two key scenarios, depending on the transmit power and application requirements. First, it is used for local, short-distance communication services that demand high-speed and low-latency connections, such as in homes and factories. Second, THz communication is used in ground stations to support low- and high-altitude communication. Therefore, it is crucial to specify the application scenario when deploying THz-based ISAC systems. Considering the diverse applications of THz sensing, one possible approach is to deploy the THz sensing band independently. This involves deploying THz communication, THz sensing, and THz ISAC in a localized area to meet communication requirements while minimizing cross-interference through frequency-division multiplexing.

2.2 Networking Resource Allocation

Efficient allocation of resources is crucial for ISAC systems to meet the KPI requirements of both communication and sensing services. Orthogonal ISAC resource allocation has emerged as a prominent research area focusing on orthogonal time division, frequency division, and space division. This approach ensures that sensing and communication resources are orthogonal to each other in different dimensions to prevent interference. By contrast, non-orthogonal ISAC resource allocation utilizes the same time-frequency resources for sensing and communication functions. This method is suitable when sensing and communication services share spatial consistency. For instance, when a BS sends a downlink signal to a device, it can also sense surrounding targets between the BS and the device using the echo signal of the downlink signal. In ISAC resource allocation, both sensing and communication performance need to be considered. One primary allocation mode involves optimizing the performance of one function while maintaining the baseline performance of the other [8]. Another approach is to optimize ISAC KPIs such as ISAC utility or ISAC efficiency [9], which is one indicator considering both sensing and communication performance.

In both monostatic and bistatic sensing modes, the reception of sensing signals can be affected by interference from various sources. The interference includes sensing

interference between sensing targets, communication interference from the local and neighboring cells. The figure below illustrates how communication interference affects the BS that receives sensing signals. The interference includes uplink interference from its local cell and downlink and uplink interference from neighboring cells. When designing the network and allocating sensing resources, it is essential to consider communication interference from the local and neighboring cells to ensure optimal performance in target sensing.
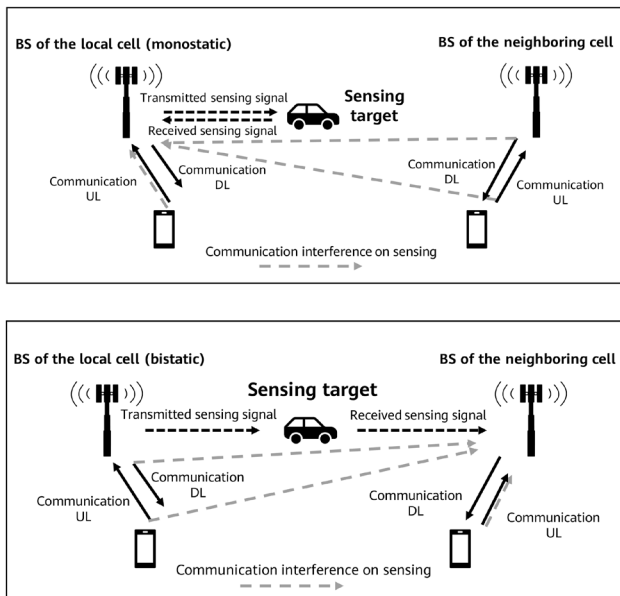
Figure 1 BS sensing inference in monostatic and bistatic sensing modes

When designing an ISAC network, it is important to carefully coordinate sensing resources between local and neighboring BSs. In monostatic mode, neighboring BSs must avoid scheduling their downlink or uplink communication on the sensing resource, as this could interfere with the local BS's ability to receive the sensing signal. In bistatic mode, coordination is necessary to prevent transmitting and receiving conflicts at a BS on the sensing resources when a neighboring BS is transmitting sensing signal to it. Meanwhile, coordinating interferences between sensing and communication becomes more challenging when monostatic and bistatic sensing coexist on the network. To address this challenge, we propose two approaches for efficiently coordinating sensing resources between BSs.

1. Grouping BSs to coordinate sensing resources

This approach involves dividing BSs into groups based on their distance from each other. To prevent interference between sensing signals of different BS groups, sensing resources are organized into different pools which are

associated to different BS groups and are mutually orthogonal in time and frequency domains. To avoid interference between sensing and communication signals, a separate communication resource pool, orthogonal to the sensing resource pools, is used to allocate communication resources among BS groups. Within each BS group, a BS informs its reserved sensing resources from the corresponding sensing resource pool to the target neighboring BSs and sends the sensing signal after receiving an acknowledgment from the neighboring BSs.
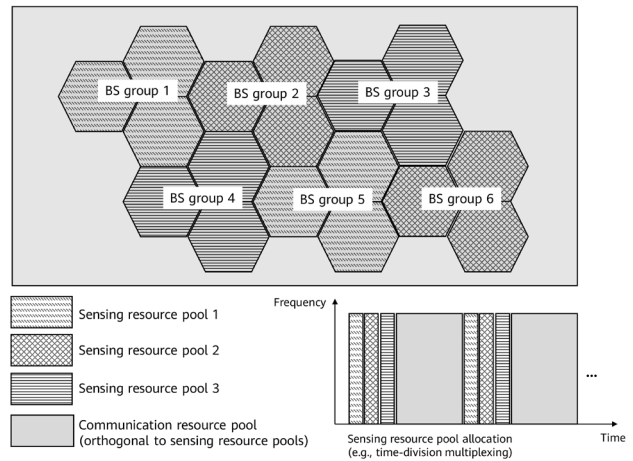
Figure 2 BS groups with orthogonal sensing resource pools

2. Coordinating sensing resources through core network

In this approach, the communication and sensing resources of BSs are uniformly managed by a core network element. Upon receiving a sensing service request from a BS, the core network element schedules and allocates sensing resources. This prevents sensing conflicts between BSs and interferences between sensing and communication as well. The sensing service request includes the sensing performance requirements, such as distance accuracy, sensing range, and sensing latency. Based on this information, the core network element determines the transmission period, sensing bandwidth, level of sensing transmitting power, sensing service priority, and other necessary details for allocating sensing resources.

2.3 Beam Management

5G NR systems have a unified air interface design for both high- and low-frequency bands. These systems use digital and analog hybrid beamforming technology. Beam alignment ensures that the transmitter and receiver communication beams are aligned, allowing for effective communication

between the UE and BS. Communication beam management encompasses beam measurement reporting, beam indication, and beam failure recovery [10]. For ISAC applications, a sensing beam management mechanism is also required for transmitting and receiving sensing signals. The integration of communication and sensing beams can vary depending on the sensing mode used.

1. Monostatic BS sensing

In this mode, the sensing BS incorporates an independent sensing receiver antenna to reduce self-interference. The BS periodically sends sensing beams in different directions, similar to sending synchronization signal blocks (SSBs), to sense targets in the vicinity of the BS. The sensing beam and the BS downlink communication beam can be integrated through time-division or space-division multiplexing techniques.

2. Bistatic sensing between the BS and UE

In this sensing mode, the sensing service operates independently from the communication service. Only sensing requirements are considered when allocating sensing resources and indicating sensing beams. Beam measurement and reporting of the sensing service are performed only within the sensing range where sensing targets are located. This differs from the communication service, where the beam with optimal channel quality is reported. When the sensing service and communication service are spatially consistent, a periodic or aperiodic communication signal can be reused to perform sensing measurement on the UE or BS side. To enhance the sensing range or accuracy while ensuring communication, we can use the repetitive transmission mechanism to indicate different beams in different transmission symbols or timeslots. The figure below illustrates the process of downlink repetitive transmission using distinct transmission beams to indicate various transmission opportunities. This supports sensing under different beams. This way, sensing and communication are closely integrated.
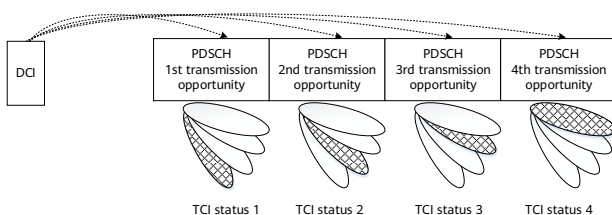
Figure 3 Beam indications of different transmission opportunities for downlink repetitive transmission

3. Bistatic BS sensing

In bistatic BS sensing mode, the transmitter BS must indicate the sensing transmission beam to the receiver BS to support sensing receiving measurement. Because the BSs are located in different geographical locations to cover different sensing areas, the sensing transmission beams going to different sensing receiver BSs from the same transmitter BS can be different. To determine its sensing transmission beam, the transmitter BS should measure the sensing beam in bistatic BS sensing mode.

Additionally, unlike communication services that cover entire areas, sensing services are only allowed in licensed areas due to regional security concerns. Therefore, the range of sensing beams must be limited. On the transmitter BS side, the range of sensing beams is controlled by adjusting the power of the sensing transmission beam and the phase configuration of the phased antenna array.

2.4 Measurement Procedure Design

During communication measurement, the receiver UE uses a reference signal to perform channel measurement and sends the measurement result to the BS or transmitter UE (in a sidelink communication scenario). For sensing measurement, the sensing signal receiver conducts sensing measurement and reports the result to the BS (in a UE measurement scenario) or the sensing result processing network element of the core network.

Table 1 analyzes the configuration and reporting of sensing measurement in different sensing modes. The sensing measurement scenarios include uplink sensing measurement, downlink sensing measurement, monostatic BS sensing measurement, bistatic BS sensing measurement, monostatic UE sensing measurement, and bistatic UE sensing measurement. Depending on the correlation between the sensing service and communication service, the resources used for sensing signal transmission can be independent sensing reference signals or reused communication reference and data signals. Additionally, the transmission priority of each sensing signal needs to be determined to prevent conflicts with communication signals or other sensing signals. As the standardization of 6G reference signals and signaling flow has not yet commenced, this discussion uses 5G-A reference signals and signaling flow as an illustrative example.

Sensing measurement reporting refers to either of these processes: (1) The UE reports downlink, monostatic UE,

Table 1 Sensing measurement configuration and reporting in different sensing measurement scenarios

Sensing Measurement Scenario	Sensing Measurement Configuration	Sensing Measurement Reporting
Uplink sensing measurement	SRS, PUSCH (DM-RS), dedicated uplink sensing signal configuration	From the BS to the sensing result processing network element of the core network
Downlink sensing measurement	SSB, CSI-RS, PDSCH (DM-RS), dedicated downlink sensing signal configuration	From the UE to the BS, reusing the CSI measurement reporting flow
Monostatic BS sensing measurement	Multiplexed downlink data or reference signal configuration, dedicated monostatic BS sensing signal configuration	From the BS to the sensing result processing network element of the core network
Bistatic BS sensing measurement	Data or reference signal configuration in IAB mode, dedicated bistatic BS sensing signal configuration	From the sensing receiver BS to the sensing result processing network element of the core network or the sensing transmitter BS
Monostatic UE sensing measurement	Multiplexed sidelink data or refer-ence signal configuration, dedicated monostatic UE sensing signal configuration	From the UE to the BS (the BS sends a measurement reporting request to the UE)
Bistatic UE sensing measurement	SL-CSI-RS, PSSCH (DM-RS), dedicated sidelink sensing signal configuration	From the UE to the BS (the BS sends a measurement reporting request to the UE) or from the sensing receiver UE to the sensing transmitter UE

or bistatic UE sensing measurement results to the BS; (2) The BS reports the monostatic or bistatic BS sensing measurement results to the sensing result processing network element of the core network or the sensing transmitter BS. The receiver of the sensing measurement result can also transmit the sensing result to the sensing result requester, which can be a UE, BS, or core network element based on a high-level procedure. A sensing measurement result includes channel state information and new sensing measurements, such as sensing distance, azimuth, pitch angle, and Doppler spectrum estimation.

For downlink sensing measurement and reporting from the UE to the BS, the procedure for measuring and reporting communication channel state information can be reused. This procedure supports periodic, semi-persistent, and aperiodic measurement reporting. Reporting resources are configured at a higher level according to the specific sensing reporting type, and reporting procedure is performed with reference to MAC signaling activation and PHY DCI signaling triggering. The measurement result needs to be associated with the corresponding sensing signal or sensing session. To implement sidelink sensing measurement and reporting from the UE to BS, a proper signaling flow is required to allow the BS to trigger reporting from the UE. The BS can trigger the UE to perform sidelink sensing measurement and reporting according to a high-level request, where the trigger signaling contains the sidelink range or sidelink

destination identifier. The UE that receives the trigger signaling obtains the required sidelink sensing measurement result through active sidelink sensing measurement (sending a sensing signal to another UE) or passive sidelink sensing measurement (receiving a sensing signal from another UE) and then reports it to the BS.

In a regional sensing scenario, it is possible to implement a shared resource pool for sensing measurement and reporting to reduce sensing measurement and reporting overheads on the UE side. This can avoid the issue of frequent resource allocation and deallocation for sensing measurement and reporting with UE dedicated signaling, especially in scenarios involving frequent UE changes due to high-speed movement.

3 Conclusion

In this paper, we examined potential technologies for future air interfaces in ISAC systems, including frequency band and bandwidth configuration, networking resource allocation, beam management, and measurement procedure design. We analyzed candidate deployment schemes for ISAC in low, mmWave, and THz frequency bands, considering their unique characteristics. To reduce communication interference on the sensing receiver BS sourced from

neighboring cells, we proposed two strategies for resource allocation: classifying BSs into groups and using the core network to coordinate sensing resources. Regarding beam management, we analyzed the multiplexing of sensing and communication beams in different sensing modes and proposed a multiplexing mechanism that leverages repetitive transmission for bistatic sensing between the BS and UE. Regarding the design of measurement procedures, we discussed sensing resource configuration and reporting mechanisms in different sensing measurement scenarios.

However, there are areas of interest that require further research. Current most sensing solutions implement monostatic or bistatic sensing measurement by designing a dedicated sensing signal based on the OFDM waveform. In addition to the key technologies discussed in this paper, the frame structure and carrier aggregation scheme also need to be considered for ISAC networking and applications. For instance, it is important to determine whether the same frame structure is used in different sensing modes and how to design the guard interval and cyclic prefix (CP) based on frequency bands and sensing applications. Allocating contiguous bandwidth from a single band remains a challenge, so technologies such as carrier aggregation and joint deployment of high and low frequencies need to be introduced to support better communication and sensing applications. Therefore, we suggest analyzing the impact of carrier frequency offset on sensing performance and hardware design. Furthermore, the introduction of new sensing waveforms will depend on whether the current OFDM-based waveform design can be adapted to future ISAC applications.

References

[1] IMT-2030(6G) Promotion Group, "6G Vision and Candidate Technologies," June 2021.

[2] IMT-2030(6G) Promotion Group, "6G Usage Scenarios and Key Capabilities," July 2022.

[3] IMT-2030(6G) Promotion Group, "Study on Integrated Sensing and Communication (the 2nd version)," Nov. 2022.

[4] IMT-2020(5G) Promotion Group, "Study on Network Architecture of 5G-Advanced Integrated Sensing and Communication," Nov. 2022.

[5] L. Gaudio et al., "On the effectiveness of OTFS for joint radar parameter estimation and communication," IEEE Transactions on Wireless Communications, vol. 19, no. 9, pp. 5951–5965, 2020.

[6] J. He et al., "6G integrated sensing and communication: Wireless sensing and sensing assisted communication," Information and Communications Technology and Policy, 2022,48(9):9–17.

[7] T. Hu et al., "New Opportunities for Multi-Band Collaborative Communication Integrated Communications and Sensing at Terahertz Band," ZTE Communications, 2022, 28(4):14–18.

[8] Z. Wang et al., "Achieving the Performance Bounds for Sensing and Communications in Perceptive Networks: Optimal Bandwidth Allocation," in IEEE Wireless Communications Letters, vol. 11, no. 9, pp. 1835–1839, Sept. 2022.

[9] J. Jiang et al., "Performance Metric for 6G Integrated Sensing and Communication System," ZTE Communications, 2022,28(05):39–45.

[10] X. Liu et al., "5G Wireless System Design and International Standard," Posts & Telecom Press, Sept. 2019.

Integrated Sensing and Communication: Another Promising Opportunity for Advancing Wireless Networks

Dajie Jiang, Jian Yao, Jianzhi Li, Shengli Ding, Yannan Yuan, Baolong Chen, Fei Qin
vivo Mobile Communication Co., Ltd.

Abstract

Upon its emergence in the communications industry, the wireless network underwent a significant growth phase. Now, it is poised to undergo another significant phase of progress with the advent of integrated sensing and communication (ISAC). This development is expected to result in a wide range of applications for sensing services. 6G, the forthcoming generation of wireless communications, is expected to function on the millimeter wave (mmWave) or even terahertz frequency, which will offer a larger bandwidth and enhanced penetration capability. 6G will also witness the implementation of larger-scale antennas and denser networks. Collectively, these factors will enhance the ability of the 6G system to deliver sensing services with higher accuracy and resolution. This paper provides an overview of the use cases, performance indicators, efficiency indicators, degree of integration, sensing modes, and key technologies related to ISAC. It also covers the development and testing of the ISAC prototype, and then concludes by addressing four challenges associated with ISAC.

Keywords

6G, ISAC key technology, performance indicator, sensing mode, prototype verification

1 Introduction

As emerging services, such as immersive extended reality (XR), holographic telepresence, interactive 3D virtual human (3DVH), collaborative robot, unmanned driving, multi-sensory interconnection, and even metaverse are gaining in popularity [1–3], tougher demands are being placed on communications, sensing, and computing (including artificial intelligence, i.e., AI). 6G is bound to become the cornerstone of society's efficient and sustainable network information development. Its inherent capacity to seamlessly integrate communications, sensing, and computing services will serve as the catalyst for a wide range of innovative services. Looking to the future, communications, sensing, and computing, the three core capabilities, will play a pivotal role in establishing seamless connections between the physical and digital worlds in the times ahead. In addition to providing communication capabilities for seamless physical-digital connections, 6G must also be able to sense the physical world in order to construct the corresponding digital world. Furthermore, to process data between these two worlds, 6G requires computing capabilities, including computing power, storage, and intelligence. Only with these capabilities can 6G ensure the comprehensive integration and efficient interaction of the physical and digital worlds, thereby realizing the vision of a fully connected physical-digital integrated world [4].

Both wireless communications and wireless sensing rely on the principles of electromagnetic wave theory. At the transmitting end, an electromagnetic wave signal is modulated to incorporate source information within it. However, as this signal propagates, it is inevitably influenced by the surrounding radio environment, thereby acquiring additional information pertaining to the environment. Consequently, the receiving end can extract from this signal not only source information, but also sensing information that reflects the characteristics of the propagation environment. In essence, electromagnetic wave signals possess inherent capabilities for both communication and sensing, making it possible to implement ISAC. Upon its emergence in the communications industry, the wireless network underwent a significant growth phase. Now, it is poised to undergo another significant phase of progress with the advent of ISAC. This development is expected to result in a wide range of applications for sensing services. When compared with independent sensing and communication systems, the ISAC system has a number of advantages, for example, lower costs, smaller device size, reduced power consumption, improved spectral efficiency, and minimized mutual interference between communication and sensing.

6G, the forthcoming generation of wireless communications, is expected to function on the millimeter wave (mmWave) or even terahertz frequency, which will offer a larger bandwidth and enhanced penetration capability. 6G will also witness the implementation of larger-scale antennas and denser networks. Collectively, these factors will enhance the ability of the 6G system to deliver sensing services with higher precision and resolution. This paper presents an overview of the use cases and performance indicators pertaining to ISAC in Section 1. Then in Section 2, it describes the level of integration, sensing modes, and key technologies linked to ISAC. It proceeds to discuss the development and testing of the ISAC prototype in Section 3 and explores how to address four ISAC-related challenges in Section 4. Finally, the paper concludes by summarizing the key findings.

2 Use Cases and Indicators

The role of the mobile communications system is poised to undergo significant changes as it evolves towards 6G. These changes will help ensure that the system can better support emerging services like immersive XR, holographic telepresence, interactive 3DVH, collaborative robot, unmanned driving, multi-sensory interconnection, and even the metaverse. In addition to functioning as an information conveyor, 6G will also serve as an information producer and processor. As an information producer, 6G system will provide sensing services, for example, using radio waves to sense and obtain the results of moving speeds, directions, materials, and imaging of various objects in the physical world. And as an information processor, the system will provide computing and AI services, such as imaging computing and AI model training, to perform preprocessing, storage, and analysis on related data. 6G will provide comprehensive information services that encompass transmission, production, and processing of information thanks to its inherent capabilities in communications, sensing, and computing, enabling a wide range of diverse services in the coming years.

Sensing is not a new concept. It is already used, for example, to detect the location (positioning) of user equipment (UE) in 4G or 5G networks. As we move towards 5.5G and 6G, the range of sensing will extend beyond connected targets (e.g., UEs) to include unconnected

targets, such as humans and weather conditions. Furthermore, the scope of sensing will go beyond location information to encompass breathing and heartbeat frequencies, object materials, moving speeds, gestures, postures, rainfall amount, and more. The 6G ISAC scenario (shown in Figure 1) defined by the International Telecommunication Union (ITU) will enable the 6G system to sense the spatial information, motion status, and environmental details of both connected and unconnected targets [5]. This usage scenario encompasses various use cases related to navigation, motion detection, and tracking, including the identification of posture, detection of falls, and identification of vehicles and pedestrians. Additionally, it includes use cases pertaining to environment monitoring such as detecting rain and pollution. It also encompasses use cases associated with environmental data in the field of AI, XR, and digital twin applications. In addition to communication indicators, the usage scenario requires advanced capabilities in precise locating and sensing. These capabilities include estimating the range, velocity, and angle information, detecting the existence of objects, and performing tasks such as localization, imaging, and mapping.
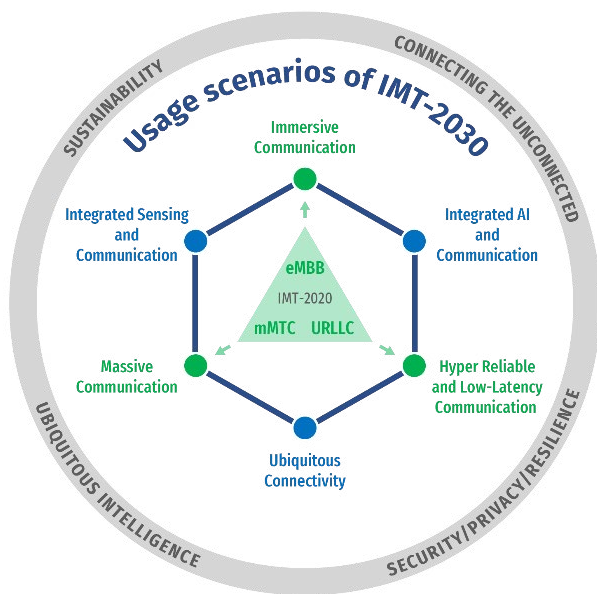
Figure 1 Six usage scenarios of 6G (IMT-2030) defined by ITU [5]

2.1 Use Cases

Sensing, in a narrow sense, refers to the sensing using radio wave signals. During the transmission of these signals, various factors in the propagation environment cause changes in their characteristics, such as amplitude and

phase. By analyzing these changes, the receiver can gather information about the transmitter carried within the signals. The receiver can also extract information that reflects the specific characteristics of the propagation environment. This process enables the receiver to obtain information regarding target objects and the propagation environment, including location, speed, direction, material composition, imaging, and weather conditions. Radio wave signals can be used to sense unconnected targets and connected targets (for example, UEs and backscatter tags).

Broadly defined, sensing can be categorized as contact sensing or contactless sensing based on whether a sensing device (such as a sensor) is physically in contact with the sensing target or object, as referenced in [6]. Contact sensing involves the use of various sensors, which are installed on the sensing target to gather specific sensing information. Such sensors include thermometers, hygrometers, barometers, gyroscopes, accelerometers, and gravity sensors. Contactless sensing relies on mediums like light, sound, or radio waves to carry out the sensing process. It can be categorized into light sensing (such as sensing using cameras, infrared cameras, or laser radars that utilize visible light or infrared rays for sensing), sound sensing (such as sonar sensing based on mechanical or ultrasonic waves), and radio wave sensing (such as sensing using mmWave radars). In [6], sensing can also be categorized into radio frequency (RF) sensing and non-RF sensing, where non-RF sensing can be further classified into contact sensing, light sensing, and sound sensing. RF sensing, also known as radio wave sensing, allows sensing devices to be installed remotely from the targets to be sensed. In contrast with non-RF sensing, RF sensing is less susceptible to the impact of weather and lighting conditions, offers a broader range of sensing, and provides greater flexibility. It outperforms camera-based sensing in terms of privacy and security. RF sensing and non-RF sensing can work together in a complementary manner.

Figure 2 Categorization of sensing

In [7], sensing is further expanded into basic information services, which include wireless sensing, enhanced network information exposure, and industry public information services. Unlike 5G, which provides only limited information services such as UE positioning and network information provisioning, 6G will enhance these capabilities in addition to providing wireless sensing. Moreover, 6G is expected to collect diverse sensor information and industry public data, such as Geographic Information System (GIS) information.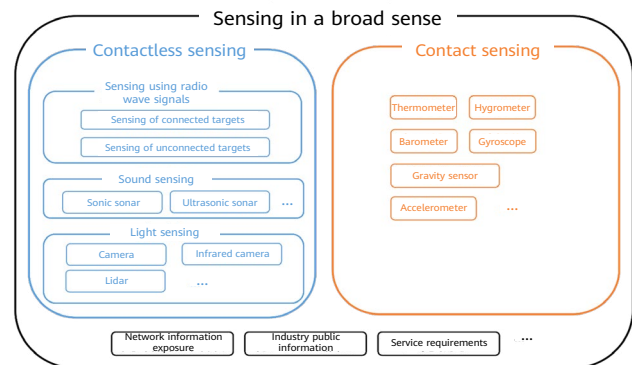 This can avoid the duplicated collection of the information by different industry applications. In [8], the concept of sensing is further broadened to include the sensing of service requirements (such as transmission data rate, communication latency, and computing power) and both physical and digital-space network status (including network elements operation status, resource utilization, AI training model library status, and digital twin network operation status). In essence, both [7] and [8] incorporate the collection/acquisition of valuable information into the scope of sensing, i.e., sensing in a broad sense. Figure 2 shows the categorization of sensing.

In ISAC, sensing mainly refers to radio wave sensing. The 3rd Generation Partnership Project (3GPP) TSG SA WG1 (SA1) provides 32 use cases pertaining to sensing in [9]. These use cases cover intruder detection in smart homes, pedestrian/animal intrusion detection on highways, rainfall monitoring, vehicles sensing for advanced driver assistance systems (ADAS), automated guided vehicle (AGV) detection and tracking in factories, unmanned aerial vehicle (UAV) flight trajectory tracing, sensing with/without obstacles at crossroads, contactless sleep monitoring, and more.

2.2 Sensing Performance Indicators

Sensing indicators can be classified into two categories: performance indicators and efficiency indicators. Performance indicators measure the performance of the sensing service and have a direct impact on users' service experience. In addition to covering common indicators used in most sensing-related use cases, such as service latency and refreshing rate, performance indicators encompass indicators specific to a particular use case, such as precision, resolution, detection accuracy, and missed detection. Table 1 describes the sensing performance indicators.

Table 1 Sensing performance indicators [9]

Indicator	Definition
Detection accuracy (also known as detection probability)	The probability of correctly detecting the target status.
Missed detection	The probability of missing the detection of the sensing result that represents the characteristics of a target object or environment when the sensing system attempts to acquire a sensing result.
False alarm	The probability of detecting an event that does not represent the characteristics of a target object or environment when the sensing system attempts to acquire a sensing result.
Accuracy of positioning estimate	The closeness of the measured sensing result (i.e., position) of the target object to its true position value. It can be divided into a horizontal sensing accuracy (referring to the sensing result error in a 2D reference or horizontal plane) and a vertical sensing accuracy (referring to the sensing result error on the vertical axis or altitude).
Accuracy of velocity estimate	The closeness of the measured sensing result (i.e., velocity) of the target object's velocity to its true velocity.
Range resolution	The minimum difference in distance between target objects to have a measurably different range.
Velocity resolution	The minimum difference in velocity between target objects to have a measurably different velocity.
Recognition accuracy	The probability of accurately identifying the type/meaning of a sensing target.
Refreshing rate	The rate at which the sensing result is generated by the sensing system. It is the inverse of the time elapsed between two successive sensing results.
Maximum sensing service latency	The time elapsed between the event triggering the determination of the sensing result and the availability of the sensing result at the sensing system interface.
Confidence level	The percentage of all the possible measured sensing results that can be expected to include the true sensing result considering the accuracy.

The 5G-oriented 3GPP report of ISAC feasibility study [9] also provides the requirements on these indicators in diverse sensing use cases. For example, the performance requirements of sensing results for intruder detection in smart homes are as follows:

- Confidence level: 95%

- Accuracy of horizontal positioning estimate by sensing: ≤ 10 m

- Accuracy of vertical positioning estimate by sensing: ≤ 10 m

- Maximum sensing service latency: < 1000 ms

- Refreshing rate: < 1s

- Missed detection: < 5%

- False alarm: < 2%

The 6G key capabilities [5] defined by the ITU include the positioning accuracy of 1–10 cm.

2.3 Sensing Efficiency Indicators

Sensing efficiency indicators reflect the cost and efficiency involved in provisioning sensing services on a network. Defining reasonable efficiency indicators, which encompass spectral efficiency, energy efficiency, and cost efficiency [7], is crucial for promoting sustainable network development. The spectral efficiency, energy efficiency, and cost efficiency associated with sensing denote the quantity of time-frequency resources, energy, and cost required to accomplish a sensing task, respectively. Research and development as well as system design related to sensing should consider both sensing performance indicators and efficiency indicators.

3 Degrees of Integration, Sensing Modes, and Key Technologies for ISAC

This section describes the degrees of integration and sensing modes, and provides a brief introduction to the key technologies involved in ISAC.

3.1 Degrees of Integration

The implementation of both wireless communications and wireless sensing relies on the reception and processing of radio signals. ISAC involves the sharing of various resources, such as spectrum, hardware, baseband, and signals, between communication and sensing services. Additionally, it requires cooperation or assistance between these two types of services. An example of this is sensing-based beam tracking, which is a sensing-assisted communications technology. Signal sharing requires the highest degree of integration and has the most stringent demands. Specifically, a signal utilized for communication and sensing services must deliver the optimal overall communication and sensing performance through waveform and signal design. Such waveforms and signals are applicable to heavy-load scenarios or delay-sensitive scenarios where communication and sensing services are operating concurrently. However, such scenarios are relatively rare. In most general sensing scenarios, it is sufficient to use time division multiplexing for communication signals and sensing signals.

3.2 Sensing Modes

Based on the directions in which sensing signals are transmitted and received, there are six sensing modes (as shown in Figure 3): Base station (BS) monostatic mode (Mode 1), BS bistatic mode (Mode 2), BS-to-UE bistatic mode (Mode 3), UE-to-BS bistatic mode (Mode 4), UE monostatic mode (Mode 5), and UE bistatic mode (Mode 6). In Mode 1 and Mode 5, achieving clock synchronization between the transmitter and receiver is easy due to both parties being co-located. However, these two modes require the BS or UE to have full-duplex capabilities. In the other four modes, the transmitter and receiver are located remotely from each other, eliminating the need for full-duplex capabilities. However, achieving clock synchronization between the transmitter and receiver is challenging in distance/latency measurement sensing. In Mode 2, a BS transmits a sensing signal, upon reception of which another BS performs sensing measurement. In this sense, the interface used in Mode 2 is a new interface that goes beyond the scope of the existing protocols. Mode 3 and Mode 4 are implemented by using downlink and uplink signals. This allows for the reuse of uplink or downlink reference signals or the use of a novel sensing-specific reference signal for sensing. Alternatively, these two modes can also use a data-carried communication signal for sensing, but only if processing of the sensing needs to be performed after the communication data is decoded. It is important to note that, if a decoding error occurs, sensing performance degrades. In Mode 6, the sensing process

relies on sidelinks. This allows Mode 6 to use the existing reference signal of sidelink communications, new sensing reference signal or a data signal of sidelink communications for sensing. By selecting appropriate terminals from a diverse range of distributed terminals, Mode 3 can improve the signal-to-noise ratio (SNR) of the sensing channel by shortening the propagation path. Table 2 lists the advantages and disadvantages of the six sensing modes.

3.3 Key Technologies

Figure 4 shows nine key ISAC technologies that empower a variety of sensing services with favorable performance and efficiency indicators pertaining to sensing. Among these technologies, ISAC waveform and signal design form the fundamental technology for meeting the requirements of sensing indicators and communication indicators. Currently, ISAC waveforms have three technical options: integrated waveform based on communication waveforms, integrated waveform based on sensing waveforms, and integrated waveform based on integration of sensing and

communication. In terms of multiple-antenna technology for sensing, there are two types of technology: virtual array-based sensing technology and beamforming-based sensing technology. Both of these technologies can improve sensing accuracy and sensing resolution. Compared with a conventional single-node ISAC system, the coordinated multi-point (CoMP) ISAC system offers enhanced capabilities by incorporating sensing collaboration among distributed transmitting and receiving nodes. This approach effectively mitigates the issue of self-interference deletion commonly encountered in single-node sensing systems. Furthermore, it expands the sensing range, achieves continuous sensing coverage, and enhances sensing accuracy through joint processing. In ISAC applications, factors related to components and hardware circuits, including timing offsets, carrier frequency offsets, channel inconsistency, and phase noise, have a significant impact on measurement accuracy. It is therefore necessary to research technologies that can effectively mitigate the impact of these non-ideal factors.

ISAC mobility management can enhance communication mobility management by leveraging information gathered through sensing, thereby improving user experience and
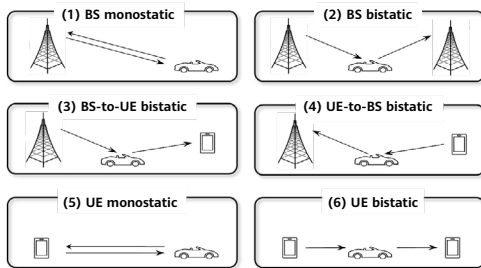
Figure 3 Sensing modes
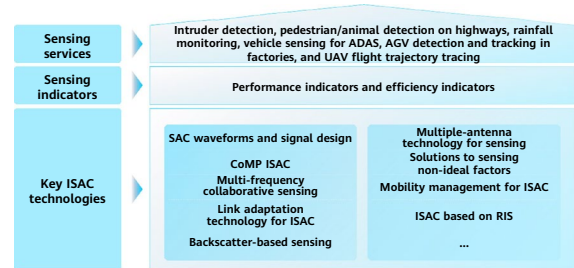
Figure 4 Key technologies

Table 2 Advantages and disadvantages of the sensing modes

Sensing Mode	Advantage	Disadvantage
BS monostatic mode (Mode 1)	Bistatic synchronization issues do not occur in this mode.	The BS must support full duplex.
BS bistatic mode (Mode 2)	Full duplex is not required.	Bistatic synchronization errors have impact on the sensing performance.
BS-to-UE bistatic mode (Mode 3)	Existing signals can be reused or enhanced to support sensing services. Full duplex is not required. The sensing coverage can be improved by selecting appropriate terminals.	The accuracy of estimating terminals' location, direction, and speed, as well as bistatic synchronization errors, have a direct impact on sensing performance.
UE-to-BS bistatic mode (Mode 4)	Existing signals can be reused or enhanced to support sensing services. Full duplex is not required. The sensing coverage can be improved by selecting appropriate terminals.	The accuracy of estimating terminals' location, direction, and speed, as well as bistatic synchronization errors, have a direct impact on sensing performance.
UE monostatic mode (Mode 5)	Bistatic synchronization issues do not occur in this mode, and unconnected sensing is supported.	Terminals must support full duplex.
UE bistatic mode (Mode 6)	Existing reference signals can be reused or enhanced for the purpose of sensing and full duplex is not required. Unconnected sensing is supported.	The accuracy of estimating terminals' location, direction, and speed, as well as bistatic synchronization errors, have a direct impact on sensing performance.

reducing system resource overhead. Moreover, ISAC mobility management can be used to ensure the continuity of sensing services. Similar to technologies like power control and bandwidth adaptation in a communications system, ISAC link adaptation can dynamically adjust sensing resource configurations based on information such as sensing measurements, thereby improving ISAC efficiency indicators. When utilized in sensing applications, reconfigurable intelligent surface (RIS) [10] can offer a wide range of technical possibilities, including RIS-assisted positioning and sensing, as well as sensing-assisted RIS communication systems. Wireless sensing based on radio frequency identification (RFID) or backscatter communication, when compared with device-free wireless sensing, not only has fundamental sensing capabilities, but can also acquire additional information related to the sensing targets. [6] and [12–14] describe some key ISAC technologies, and [11] describes the preceding nine key ISAC technologies.

4 Prototype Verification

This section describes the three ISAC prototypes and corresponding verification results.

4.1 Respiration Monitoring Based on Bistatic Sensing

As an important vital sign, respiration is closely related to human health, and it is important to monitor a person's respiration during sleep. The traditional means of respiration monitoring based on wearable devices has certain drawbacks, including compromising comfort, restricting free movement, and having limited applicability. To address these drawbacks, ISAC leverages channel state information derived from the analysis of received wireless signals to acquire characteristic data related to human health, such as breathing data. Such a wireless sensing approach has advantages of being non-contact and cost effective.

Figure 5 shows the impact of human breathing on a radio channel. Assume that a line-of-sight (LOS) propagation path and other static reflection paths exist between the transmitter and the receiver, and denote them as H_s. The reciprocating motion of the human chest brought by breathing corresponds to a dynamic reflection path H_d in a channel. Since the motion is subtle, we can overlook the amplitude change and focus on the phase changes of H_d.

The received signal at the receiver is a superposition of signals of all paths. The channel state information (CSI) obtained through the received signal's channel estimation is subjected to the phase changes of the dynamic reflection path H_d, with both its amplitude and phase periodically changing with respiration.
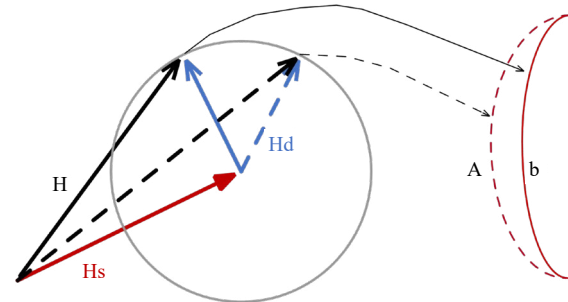
Figure 5 Respiration monitoring principle

Non-ideal factors such as frequency offset and phase noise often exist due to the difference between the clocks of the transmitter and the receiver, which can affect the performance of respiration monitoring based on the amplitude or phase of CSI. Since the impact of frequency offset and phase noise on different receiving antennas are basically the same, we can mitigate the foregoing non-ideal factors by using the CSI ratio or conjugate multiplication of different receiving antennas. The change of the signal propagation path caused by respiration leads to the phase of the dynamic reflection path changing periodically within a certain range. The position of the human body relative to the transmitter and receiver, the selection of transmitting and receiving antennas, and the selection of subcarriers all have an impact on the phase change range of the dynamic reflection path. The amplitude or phase of the complex result of the CSI ratio or conjugate multiplication varies with the phase change ranges, thereby affecting the performance of the respiration monitoring. To enhance the stability of respiration monitoring, it is possible to filter and combine the processing results of CSI ratios or conjugate multiplications of different antennas and different subcarriers based on the monitored SNR — a ratio of the power of a signal reflection path associated with human respiration to the power of noise. Considering that there are differences in the effects of respiration on the amplitude and phase of the complex result of the CSI ratio or the conjugate product, that is, there are differences in the effects on the in-phase (I) and quadrature (Q) components. Moreover, the I-channel data and the Q-channel data generally have complementary characteristics, for example, if the I-channel data

performance is satisfactory, the Q-channel data performance may be poor, and vice versa. Therefore, we can combine I- and Q-channel data via projection to obtain final time-domain detection data [15]. For instance, a CSI ratio result is expressed as $Ae^{-i\theta} = a + bi$, where a and b represent the real part (I-channel) and the imaginary part (Q-channel). The result is then projected onto the vector $[cos\theta\ \ sin\theta]$ to obtain $acos\theta + bsin\theta$, with the value of θ being selected based on the corresponding SNR estimation result. Finally, the respiration rate is calculated either by converting the time-domain detection data to the frequency domain or by using time-domain detection algorithms such as peak detection.

The respiration monitoring prototype, as shown in Figure 6, is developed based on the 5G NR system. It operates at a center frequency of 3.6 GHz and utilizes a bandwidth of 100 MHz. The BS-to-UE bistatic mode (Mode 3 in Figure 3) is used, which involves the configuration of one transmitting antenna and four receiving antennas. The prototype uses the channel state information reference signal (CSI-RS) as the sensing signal. The CSI-RS is transmitted at an interval of 20 ms, with the CSI-RS density equal to 1. In this case, the time-domain sampling rate of the sensing data is 50 Hz. Reusing the downlink reference signal of the 5G NR system for sensing does not additionally increase the sensing signal resource overhead and has no impact on uplink and downlink communication services.

Receiving antenna   UE (receiving end)   BS (transmitting end)   Transmitting antenna   Curve of CSI ratios   Respiration rate
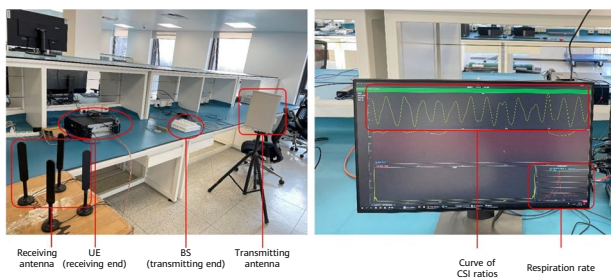
Figure 6 Respiration monitoring prototype based on the 5G NR system

The overall data processing flow is as follows:

- Obtain channel information. Retrieve the raw channel information from the received CSI-RSs through least squares (LS)-based channel estimation, and then perform noise suppression processing and calculate the CSI ratio.

- Eliminate outliers. In hardware systems, the CSI ratio result often contains amplitude outliers that need to be eliminated through filtering, such as Hampel filtering [16]. In Hampel filtering mode, a sliding window is used to detect and replace outliers based on medians and

median absolute deviation (MAD). The filter response can be expressed as:

$$y_k = \begin{cases} x_k & |x_k - m_k| \le tS_k \\ m_k & |x_k - m_k| \ge tS_k \end{cases}'$$

where m_k is the median value corresponding to the sliding window $m_k = \{x_{k-K}, \ldots, x_k, \ldots, x_{k+K}\}$ and S_k is the MAD scale estimation.

- Band-pass filtering and smoothing filtering. Perform bandpass filtering and smoothing filtering on the data according to the range of human respiration rate. The smoothing approach adopts the Savitzky-Golay filter [17]. This filter, a method based on local polynomial fitting by LS in the time domain, is widely used in data stream smoothing and denoising. It can effectively filter out noise while keeping the characteristics of the signal distribution (e.g., the relative maximum, minimum, and width) unchanged.

- Calculate the respiration rate. Calculate the combined data of the I- and Q-channel after projection based on the preprocessed CSI ratio results of various antennas and subcarriers as well as the corresponding SNRs. Then select the optimal antennas, subcarriers, and projection data for time-domain detection, and calculate the respiration rate.

4.2 Trajectory Tracking Based on CoMP Bistatic Sensing

Another typical ISAC application is the positioning and trajectory tracking of moving targets. The network can utilize the sensing modes, such as BS monostatic, BS-to-UE bistatic, or UE-to-BS bistatic, to monitor and track the trajectory of pedestrians, vehicles, UAVs, and other moving targets. To verify the feasibility of ISAC in bistatic mode, vivo developed a prototype for trajectory tracking. The prototype was built based on vivo's self-developed 6G verification platform.

Trajectory tracking can be implemented by using any of the following methods:

- The sensing node works in monostatic mode to continuously obtain the angle and range of the target and then continuous positioning and trajectory tracking are realized.

- The sensing nodes work in bistatic mode to acquire the angle of the target and the propagation distance of the reflected signal of the target; then the target's location

and trajectory are determined based on the location of the sensing nodes.

- The sensing nodes work in bistatic mode, in which multiple sensing receiving nodes obtain the delays and/ or Doppler frequencies of the reflected signal of the target. Then the target's velocity, moving direction, location, and trajectory are calculated.
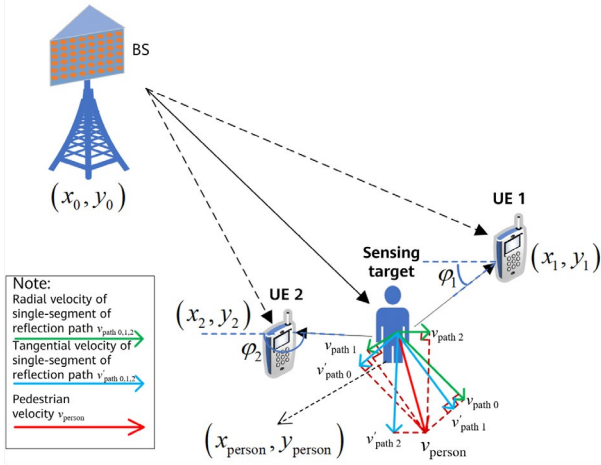
- A combination of the abovementioned methods.

Figure 7 Schematic diagram of ISAC prototype for trajectory tracking based on CoMP bistatic sensing

Currently, the research conducted by vivo focuses on methods 2 and 3. Figure 7 illustrates the principle of an ISAC prototype used for CoMP trajectory tracking in bistatic mode. The following uses a pedestrian as a sensing target for further explanation. In this scenario, the BS transmits sensing signal, which is then reflected by the human body and received by UE 1 and UE 2. By multiplying the Doppler frequencies obtained by the UEs with the signal wavelength, the change rates of the total length of the reflection paths corresponding to BS-pedestrian-UEs can be determined. And for the reflection path from the pedestrian to the BS, v_{person} can be decomposed into the radial velocity v_{path0} and the tangential velocity v'_{path0}. For the reflection path from the pedestrian to UE 1, v_{person} can be decomposed into the radial velocity v_{path1} and the tangential direction v'_{path1}. For the reflection path from the pedestrian to UE 2, v_{person} can be decomposed into the radial velocity v_{path2} and the tangential velocity v'_{path2}. Given the BS location (x_0, y_0), UE 1 location (x_1, y_1), UE 2 location (x_2, y_2), and the current pedestrian location (x_{person}, y_{person}), we can easily obtain the direction of the radial velocity vector v_{path0}, v_{path1}, v_{path2}. We assume that the location vector of the pedestrian is $p_{person} = \langle x_{person}, y_{person} \rangle$, and the location vectors of the BS, UE 1 and UE 2 that participate in sensing are $p_0 = \langle x_0, y_0 \rangle$ and $p_u = \langle x_u, y_u \rangle$, $u = 1, 2$,

respectively. For the reflection path from the pedestrian to UE 1, we have:

$$\frac{\boldsymbol{v}_{\text{person}} \cdot (\boldsymbol{p}_{\text{person}} - \boldsymbol{p}_1)}{\|\boldsymbol{p}_{\text{person}} - \boldsymbol{p}_1\|} = v_{\text{path1}}, \qquad (1)$$

For the reflection path from the pedestrian to UE 2, we have:

$$\frac{\boldsymbol{v}_{\text{person}} \cdot (\boldsymbol{p}_{\text{person}} - \boldsymbol{p}_2)}{\|\boldsymbol{p}_{\text{person}} - \boldsymbol{p}_2\|} = v_{\text{path2}}, \qquad (2)$$

and for the reflection path from the pedestrian to the BS, we have:

$$\frac{\boldsymbol{v}_{\text{person}} \cdot (\boldsymbol{p}_{\text{person}} - \boldsymbol{p}_0)}{\|\boldsymbol{p}_{\text{person}} - \boldsymbol{p}_0\|} = v_{\text{path0}}. \qquad (3)$$

In Equations (1) and (2), we assume the locations of the BS, UE 1, UE 2, and pedestrian (i.e., vectors p_0, p_1, p_2, p_{person} respectively) are known. The change rates of the total length of the reflection paths, i.e., v^1_{path}, v^2_{path} can be calculated based on the estimated Doppler frequencies, which are obtained by multiple signal classification (MUSIC) or other high-precision parameter estimation algorithms. The calculation can be performed based on the downlink sensing signal received by the UEs or the uplink sensing signal received by the BS. The relationship between the Doppler frequencies f_d^1, f_d^2 estimated by UE 1 and UE 2 and the change rates v^1_{path}, v^2_{path} of the total length of the reflection path is as follows:

$$\begin{cases} v^1_{\text{path}} = v_{\text{path0}} + v_{\text{path1}} = \lambda f_d^1 \\ v^2_{\text{path}} = v_{\text{path0}} + v_{\text{path2}} = \lambda f_d^2 \end{cases} \qquad (4)$$

Where λ denotes the signal wavelength. In Equation (4), only the radial velocities v_{path0}, v_{path1}, v_{path2} need to be concerned about, and the tangential velocities v'_{path0}, v'_{path1}, v'_{path2} are ignored because they have no effect on the length of the reflection paths. There are five scalar unknowns in Equations (1)–(4), specifically, three unknowns v_{path0}, v_{path1}, and v_{path2} related to the radial velocity (i.e., change rate of the segments of reflection paths) and two scalar unknowns corresponding to v_{person} (i.e., the projection component of the velocity vector on the x axis and y axis, respectively). The result of v_{person} can be obtained by substituting Equations (1)–(3) into Equation (4).

Assuming that the interval of two consecutive measurements of pedestrian movement is short enough (e.g., 5–10 ms), the pedestrian can be considered to move at a constant velocity during this period. Therefore, it is

possible to predict the pedestrian location at the next moment. By continuously performing N times Doppler measurements along the time dimension, the trajectory of the pedestrian can be obtained, with a given initial position.

The ISAC prototype, developed by vivo for trajectory tracking based on CoMP bistatic sensing, adopts one 4-channel universal software radio peripheral (USRP) device. This device acts as a BS to transmit sensing signal (i.e., the CSI-RS) using one transmitting antenna port. The carrier frequency is 3.5 GHz, and the bandwidth is 100 MHz. Moreover, another 4-channel USRP device is used to simulate two UEs, each of which has two receiving antennas. During testing, the two UEs continuously obtain the downlink CSI and send it to a computer. After receiving the CSI, the computer performs signal processing to obtain the sensing result. Specifically, the computer estimates the Doppler frequencies of the dynamic reflection paths caused by the pedestrian's movement based on the CSI samples obtained by UEs in continuous time, Then, the computer calculates the pedestrian's trajectory based on the estimated Doppler frequencies. Figure 8 shows the layout and the photo of the prototype.
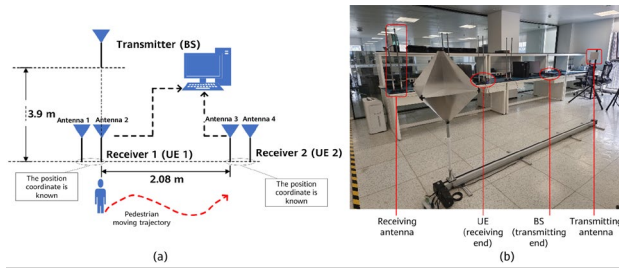

Figure 8 ISAC prototype for trajectory tracking based on CoMP bistatic sensing: (a) Layout of the prototype; (b) Prototype

A high-precision parameter estimation algorithm such as MUSIC is used to acquire the Doppler frequency of the dynamic reflection path. In each estimation, 50 consecutive CSI samples are used, with the CSI-RS period T_{CSI-RS} of 5 ms. Given an estimation, the channel matrix formed by a group of channel estimation vectors obtained by the UEs is:

$$\widehat{\mathbf{H}}_{LS} = \left[\widehat{\mathbf{h}}_0^{0:N_f-1}, \ldots, \widehat{\mathbf{h}}_{N_t-1}^{0:N_f-1} \right] = \left[\widehat{\mathbf{g}}_{0:N_t-1}^0; \ldots; \widehat{\mathbf{g}}_{0:N_t-1}^{N_f-1} \right], \qquad (5)$$

with a dimension of $N_f \times N_t$. In Equation (5), N_f represents the number of subcarriers in frequency domain used for channel estimation, N_t represents the number of CSI samples used to estimate the Doppler frequency, and $\widehat{\mathbf{h}}_l^{a:b}$ represents the channel estimation vector between pilot position a and pilot position b in the frequency domain at the lth moment. Similarly, $\widehat{\mathbf{g}}_{c:d}^m$ represents the channel estimation vector of the mth subcarrier between pilot

location c and pilot location d in the time domain. To fully utilize the frequency-domain channel estimation data and improve the accuracy of parameter estimation, the modified correlation method (i.e., the forward-backward method) [18] is used in the time domain to calculate the time domain channel correlation matrix, together with the averaging operation in frequency domain. Given the time domain channel correlation matrix $\widehat{\mathbf{R}}_{L \times L}^t$, the foregoing preprocessing process is:

$$\widehat{\mathbf{R}}_{L \times L}^t = \frac{1}{N_f} \sum_{k=0}^{N_f} \frac{1}{2} \left[\widehat{\mathbf{\Psi}}(k) + \mathbf{J}\widehat{\mathbf{\Psi}}(k)^H \mathbf{J} \right], \qquad (6)$$

$$\widehat{\mathbf{\Psi}}(k) = \frac{1}{N_t - L + 1} \sum_{l=0}^{N_t-L+1} \left(\widehat{\mathbf{g}}_{l:l+L-1}^k \right) \left(\widehat{\mathbf{g}}_{l:l+L-1}^k \right)^H, \qquad (7)$$

where L is a window length for performing moving averaging on the channel estimation matrix in the time domain. It is generally calculated using $L = \lfloor 0.5(N_t+1) \rfloor$, where $\lfloor * \rfloor$ represents rounding down. The matrix $\mathbf{J}$ is the exchange matrix with ones on the antidiagonal and zeros elsewhere, and is with a dimension of $L \times L$. $\widehat{\mathbf{\Psi}}(k)$ is the average time domain correlation matrix, which is obtained by calculating the channel estimation vectors within the moving average window of the kth subcarrier.

It is noted that due to the independent clock at transmitter and receiver, there exists a local frequency offset, which causes carrier frequency offset (CFO). The superposition of CFO and Doppler frequency offset results in trajectory estimation errors. In our prototype, the impact of CFO on Doppler frequency estimation can be eliminated by using the CSI ratio method [19]. In this method, the channel matrix $\widehat{\mathbf{H}}_{LS}$ in Equation (5) is replaced with the so-called CSI ratio matrix $\widehat{\mathbf{H}}_{Ratio}$. $\widehat{\mathbf{H}}_{Ratio}$ is obtained by dividing the channel matrices, $\widehat{\mathbf{H}}_{LS1}$, $\widehat{\mathbf{H}}_{LS2}$ obtained from the UE's two receiving antennas, respectively. In the case that the UE has only one receiving antenna, the CFO can be eliminated by utilizing the reference path (e.g., the LOS path) [20].

Once $\widehat{\mathbf{R}}_{L \times L}^t$ is calculated, the Doppler frequency can be estimated using the MUSIC algorithm. Assuming that there are two types of multipaths (i.e., the dynamic reflection paths caused by human movements, and the static LOS path and reflection paths caused by other static objects in the environment), the number of sources (i.e., estimation targets) is set to 2. Eigen value decomposition (EVD) is performed on $\widehat{\mathbf{R}}_{L \times L}^t$, that is:

$$\widehat{\mathbf{R}}_{L \times L}^t = \mathbf{U}_s \mathbf{\Sigma}_s \mathbf{U}_s^H + \mathbf{U}_n \mathbf{\Sigma}_n \mathbf{U}_n^H, \qquad (8)$$

where U_s denotes the signal subspace of the correlation matrix $\widehat{\mathbf{R}}_{L \times L}^t$, whose dimension is $L \times 2$, with each column corresponding to one eigenvector of $\widehat{\mathbf{R}}_{L \times L}^t$. Σ_s denotes the diagonal matrix, of which the diagonal elements are eigenvalues corresponding to eigenvectors and are arranged in descending order. And U_n denotes the noise subspace of the correlation matrix $\widehat{\mathbf{R}}_{L \times L}^t$, whose dimension is $L \times (L-2)$. In this case, the MUSIC pseudo-spectrum for Doppler frequency is:

$$P_{\text{MUSIC}}(f) = \frac{1}{\mathbf{s}^H(f)\,\mathbf{U}_s\mathbf{U}_s^H\mathbf{s}(f)}, \qquad (9)$$

where,

$$\boldsymbol{s}(f) = \left[1, e^{j2\pi f \times T_{\text{CSI-RS}}}, e^{j2\pi f \times 2T_{\text{CSI-RS}}}, \ \ldots, \ e^{j2\pi f \times (N_t - L + 1)T_{\text{CSI-RS}}}\right]$$

denotes the Doppler steering vector. The obtained MUSIC pseudo-spectrum has two significant spectral peaks. One is located at 0 Hz, and the other is located at a non-zero frequency, which corresponds to the estimated Doppler frequency of the reflection path caused by human movements. According to Equations (5)–(9), for each Doppler measurement, we can determine the Doppler frequencies, denoted as f_d^1 and f_d^2, of the dynamic reflection paths of BS-pedestrian-UE 1 and BS-pedestrian-UE 2, respectively.

Figure 9 illustrates a MUSIC pseudo-spectrum obtained from the measured CSI ratio, with the horizontal axis representing time (a total duration of about 6 seconds) and the vertical axis representing the Doppler frequency. It is shown that, in most instances, there are two spectral peaks. One peak consistently appears at 0 Hz, corresponding to the static reflection path and LOS path in the environment. The other peak, however, its location varies over time due to the variation of human velocity and the change of the angle between the human moving direction and the signal incidence direction. Figure 10 shows examples of the test results for different pedestrians' trajectories. The red curve represents the actual trajectory, whereas the blue curve represents the estimated trajectory. The red triangle indicates the position of the transmitter (i.e., the simulated BS), and the green and red triangles indicate the position of the two receivers (i.e., the two simulated UEs). In the test, the pedestrian walked in three trajectories: straight line, arcs, and "S" curve. In all three cases, the estimated trajectories closely align with the actual trajectories. The initial location of the pedestrian can be determined by a number of methods. One such method involves leveraging a BS or UE equipped with multiple antennas to perform angle and range measurements of the pedestrian for initial position estimation [21].
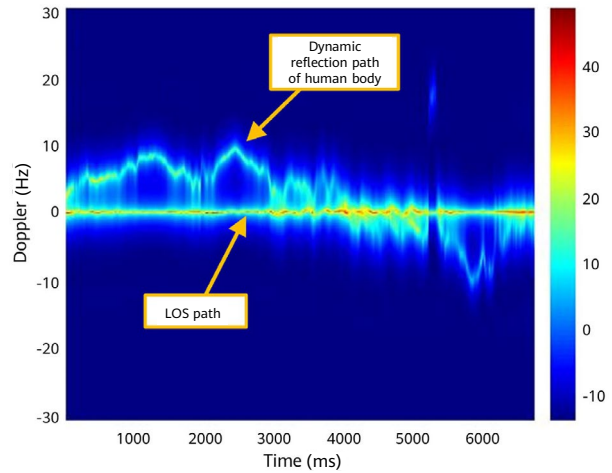

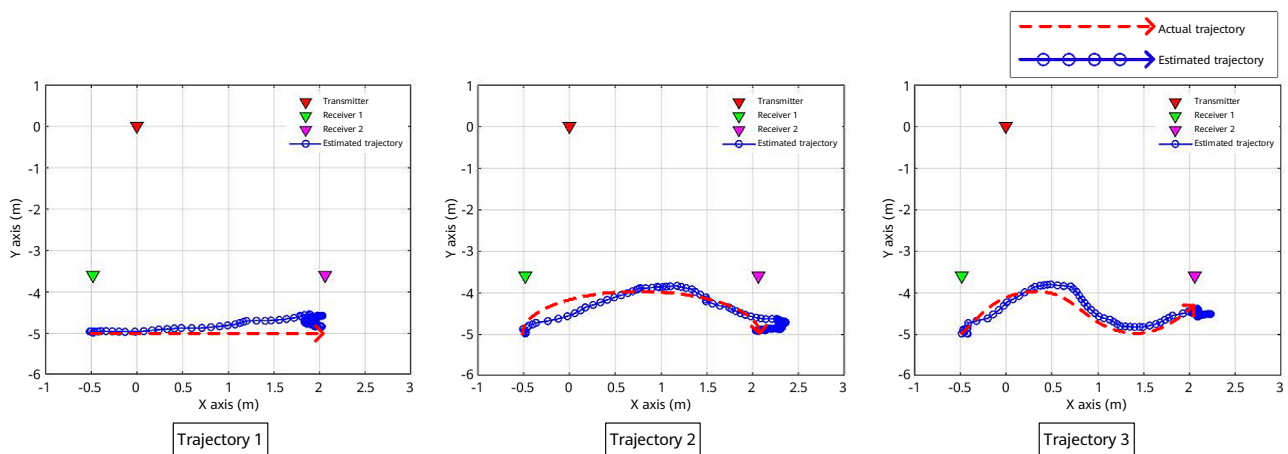**Figure 9** An example of MUSIC pseudo-spectrum based on measurements

Figure 10 Examples of trajectory tracking results

4.3 Range and Velocity Measurements for 6G Monostatic Sensing

vivo has developed an ISAC prototype based on monostatic sensing, which is used to verify the feasibility of sensing pedestrians or other moving objects in the wireless environment by receiving sensing echo signals through the BS when the communication BS is executing communication services with UEs. Figure 11 shows the architecture and graphic demonstration of the prototype, and detailed configurations can be found in [22].

The transmitter and receiver of the sensing signal share a USRP as the RF and baseband module. To prevent the receiver saturation caused by self-interference of the direct path of the sensing signal from the transmitting antenna to the receiving antenna, the two antennas need to be spatially isolated from each other.
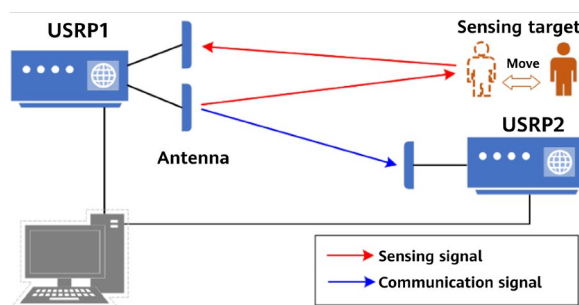
(a) Prototype architecture

(b) Graphic demonstration

Figure 11 Monostatic ISAC prototype

The receiver extracts sensing signals based on configured parameters and performs channel estimation to obtain channel matrices. Channel estimation and pre-processing before the channel estimation are performed in the same way as the reference signal processed in the NR system. Then, the sensing signals are processed based on the channel matrices, as shown in Figure 12. For a moving target, the prototype performs signal detection directly in the delay-Doppler domain and then calculates the range and velocity. For a static target, the prototype uses the clutter cancellation method to suppress the influence of environmental clutter, then performs the signal detection, and outputs the range of the static target.
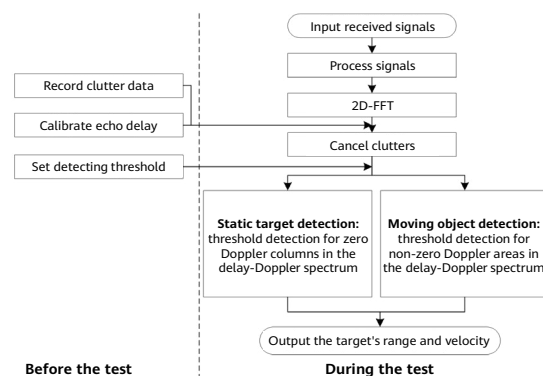
Figure 12 Sensing signal processing flow of the monostatic ISAC prototype

It is important to note that, prior to performing the measurement, the system needs to be calibrated to ensure that the sensing target can be correctly detected and the range and velocity of the sensing target can be accurately acquired. This involves three main steps. Firstly, in order to conduct the clutter cancellation, it is necessary to collect and record the clutter data before adding the sensing target to the environment. Secondly, it is necessary to calibrate the sensing echo delay in order to eliminate the impact of additional delay caused by hardware and a series of processing, which will influence the measurement of the range of the sensing target. Specifically, the additional delay can be accurately obtained by measuring the range of a sensing target whose range is already known. Furthermore, to minimize the false alarm rate and the missing detection probability, an appropriate signal detection threshold must be set. In complex indoor multipath environments, a fixed threshold cannot deliver satisfactory performance. Therefore, adaptive thresholds need to be set for signal detection according to the constant false alarm rate (CFAR) method and analysis on actual data.

The prototype has passed the testing and certification of the IMT-2030 (6G) Promotion Group. Figure 13a shows the range measurement results when the static targets are placed at 10 different positions varying from 2 m to 4.6 m away from the receiving antenna of the sensing signal, with the root mean square error (RMSE) of the measured range being approximately 0.1935 m. Figure 13b shows the velocity measurement results when the velocity of the moving target toward the receiving antenna of the sensing signal varying between 0.2 m/s and 2 m/s, with the RMSE of the measured velocity being approximately 0.0643 m/s. It is important to note that sensing and communication are performed concurrently. Figure 13c shows a measured 64 quadrature amplitude modulation (64QAM) constellation diagram, with the measured communication throughput being approximately 1.213 Gbit/s. The results of range measurement, velocity measurement, and communication throughput are consistent with the theoretical prediction of configuration parameters.

(a) Range measurement RMSE　　(b) Velocity measurement RMSE

(c) Actual 64QAM constellation diagram

Figure 13 Test results of the monostatic ISAC prototype

The range resolution and velocity resolution of the sensing signal are limited due to the limitations imposed by the sensing signal bandwidth and sensing frame length. For example, in monostatic sensing mode, the range resolution 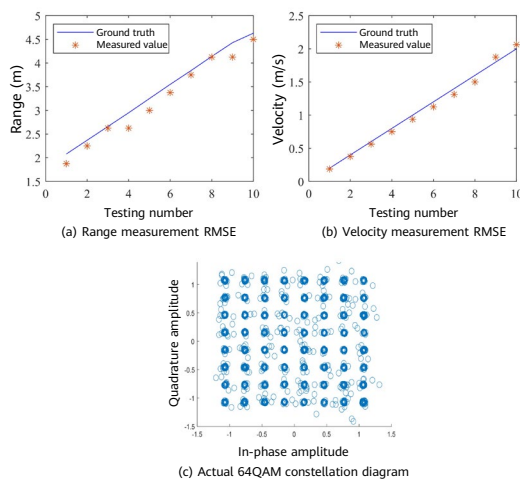corresponding to a 400 MHz sensing signal bandwidth is 0.375 m, which usually cannot meet the sensing demands in numerous indoor scenarios. There are many high-resolution algorithms in the literature, such as MUSIC and estimation of signal parameters via rotational invariant techniques (ESPRIT). However, these high-resolution algorithms typically require substantial computational resources and prolonged operation durations, making them impractical for real-time sensing services. This limitation becomes particularly evident when dealing with terminals that have limited computing capabilities.

To resolve the preceding issues, the following methods can be used:

1. Two-step spectral analysis

   This method consists of the following two steps:

   - Step 1: Two-dimensional Fourier transform (2D-FFT) as well as signal detection on the 2D-FFT processing result. This makes it possible to obtain a range unit and velocity unit where the sensing target is located.

   - Step 2: One-dimensional MUSIC operation on the column of the range unit or the row of the velocity unit in which the sensing target is located. This makes it possible to obtain the high-resolution range and velocity measurements.

The two-step spectral analysis method is similar to the 2D MUSIC algorithm in its ability to acquire the range and velocity resolution of a sensing target. However, it distinguishes itself by requiring significantly fewer computational resources and shorter operation times due to the utilization of one-dimensional MUSIC operations. The results of this method are presented in Figure 14b, showing that the resolution achieved using this method is significantly greater than that of the 2D-FFT technique.

(a) Results of step 1　　　　　(b) Results of step 2
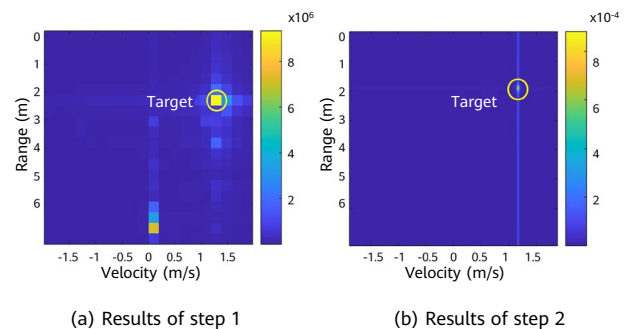
Figure 14 Processing results of the two-step spectral analysis method

2. Amplitude-phase method

   As the name suggests, this method leverages the amplitude and phase characteristics of the sensing signal to extract the fractional resolution units of range and velocity measurements. In the 2D-FFT processing result, the signal energy of the sensing target typically falls into two adjacent resolution units in the range or velocity dimension. For example, the velocity resolution of the prototype described in this section is 0.1875 m/s. Therefore, the signal energy of the sensing target moving at a velocity of 1.2 m/s resides within the adjacent resolution units of 1.125 m/s and 1.3125 m/s. The signal amplitudes

and phases on two adjacent resolution units have the following relationships:

$$\frac{|X[k]|}{|X(k+1)|} \approx \frac{\mathrm{sinc}(\Delta)}{\mathrm{sinc}(\Delta-1)} \qquad (10)$$

$$\angle X[k] - \angle X[k+1] \approx \pi \qquad (11)$$

Where, $X(k)$ and $X(k+1)$ respectively represent complex signals on the kth and (k+1)th resolution units, $|X|$ and $\angle X$ respectively represent amplitude and phase of the complex-value X, and Δ represents the values of the fractional resolution units. Through signal detection based on the 2D-FFT processing result, two adjacent resolution units corresponding to a sensing target can be determined based on the phase relationship. Subsequently, a fractional resolution unit can be calculated according to the amplitude relationship. The simulation result shown in Figure 15 demonstrates using the amplitude-phase method can achieve a significant improvement in resolution performance compared with using only the 2D-FFT method.
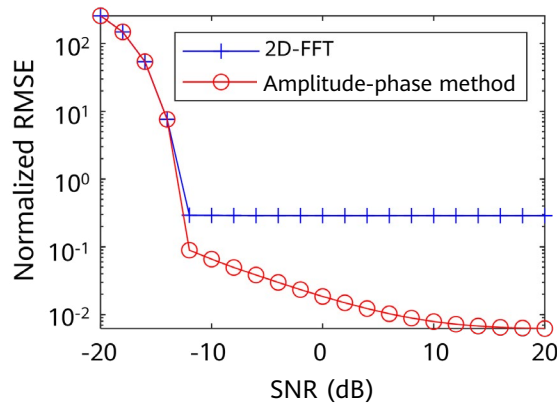
Figure 15 Comparison between the amplitude-phase method and the 2D-FFT method

5 Challenges

Mobile communication has evolved over time, progressing from 1G to 5G, and has undergone numerous verifications to establish its maturity. ISAC is a promising technology expected to unlock further advancements in wireless networks. However, it faces some implementation challenges.

5.1 Uncertainty/Ambiguity of Sensing Results

A communication service ensures the correctness of transmitted information by conducting a cyclic redundancy check (CRC) at the receiving end, thereby offering a reliable and trustworthy service. ISAC is used to detect targets in complex electromagnetic environments, such as indoor or dense urban areas. In such environments, there are usually multiple sensing targets and numerous non-target scatterers, posing challenges in accurately modeling the sensing target as a point target model. Therefore, once the underlying technical issues pertaining to the transmission, reception, and processing of sensing signals are successfully addressed, accurately identifying a sensing target from complex point cloud data becomes a challenging task, particularly in the case of an unconnected sensing target.

5.2 Diversified ISAC Use Cases

3GPP's feasibility study for 5G ISAC outlines 32 ISAC use cases [9]. It is evident that there will be a wider range of ISAC use cases specifically tailored for 6G in the future, including hundreds of to-business (2B) and to-consumer (2C) use cases. Each case involves distinct sensing algorithms, measurement quantities, results, and prior information. Devising individual optimization policies for each ISAC use case during solution design and standards development will be challenging. Therefore, we need to design a unified technical framework to meet the needs of hundreds of ISAC use cases. This is the first challenge brought by the diverse nature of sensing use cases. The second challenge arises from possible restrictions in the market volume and deployment area for each use case, hindering the establishment of a robust ISAC industry on a large scale. This limitation affects the investment and deployment of the ISAC networks in the future.

5.3 Business Model, Privacy, and Charging Issues

Many ISAC use cases are related to 2B services, such as smart transportation and low-altitude drone monitoring. In most cases, the provisioning of 2B services requires cooperation and approval from various departments. For example, in smart transportation scenarios, the deployment of ISAC base stations or terminals at intersections requires auxiliary devices to be deployed at traffic light stations, requiring authorization from transportation departments. Similarly, low-altitude drone monitoring requires approval from departments responsible for managing low-altitude operations within the city.

Compared with camera-based sensing, wireless sensing offers enhanced privacy within the LOS. Whereas camera-based sensing captures images that can be visually recognized, wireless sensing only collects patterns of sensing measurement quantities. However, wireless sensing has a broader range of detection than non-RF sensing. For example, it can detect human activities beyond the LOS (e.g., behind closed doors). Therefore, ensuring security and privacy remains a significant challenge in the field of wireless sensing. The security and privacy mechanism must effectively safeguard the privacy information acquired from sensing targets within the owner's authorization scope. It should also ensure E2E security and authentication for the wireless sensing service to prevent any potential attacks or tampering. For example, privacy and security measures can be implemented through techniques such as conducting fuzzing processing on sensing information or encrypting sensing signals.

In the ISAC business model, sensing-related charging is a crucial component that involves the charging objects, standards, and schemes. It is necessary to consider both the charging for the demand side of sensing services and the payment for the nodes that transmit or receive sensing signals (such as UEs). To ensure the future commercial viability of ISAC, it is necessary to establish flexible, reasonable, innovative, and open business models and ecosystems that encompass charging policies.

5.4 Integrated Design of External and Internal ISAC Services [23]

The 6G system can offer an ISAC service, referred to as the "external ISAC service," for objects (such as 2B and 2C) outside the 6G system. In addition to this, the sensing capability of the 6G system can also be utilized internally, serving as an "internal ISAC service" for the system itself. An example of the internal ISAC services is sensing-assisted communications, such as location-aware beamforming. The provisioning of internal ISAC services is mainly from technical considerations. These services are specifically designed to improve the performance and efficiency indicators of the 6G system, thereby enabling the system to provide better external services. The provisioning of external ISAC services is mainly from business considerations. Further study needs to be conducted on whether and how to integrate external and internal ISAC services.

6 Conclusion

Both wireless communications and sensing are developed based on the electromagnetic wave theory. At the transmitter end, source information is encoded onto an electromagnetic wave signal through modulation. As the signal propagates, it interacts with the wireless environment and therefore also carries information about the surrounding environment through modulation. Similar to how wireless networks underwent a significant growth phase upon their emergence in the communications industry, they are again poised to undergo another significant phase of progress with the advent of ISAC. This development is expected to result in a wide range of applications for sensing services. Performance indicators for sensing include common indicators (e.g., service latency and refreshing rate) and encompass indicators specific to a particular use case, such as accuracy, resolution, detection accuracy, and missing detection. Efficiency indicators include spectral efficiency, energy efficiency, and cost efficiency. Research and development, as well as system design, related to sensing should consider both types of indicators. In ISAC, key technologies include ISAC waveform and signal design, multiple-antenna technology for sensing, CoMP ISAC, elimination of non-ideal sensing factors, mobility management, link adaptation, and more.

vivo has developed three ISAC prototypes: bistatic sensing prototype for respiration monitoring, bistatic sensing prototype for CoMP trajectory tracking, and monostatic sensing prototype for target ranging and velocity measurement. These prototypes demonstrate the feasibility of ISAC and verify the performance of the involved technologies in eliminating certain non-ideal factors. Finally, we address some challenges in the implementation of ISAC, including uncertainty/ambiguity of sensing results, vast diversity of sensing use cases, ISAC business model, privacy and charging issues, and integrated design of external and internal sensing services.

References

[1] IMT-2030(6G) Promotion Group. "White Paper on 6G Typical Scenarios and Key Capabilities [R]," 2022.

[2] Next G Alliance. "6G Applications and Use Cases[R]," 2022.

[3] Hexa-X. 6G Vision, "Use Cases and Key Societal Values[R]," 2021.

[4] vivo Communications Research Institute, "6G Vision, Requirements and Challenges [R]," 2020.

[5] ITU-R recommendation, "Framework and overall objectives of the future development of IMT for 2030 and beyond," June 2023.

[6] Dajie Jiang, Jiang Yao, Jianzhi Li, *et al*., "Key technologies and challenges for integrated sensing and communication[J]," Mobile Communications, 2022,46(5): 69-77.

[7] vivo Communications Research Institute. "6G Services, Capabilities and Enabling Technologies [R]," 2022.

[8] China Institute of Communications. "Integrated Sensing, Communications and Computation Network Frontier Report [R], "2022.

[9] 3GPP TR 22.837 V19.0.0, "Feasibility Study on Integrated Sensing and Communication," June 2023.

[10] Kun Yang, Dajie Jiang and Fei Qin, "Overview of the Intelligent Surface for 6G Communications [J]. Mobile Communications," 2020, 44(6): 70–74.

[11] vivo, China Telecom, China Mobile, *et al*., "System Architecture and Key Technologies of Integrated Sensing and Communication [R]," 2023.

[12] Daqing Zhang, Fusang Zhang, Dan Wu, *et al*., "Practical Issues and Challenges in CSI-based Integrated Sensing and Communication[J]. Mobile Communications," 2022, 46(05): 9–16.

[13] Mugen Peng, Xiqing Liu, Zile Liu, *et al*., "Principles and techniques in communication and sensing integrated 6G systems [J]. Control and Decision," 2023, 38(1):22–38.

[14] Andrew Zhang J, Wu K, Huang X, *et al*., "Integration of Radar Sensing into Communications with Asynchronous Transceivers[J]," 2022,60(11): 106–112.

[15] Zeng Y, Wu D, Xiong J, *et al*., "FarSense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas[J]," in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2019, 3(3): 1–26.

[16] Pearson, R.K., Neuvo, Y., Astola, J. *et al*., "Generalized Hampel Filters," in *EURASIP J. Adv. Signal Process*, 2016, 87 (2016).

[17] R. W. Schafer, "What Is a Savitzky-Golay Filter? [Lecture Notes]," in *IEEE Signal Processing Magazine*, vol. 28, no. 4, pp. 111–117, July 2011.

[18] Wong, Ian C., and Brian L. Evans. "Sinusoidal modeling and adaptive channel prediction in mobile OFDM systems," in *IEEE Transactions on Signal Processing*, 56.4 (2008): 1601–1615.

[19] Zeng Y, Wu D, Xiong J, *et al*., *FarSense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas[J]*, in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2019, 3(3): 1–26.

[20] Songqian L, Chenhao L, Aimin T, *et al*., "Integrating Passive Bistatic Sensing into mmWave B5G/6G Networks: Design and Experiment Measurement [C],"in *2023 IEEE International Conference on Communications (ICC). IEEE*, 2023: 1–6.

[21] Li, Xiang, *et al*., "IndoTrack: Device-free indoor human tracking with commodity Wi-Fi," in *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1.3 (2017): 1–22.

[22] S. Ding, *et al*., "Integrated Sensing and Communication: Prototype and Key Processing Algorithms," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, pp. 73–78, May 2023.

[23] Dajie Jiang, Yannan Yuan, Tong Zhou, *et al*., *Services, System Architecture and Key Technologies for 6G Integrated Communication*, in *Sensing and Computing[J], Mobile Communications*, 2023,47(03): 2–13.

Standardization of WLAN Sensing: IEEE 802.11bf

Rui Du [1], Narengerile [1], Zhuqing Tang [1], Mengshi Hu [1], Yiyan Zhang [1], Yan Xin [2], Junghoon Suh [2], Osama AboulMagd [2], Stephen McCann [2], Michael Montemurro [2], Rojan Chitrakar [3], Lei Huang [3], Xiao Han [1]

[1] Wireless Technology Lab
[2] Ottawa Advanced Wireless Technology Lab
[3] Singapore Short-Range Wireless Lab

Abstract

With the development of integrated sensing and communication, wireless local area network (WLAN) or wireless fidelity (Wi-Fi) technology has been successfully utilized to enable sensing functionalities such as detection, localization, and recognition. However, the existing Wi-Fi standards were primarily developed to ensure communication quality and improve communication performance and may not be able to meet the stringent sensing requirements of emerging applications. To address these limitations, a new task group (TG), IEEE 802.11bf, has been established by the IEEE 802.11 working group. The objective of this group is to create a new amendment to the existing Wi-Fi protocols that maximizes support for sensing requirements in various applications. This paper provides a comprehensive overview of the first international sensing standard[1], IEEE 802.11bf (WLAN sensing).

Keywords

integrated sensing and communication, WLAN, sensing, IEEE 802.11bf

---

[1] To ensure the consistency with the standard naming, WLAN sensing is used for content related to the IEEE 802.11bf, and Wi-Fi sensing is used for content not related to the IEEE 802.11bf.

1 IEEE 802.11bf: Definition, Scope, Formation, and Timeline

Since its standardization in 1997, wireless fidelity (Wi-Fi) has become a widely used technology worldwide. According to the Wi-Fi Alliance (WFA), the global economic value provided by Wi-Fi is expected to reach nearly US$5 trillion by 2025. In recent years, Wi-Fi sensing has gained increasing attention from the academic community and industry as an emerging technology. However, existing academic research on Wi-Fi sensing mainly focused on solutions in scenarios with individual or a few devices without fully using the advantageous features of Wi-Fi. For instance, many studies still use channel state information (CSI) of 20/40 MHz, while Wi-Fi 6 already supports a bandwidth of 160 MHz. Moreover, existing Wi-Fi standards lack a dedicated sensing procedure, making it inefficient to establish a sensing application. To address these limitations and better support various Wi-Fi sensing applications while minimizing the impact of sensing on Wi-Fi communication performance, the IEEE 802.11 working group established a new task group (TG), IEEE 802.11bf, in 2020. This TG aims to create a new amendment to the WLAN standard by 2025.

1.1 Definition and Scope

IEEE 802.11bf is a technology that uses wireless signals received from 802.11 stations (STAs) to determine the features of intended targets, including range, velocity, angle, target detection, and image. This technology can be used to detect objects, humans, animals, and environments in various scenarios, such as rooms, houses, vehicles, and enterprises. The Project Authorization Request (PAR) and Criteria for Standards Development (CSD) specify that

IEEE 802.11bf aims to develop an amendment that defines modifications to the IEEE 802.11 medium access control (MAC), the directional multi-gigabit (DMG) physical (PHY), and enhanced DMG (EDMG) PHY. This amendment enhances wireless local area network (WLAN) sensing operation in license-exempt frequency bands between 1 GHz and 7.125 GHz and above 45 GHz. This amendment aims to support various sensing applications in everyday life while ensuring backward compatibility and coexistence with existing or legacy IEEE 802.11 STAs operating in the same band.

1.2 Formation and Timeline

The IEEE 802.11 Wireless Next-Generation Standing Committee (WNG SC) first discussed forming a WLAN sensing project in July 2019. During this meeting, the feasibility of WLAN to support sensing use cases and their requirements were evaluated. In October 2019, the WLAN sensing topic interest group (TIG) was established, followed by the approval for the formation of the WLAN sensing study group (SG) by the New Standards Committee (NesCom) in the IEEE Standards Association (SA) in November 2019. The WLAN sensing TIG and SG team members worked together to develop the PAR and CSD, which were approved, leading to the official establishment of a new TG, IEEE 802.11bf, in September 2020.

Figure 1 illustrates the timeline of the progress made toward the IEEE 802.11bf amendment. IEEE 802.11bf released draft 1.0 in January 2023 and, shortly after, completed the initial letter ballot (LB) and developed the basic framework for the WLAN sensing procedure. Two recirculation LBs were conducted, and the submission of draft amendment 4.0 to the IEEE SA for a subsequent SA ballot is expected in March 2024. Finally, after the recommendation of the Standards
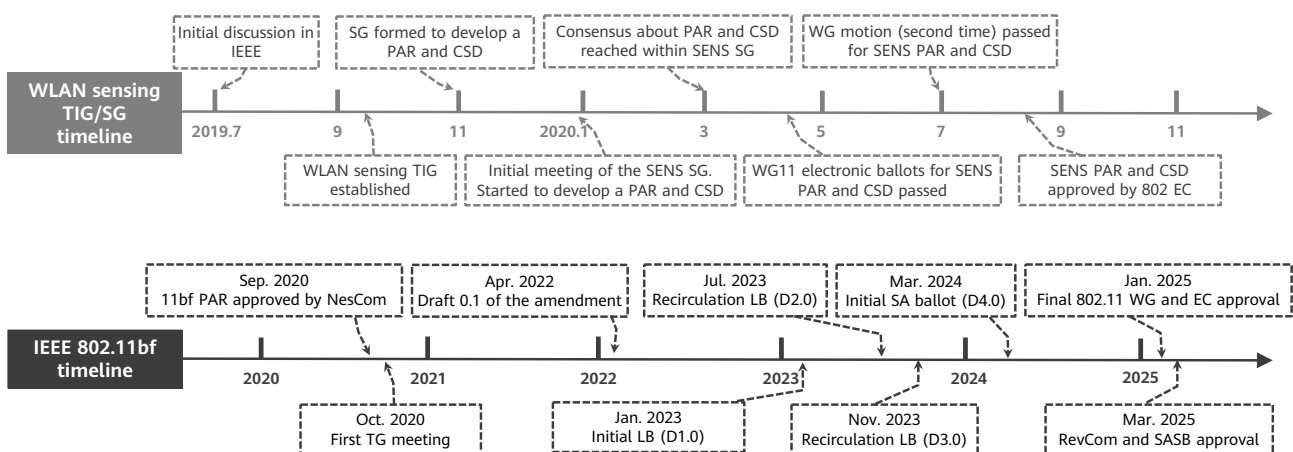

Figure 1 Timeline and progress for IEEE 802.11bf

Review Committee (RevCom) and the approval of the IEEE SA Standards Board (SASB), the project is expected to be published as an IEEE 802.11bf amendment specification in March 2025.

2 Use Cases

This section outlines the WLAN sensing use cases and identifies the key performance indicator (KPI) requirements. Due to space constraints, this section provides a brief overview of four use cases: presence detection, activity recognition, target localization and tracking, and vital signs monitoring. For a more comprehensive understanding, please refer to the IEEE 802.11bf use cases in [1].

2.1 Presence Detection

In a typical indoor setting, human presence detection plays a crucial role in creating a smart home environment. According to [1], it is important to consider a maximum range coverage of 10–15 m for effective presence detection, with a specific value being selected based on the room size. Moreover, in multi-person environments, it is essential to have a range resolution of 0.5–2 m and an accuracy of 0.2 m for precise presence detection.

2.2 Activity Recognition

Human activity recognition (HAR) is crucial in the field of human-computer interaction (HCI). Since human activities can distort wireless channels, it is possible to extract patterns such as the Doppler spectrum and target speed from the estimated CSI to detect and recognize human daily activities. The IEEE 802.11bf standard recommends using the sub-7 GHz band for recognizing large motions (such as full-body motion) and the 60 GHz band for recognizing small

motions (such as finger motion). The Doppler resolution for micro-Doppler-based recognition varies depending on the specific scheme used.

2.3 Target Localization and Tracking

Localization is the process of determining the position of a target, whereas tracking aims to confirm the target's movement trajectory by observing changes in its position over time. The geometric model-based human tracking method offers better generalization in different environments. This method utilizes the super-resolution joint multiparameter estimation algorithm to enhance multipath resolution, enabling accurate localization and tracking of human targets. However, the accuracy of parameter estimation may still be constrained by the limited number of antennas and bandwidth. The IEEE 802.11bf standard recommends a range resolution of 0.5 m and a range accuracy of 0.2 m for human target localization and tracking.

2.4 Vital Signs Monitoring

Respiration and heartbeat estimation are commonly used techniques for analyzing human health conditions, including sleep quality analysis. By observing the vibration pattern of the phase and/or amplitude of the CSI, it is possible to estimate the rate of respiration or heartbeat. Regarding heartbeat detection, the IEEE 802.11bf standard recommends using the 60 GHz band. It is important to highlight that vital signs like respiration and heartbeat can serve as input features to aid in the detection of human static presence.

Table 1 lists the KPIs of the preceding four use cases.

Figure 2 Main applications of WLAN sensing

Table 1 KPIs of typical use cases

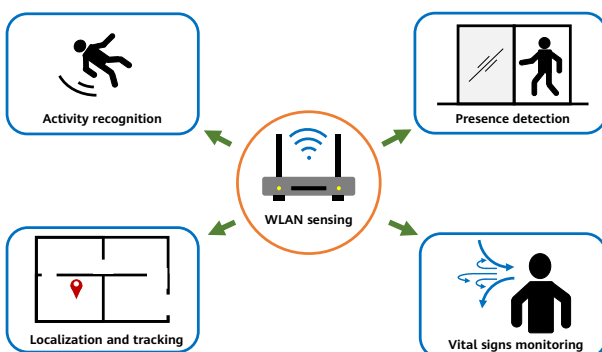| Use Case	KPI or Recommended Frequency Band
Presence detection	Range coverage: 10–15 m Range resolution: 0.5–2 m Range accuracy: < 0.2 m
Activity recognition	Large motion: sub-7 GHz Small motion: 60 GHz
Target localization and tracking	Range resolution: 0.5 m Range accuracy: < 0.2 m
Vital signs monitoring	Heartbeat detection: 60 GHz

3 IEEE 802.11bf: Sensing Framework and Procedures

This section outlines the main framework and procedures of the IEEE 802.11bf standard. It also provides the fundamental definitions and typical procedures of the transceiver roles in sub-7 GHz sensing and 60 GHz DMG sensing.

3.1 Transceiver Roles and Specific Definitions

• Sensing initiator: an access point (AP) or non-AP STA that initiates a sensing procedure by sending a request frame for sensing measurement.

• Sensing responder: an STA that participates in a sensing procedure initiated by a sensing initiator, responding with a response frame for sensing measurement.

• Sensing transmitter: an STA that sends a sensing PLCP protocol data unit (PPDU) during a sensing procedure.

• Sensing receiver: an STA that receives the sensing PPDU sent by the sensing transmitter and obtains the measurement result.

3.2 Sub-7 GHz Sensing

The IEEE 802.11ax and IEEE 802.11be amendments facilitate sub-7 GHz sensing, supporting a bandwidth of up to 320 MHz and accommodating a maximum of eight antennas. These amendments enable high-resolution sensing measurements across the 2.4 GHz, 5 GHz, and 6 GHz bands, providing wide coverage for various sensing applications.

• Basic sensing procedure

Figure 3 illustrates the sub-7 GHz sensing procedure, which encompasses the sensing capabilities exchange, sensing measurement session, sensing measurement exchange, and sensing measurement session termination. Before sensing measurement session, sensing-capable STAs will exchange their sensing capabilities. The setup of sensing measurement session involves configuring the attributes and parameters of the sensing STAs for specific sensing applications. This includes determining the STA's role and the type of measurement report, with different sets of operational attributes identified by measurement session IDs. In the sensing measurement exchange, the sensing STA performs the actual sensing and reports the sensing measurement result based on
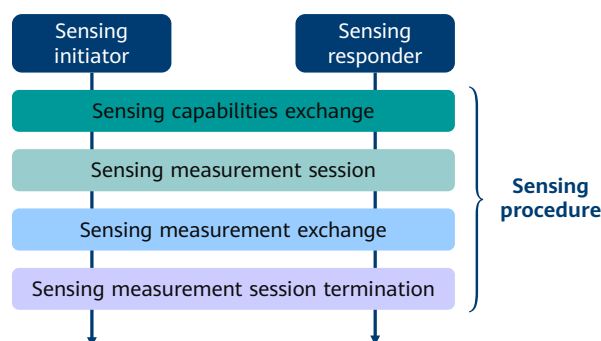
Figure 3 WLAN sensing procedure

the parameters configured in the previous phase. Each sensing measurement exchange is uniquely identified by a measurement exchange ID. If sensing tasks are no longer required, the sensing STAs can terminate the existing sensing measurement session.

Sub-7 GHz sensing supports two types of sensing measurement exchanges: trigger-based (TB) and non-TB. A TB sensing measurement exchange is initiated by an AP, whereas a non-TB sensing measurement exchange is initiated by a non-AP STA.

The TB sensing measurement exchange is a trigger-based variant of the sensing measurement exchange, where the AP is the sensing initiator and one or more non-AP STAs are the sensing responders. This allows for one-to-many sensing between an AP and multiple non-AP STAs. The TB sensing measurement exchange may comprise several phases, including polling, null data packet announcement (NDPA) sounding, trigger frame (TF) sounding, and reporting. It is important to note that the order of these phases may vary. Polling is performed to check the availability of the sensing responder that is expected to participate in the sensing measurement exchange. NDPA sounding is performed from the sensing initiator to the sensing responder, whereas TF sounding is implemented from the sensing responder to the sensing initiator. In the reporting phase, sensing measurement results from the NDPA sounding phase are reported, and the corresponding sensing measurement reporting can be either immediate or delayed.

Figure 4 shows an example of a TB sensing measurement exchange with one initiator AP and five non-AP responder STAs. In the polling phase, the initiator AP polls five STAs, where STA1, STA2, STA3, and STA 4 send a clear-to-send-to-self (CTS-to-self) frame to confirm that they can participate in the sensing measurement exchange. Based on roles in the measurement setup phase, STA3
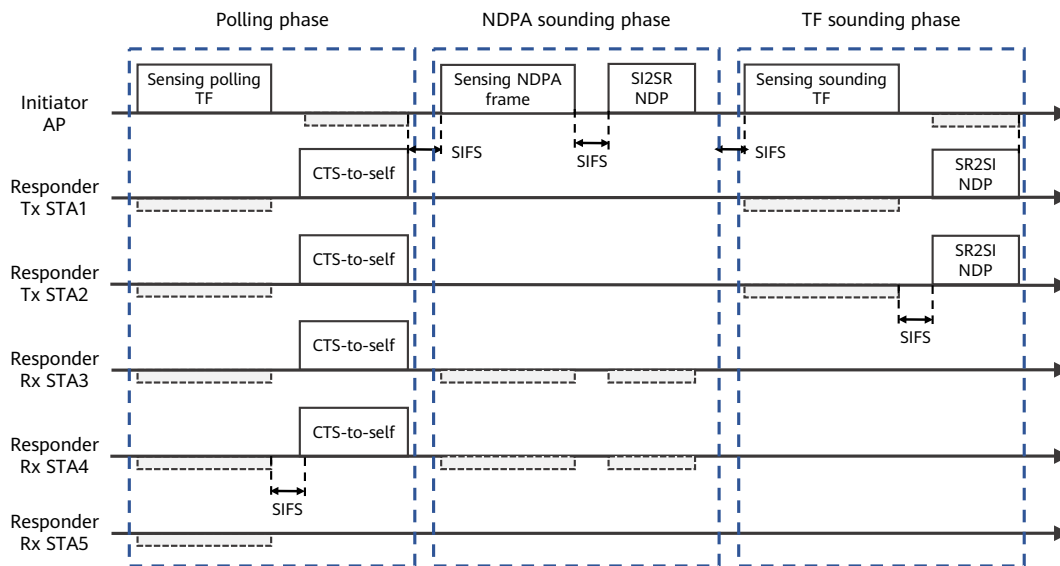
Figure 4 Example of TB sensing measurement exchange

and STA4 participate in the TF sounding phase, and STA1 and STA2 participate in the NDPA sounding phase. Note that the example in Figure 4 does not include a feedback reporting phase.

The non-TB sensing measurement exchange is a variation of the sensing measurement exchange that does not rely on triggers. It is used when a non-AP STA is the sensing initiator, and an AP is the sensing responder, allowing for one-to-one sensing between a non-AP STA and an AP. In Figure 5, the non-AP STA first sends an NDPA frame to the AP. Then, after going through a short interframe space (SIFS), the non-AP STA sends a sensing initiator to sensing responder (SI2SR) NDP frame to the AP. After receiving the SI2SR NDP frame and going through an SIFS, the AP sends a sensing responder to sensing initiator (SR2SI) NDP frame. This procedure allows the implementation of bidirectional sounding (i.e., SI2SR and SR2SI) in the non-TB sensing measurement exchange.

- Sensing by proxy

  Compared to non-TB sensing, which involves a non-AP initiator STA and a responder AP (only implementing

one-to-one sensing between a non-AP STA and an AP), the sensing by proxy (SBP) procedure allows an SBP initiator (non-AP STA) to request an SBP responder (AP) to perform WLAN sensing on its behalf. This enables the SBP initiator to obtain sensing measurement results between the AP and multiple non-AP STAs. By utilizing the SBP procedure, a non-AP STA can obtain sensing measurement results from multiple STAs in the environment, which can better support corresponding sensing applications. Additionally, the SBP procedure allows the SBP initiator to obtain sensing measurement results between an AP and one or more responder STAs that exceed its hardware limits. Figure 6 shows an example of the SBP procedure.

To establish an SBP procedure, the SBP initiator begins by sending an SBP request frame to an SBP-capable AP. Within the SBP request frame, the SBP initiator can indicate requirements for the sensing measurement setup to be performed by the SBP responder AP. The SBP

Figure 5 Non-TB sensing measurement exchange
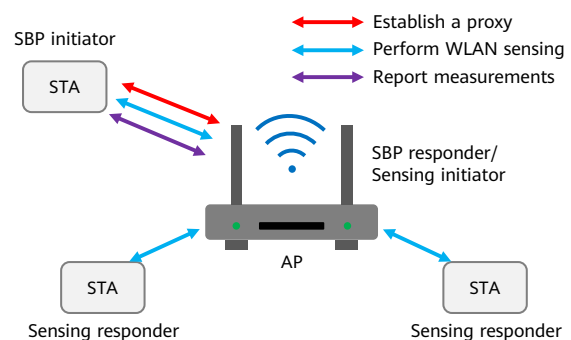
Figure 6 SBP procedure
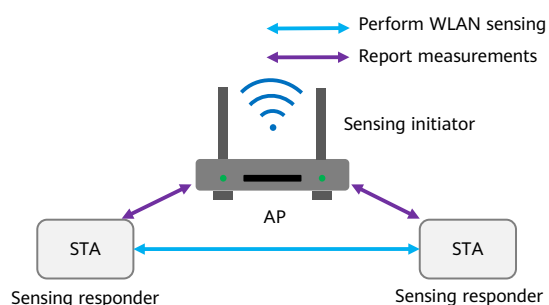
Figure 7 SR2SR sensing procedure

initiator has the capability to provide the SBP responder AP with information regarding the number of sensing responders and their preferred responder list through the SBP request frame. Additionally, the SBP initiator has the option to decide whether or not to participate in the sensing procedure. Upon receiving the SBP request frame, the SBP responder AP uses the recommended parameters to initiate sensing and obtain the sensing measurement results. After the measurements are complete, the SBP responder AP collects the sensing measurement results from the sensing responders and provides the results to the SBP initiator.

- SR2SR sensing

  The sensing measurement between an AP and a non-AP STA can be implemented through a basic sensing procedure or an SBP procedure. To obtain sensing measurements between non-AP STAs, IEEE 802.11bf provides a sensing responder to sensing responder (SR2SR) procedure, further enhancing the diversity of transmit and receive in sensing measurements. Figure 7 illustrates an example of the SR2SR sensing procedure.

  To initiate the SR2SR sounding phase, the sensing responder will be designated an SR2SR transmitter role and/or receiver role during the sensing measurement setup. In each sensing measurement exchange, the AP schedules one sensing responder as a transmitter and other sensing responders as receivers to perform the SR2SR sensing procedure. Once the measurements are completed, the sensing receivers can report the SR2SR sensing measurement results to the sensing initiator.

3.3 DMG Sensing

DMG sensing is implemented based on the IEEE 802.11ad and IEEE 802.11ay standard amendments. Unlike sub-7 GHz sensing, DMG sensing offers a wider channel bandwidth

of 2.16 GHz per channel and operates at a higher carrier frequency of 60 GHz. This allows for short-range sensing applications that require higher range resolution and Doppler sensitivity.

- DMG sensing procedure

  The DMG sensing procedure is similar to the sub-7 GHz sensing procedure, but there are three main differences. The first difference is the exchange of beam direction information. IEEE 802.11ad and IEEE 802.11ay use directional transmission and reception to compensate for the severe path loss of 60 GHz. Therefore, during the DMG sensing measurement setup, the DMG sensing initiator needs to exchange the directions of transmit and receive beams with the DMG sensing responder. By negotiating these beam directions, the DMG sensing initiator can establish measurements to detect objects in a specific direction in some applications. Another difference between DMG sensing and sub-7 GHz sensing is the concept of burst. To better support Doppler measurement in sensing, the concept of burst is introduced to DMG sensing and is defined as a set of DMG sensing measurement exchanges with equal time intervals. As shown in Figure 8, the DMG sensing initiator can configure different inter-burst interval values and intra-burst interval values in different DMG sensing measurement setups for various applications. The last difference between DMG sensing and sub-7 GHz sensing is that the DMG sensing measurement exchanges have a uniform mode, regardless of whether the sensing measurement is initiated by an AP or a non-AP STA. This is different from the TB and non-TB modes of sub-7 GHz sensing.

  Depending on the number and roles of the devices involved in sensing, there are several types of DMG sensing. These include:

  – Monostatic sensing: In this type, the sensing transmitter and sensing receiver are the same STA.

  – Bistatic sensing: Here, the sensing transmitter and the sensing receiver are two distinct STAs.

  – Monostatic sensing with coordination: In this case, one or more STAs performing monostatic sensing are coordinated by a personal basic service set (PBSS) control point (PCP)/AP.

  – Bistatic sensing with coordination: Similar to monostatic sensing with coordination, but here one or more STAs performing bistatic sensing are coordinated by a PCP/AP.
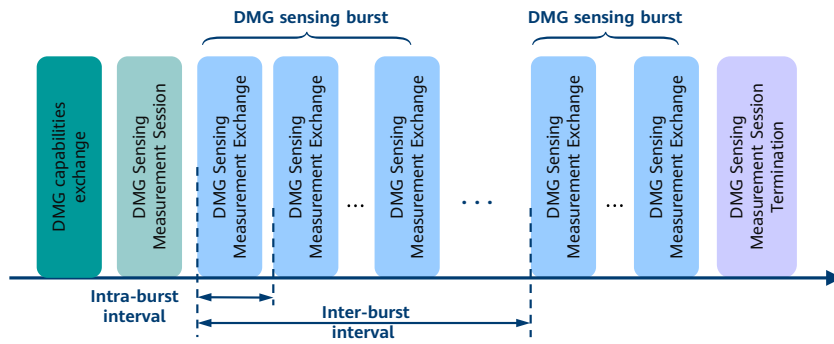
Figure 8 DMG sensing procedure

– **Multistatic sensing:** In this type, the sensing transmitter and more than one sensing receivers are distinct STAs.

– **Passive sensing:** Sensing is performed by passively receiving communication packets/signals, such as PPDUs carrying beacon frames.

These different types of DMG sensing can be configured according to the specific requirements of various sensing applications. Please refer to Figure 9 for a visual representation.

- **DMG sensing by proxy**

  DMG sensing also supports SBP. In the DMG SBP procedure, the SBP initiator configures different DMG sensing types through the DMG SBP responder AP to obtain the required sensing measurement results.
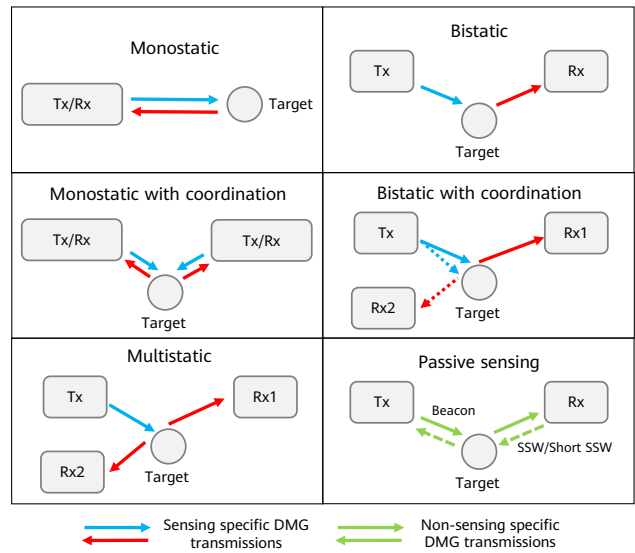
Figure 9 DMG sensing types
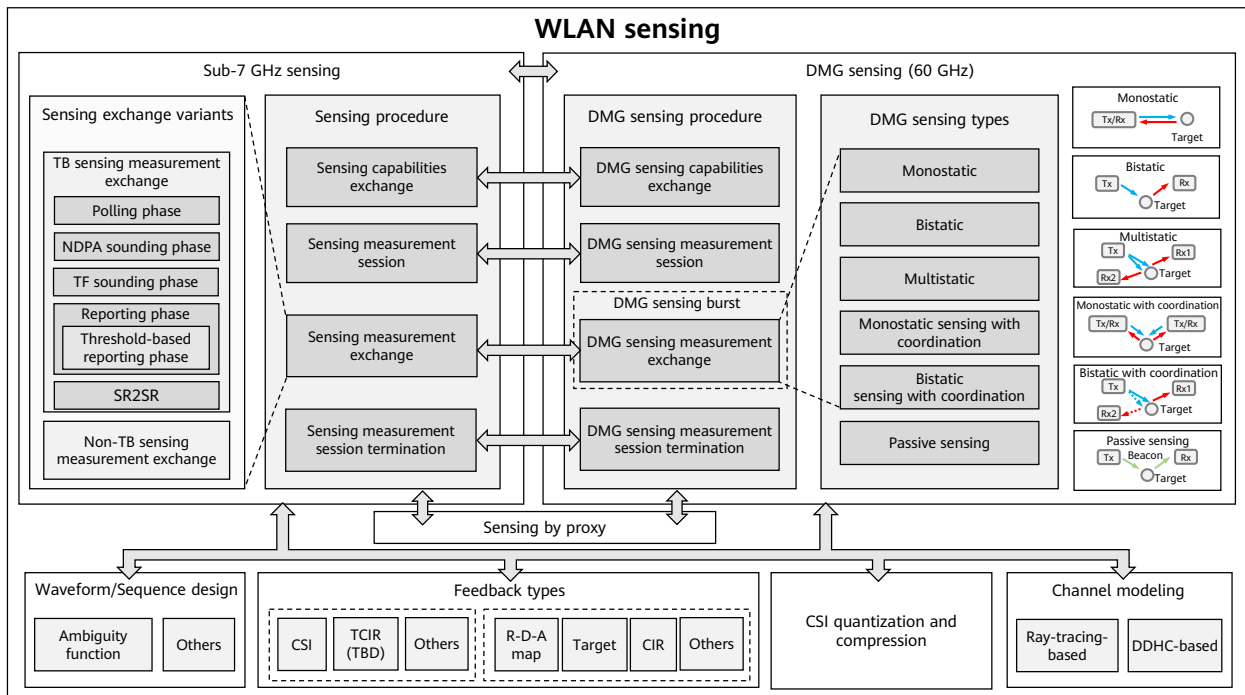
Figure 10 Overview of the WLAN sensing procedures

- DMG SR2SR sensing

The details of DMG SR2SR sensing are currently being discussed in the IEEE 802.11bf standard amendment, and the DMG SR2SR sensing procedure may be included in the final IEEE 802.11bf draft amendment.

Figure 10 illustrates the IEEE 802.11bf procedures for sub-7 GHz sensing and DMG sensing. By utilizing the sensing procedures described above, IEEE 802.11bf can implement sensing measurements between each pair of sensing STAs, leveraging transmit/receive diversity to support different sensing applications. For a more comprehensive introduction to IEEE 802.11bf, see [2].

4 IEEE 802.11bf: Candidate Technologies

During discussions within the IEEE 802.11bf standard amendment, experts proposed several candidate technologies. This section outlines some of the most important ones that were discussed. It is worth noting that while two candidate technologies are listed here, the quantization and feedback of the CSI matrix have already been determined in the standard, whereas the other one technology is still under discussion. The quantization and feedback of the CSI matrix rely on foundational technologies and are therefore also described and emphasized in this section.

4.1 CSI Feedback Types and Quantization

Since the release of IEEE 802.11ac, compressed CSI has become the sole explicit feedback type used by mainstream sub-7 GHz in IEEE 802.11 standard amendments such as the IEEE 802.11ax and IEEE 802.11be. However, the compressed CSI is primarily designed for MIMO precoding and lacks certain channel information compared to the CSI matrix, making it inadequate for supporting sensing. Therefore, sensing measurement feedback has become a crucial topic of discussion in IEEE 802.11bf. After extensive deliberation, IEEE 802.11bf has adopted the CSI matrix as the only explicit sensing measurement feedback type in sub-7 GHz sensing. In this feedback, the per subcarrier scaling has been replaced with per link scaling to mitigate the impact of unbalanced antennas during the quantization and feedback phases. For DMG sensing, IEEE 802.11bf offers three types of feedback: channel measurement, Range-Doppler-Angular map (R-D-A map), and target-related parameters (i.e., target detection results based on the R-D-A map).

4.2 Multi-Link Sensing

In IEEE 802.11bf, sensing primarily involves a single link of transmission and reception between STAs. However, with the upcoming commercial use of Wi-Fi 7 in the next few years, multi-link (ML) STAs will emerge in the market. These ML STAs will support simultaneous transmission and reception of multiple links. To further enhance sensing performance, ML is a key technology that will need to be discussed in the IEEE 802.11bf and future sensing standards. ML-sensing-capable STAs can operate at various frequencies (such as 2.4 GHz, 5 GHz, and 6 GHz) to perform measurements simultaneously and provide diversity gains in observations across frequencies. In certain scenarios, coordinating different links can offer better support for both sensing and communication. For instance, in some scenarios where a higher sensing frequency is required, sensing measurement feedback can be performed on a link separate from the "sensing link." This approach reduces the occupation of the air interface of the sensing link and effectively enhances the sensing frequency.

5 Conclusion and Future Outlook

As the work of IEEE 802.11bf progresses, the WFA is also engaging in discussions regarding initiating Wi-Fi sensing certification. We anticipate that Wi-Fi sensing products will gradually emerge in the market in the coming years, enabling efficient support for various sensing applications across different vendors. This development will further drive commercial adoption and academic exploration of existing Wi-Fi sensing technologies.

References

[1]   A. Kasher, A. Eitan, S. Trainin, Y. Sun, and R. Du, "WiFi sensing use cases," IEEE 802.11-21/1712r2, Jan. 2021. [Online]. Available: https://mentor.ieee.org/802.11/dcn/20/11-20-1712-02-00bf-wifi-sensing-use-cases.xlsx

[2]   R. Du, H. Xie, M. Hu, Narengerile, Y. Xin, S. McCann, M. Montemurro, T. X. Han, and J. Xu, "An overview on IEEE 802.11bf: WLAN sensing," 2022. [Online]. Available: https://arxiv.org/abs/2207.04859
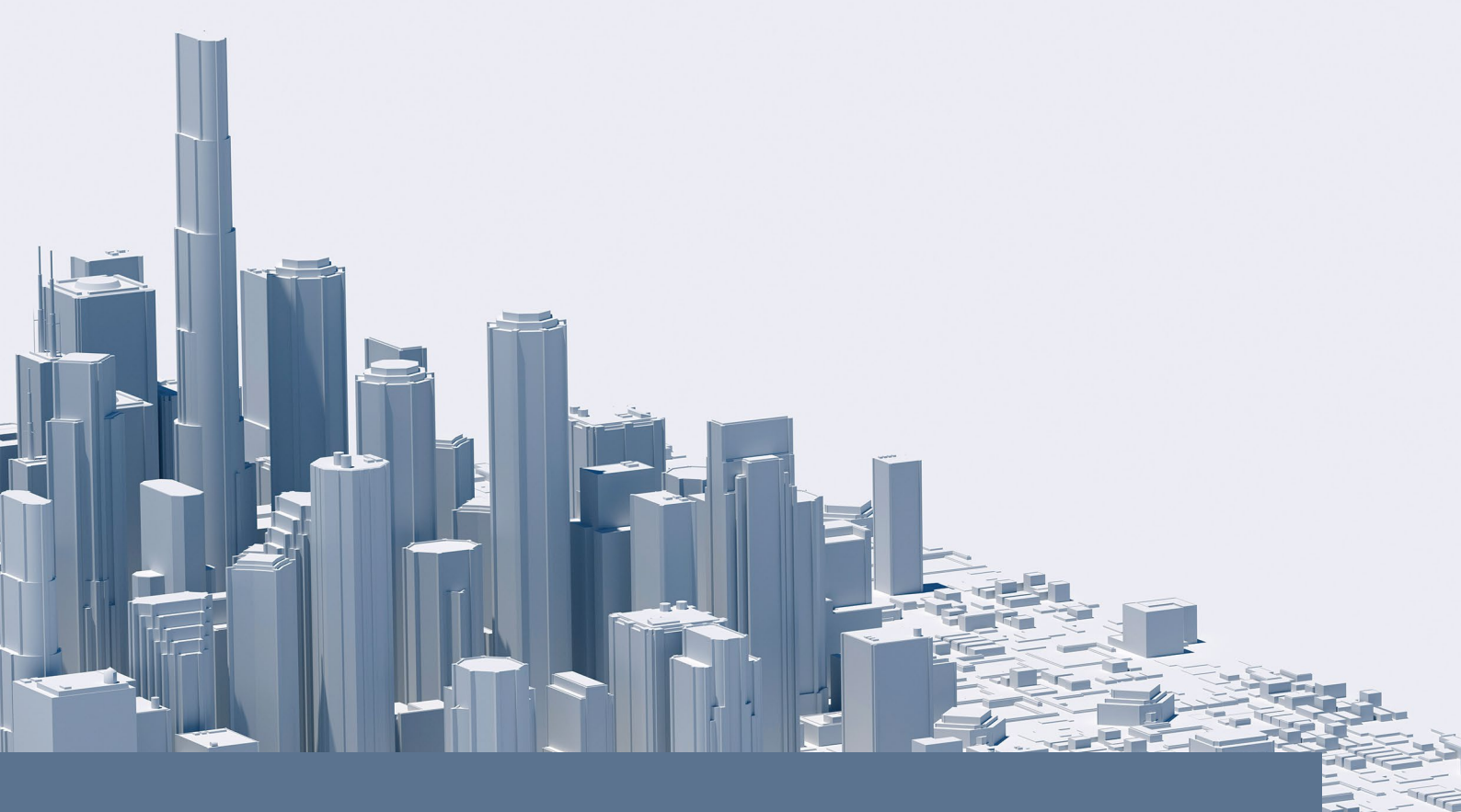
6G Integrated Sensing and Communication — Sensing-assisted Environment Reconstruction and Channel Reconstruction

Zhi Zhou [1], Jia He [1], Xiaoyan Bi [2], Ziming Yu [1], Yi Chen [1], Yan Chen [2], Guangjian Wang [1], Jianglei Ma [2], Peiying Zhu [2]

[1] Wireless Technology Lab

[2] Ottawa Advanced Wireless Technology Lab

Abstract

Wireless sensing is an important technology for various applications such as logistics detection, object identification, and environment reconstruction. Its ability to penetrate through objects and its non-ionizing nature make it an attractive alternative. In order to incorporate intelligent technologies into the upcoming 6G network, it is crucial to gather information from the surrounding environment through sensing. Integrated sensing and communication (ISAC) is considered essential for connected intelligence in 6G because it helps optimize spectrum usage, hardware resources, and computing power [1]. Additionally, by utilizing the environmental knowledge provided by sensing, communication throughput can be improved through reconstructed channel information. This paper focuses on studying environment reconstruction in a scenario where multiple transmission reception points (TRPs) are enabled by a communication network. We propose a multistatic sensing architecture based on the assumption of scatter groups to enhance sensing accuracy. Furthermore, we introduce a group trace scheme that leverages previously reconstructed environments to reconstruct communication channels. Through a validation campaign involving outdoor measurements at 10 GHz, we demonstrate that our approach can achieve sensing accuracy at the decimeter level and channel reconstruction at the degree level.

Keywords

ISAC, 6G, environment reconstruction, sensing, channel reconstruction

1 Introduction

Wireless sensing, which includes positioning, motion detection, and imaging, has traditionally been developed as a separate technology alongside mobile communication systems. However, in the context of 6G, sensing services will be expanded to encompass the detection of both active and passive objects. By utilizing sensing technology to gather information about the physical environment, a virtual cyber world can be created, enabling exploration of communication capabilities and enhancing user experience. In 6G, sensing and communications will be integrated into a reciprocal system [2], allowing network sensing to achieve intelligent connectivity. To fully leverage the communication network's role in sensing services, the concept of multi-transmission reception point (TRP) sensing has been introduced to enhance sensing performance.

There are two primary forms of integrated sensing and communication (ISAC) systems: monostatic sensing and bistatic sensing. In monostatic sensing, the sensing transmitter and receiver are co-located. This allows the ISAC signal to use the payload as a sensing pilot, without consuming any communication resources. This configuration also allows the transmitter and receiver to share hardware components, such as oscillators, eliminating non-ideal factors like phase noise. Consequently, the sensing algorithm has a lower complexity and better accuracy. Most components of the sensing signal follow one-bounce paths. This prevents ill-conditioned equations and leads to higher solution accuracy for sensing. However, the detection range based on monostatic sensing is closely related to the environmental incident angle, and the sensing field of view is significantly affected by the material and orientation of the target.

Bistatic sensing involves a receiver and transmitter being placed at different locations, requiring the use of a dedicated sensing pilot as additional overhead for the ISAC signal. However, due to differences in the transmitter and receiver systems, non-ideal factors such as synchronization and phase noise can lead to degradation in sensing performance. This requires a more complex calibration algorithm to address the issue. Additionally, bistatic sensing signals have rich components with a large number of multiple-bounce paths. In urban scenarios with rich scattering, multiple-bounce paths can introduce ill-conditioned problems, resulting in shadow space and incorrect solutions in sensing. Despite these challenges, bistatic sensing offers a unique

contribution to sensing range. As the terminal moves, the sensing field of view is also increased.

Mobile communication networks offer significant opportunities and benefits for cooperative sensing. The sensing results are shared among nodes through the mobile network, allowing multiple network nodes such as base stations (BSs) and user equipment (UE) to function as a joint sensing system. This collaboration between different nodes reduces sensing uncertainty through data fusion, resulting in improved coverage capability, sensing accuracy, and resolution. The main research challenges in achieving optimal sensing fusion results with low overhead include synchronization, joint processing, information transfer and compression, and network resource allocation.

This paper will focus on ISAC, as well as joint processing of multi-TRP sensing data for environment reconstruction based on scatter group assumption [3]. Additionally, the accurately reconstructed environment can assist in channel reconstruction, thereby facilitating communication. The paper will also discuss an ISAC prototype and validation, followed by the conclusions.

2 Environment Reconstruction

With the widespread acceptance of ISAC as one of the key technology trends for 6G, this paper takes a step further and explores the application of 6G ISAC technologies in accurate sensing-assisted communication. Our goal is to utilize the communication architecture to achieve accurate reconstruction of the environment with minimal hardware modifications [4]. By reconstructing the environment, the communication propagation channel becomes easier to reconstruct. This reconstructed environment can then be used to assist with tasks such as beam alignment and link failure mitigation, resulting in a significant reduction in communication overhead. The flowchart in Figure 1 illustrates the process of environment reconstruction and sensing-assisted communication. In this process, echo data (of the monostatic approach) is acquired through beamforming scanning. Subsequently, scattering point extraction is performed. To address the presence of multiple reflections and noise, a radius outlier removal filter [5] is employed to eliminate displaced points. The unsupervised clustering algorithm and scattering planes are combined to extract semantic information about the environment. Finally, a polygon trace based on the reconstructed environment is implemented to assist with beamforming.
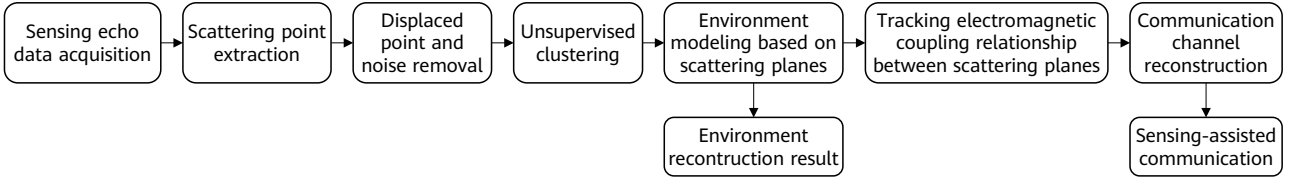
Figure 1 Environment reconstruction and sensing-assisted communication process

2.1 Clustering Algorithm

The Gaussian mixture model (GMM) is used for the electromagnetic sensing data segmentation. GMM is parameterized by two types of values, the mixture component weights and the component means and variances/covariance. Assume that there are K components, and that the k_{th} component has a mean of μ_k and covariance matrix of Σ_k. The mixture component weights ϕ_k for component C_k comply with the constraint $\sum_{k=1}^{K} \phi_k = 1$. The multi-dimensional model can be defined as:

$$p(x) = \sum_{k=1}^{K} \phi_k \mathcal{N}(\vec{x}|\vec{\mu}_k, \delta_k),$$

$$\mathcal{N}(\vec{x}|\vec{\mu}_i, \Sigma_k) = \frac{\exp\left(-\frac{1}{2}(\vec{x} - \vec{\mu})^T \Sigma_i^{-1}(\vec{x} - \vec{\mu}_i)\right)}{\sqrt{(2\pi)^K |\Sigma_i|}}.$$

(1)

The expectation maximization [6] algorithm is used to estimate the parameters of GMM. GMM is a soft decision framework. Generally speaking, GMM assigns a cluster membership score γ_{ik} to the ith point to describe the probability that the ith point belongs to the kth cluster. Therefore, we have:

$$\sum_{k=1}^{K} r_{ik} = 1$$

(2)

where K is the total number of cluster categories. The maximum number of $[\gamma_{i1}, \gamma_{i2}, \cdots, \gamma_{iK}]$ decides to which cluster the ith point belongs. Therefore, the probability based on data distribution can be referred to as belief1.

2.2 Double Belief for Scattering Plane Reconstruction

Unlike the light detection and ranging (LiDAR) point cloud, the sensing accuracy of wireless sensing data is severely affected by the beam width θ, distance d and incident angle α to scattering planes. To solve the problems mentioned above, we introduce belief2 based on the physical propagation model defined as:

$$\beta = \frac{1}{1 + e^{d\theta \sin(\alpha - \theta/2)}}.$$

(3)

Since two types of beliefs are defined, the overall belief b can be represented as:

$$b = \gamma\beta.$$

(4)

The center of the kth reconstructed plane can be defined as:

$$\boldsymbol{c}_k = \frac{\sum_i b_{ik} \boldsymbol{p}_{ik}}{\sum_i b_{ik}},$$

(5)

where $\boldsymbol{p}_{ik}$ represents the location of ith point belonging to kth cluster.

To ensure high accuracy in reconstructing the environment, only high-belief points are used while low-belief points are discarded. The normal vector δ of the kth cluster is determined using an oriented bounding box through principal component analysis (PCA) [7]. Since the point cloud obtained from monostatic sensing in this study is sparse, evaluating the accuracy of environment reconstruction using Chamfer distance or F-score based on the point cloud is not suitable. Instead, the reconstructed scattering group is considered as a whole to measure the accuracy of reconstruction. Therefore, two indicators have been defined to describe the accuracy of environment reconstruction: the normal vector error and the distance error, as illustrated in Figure 2.
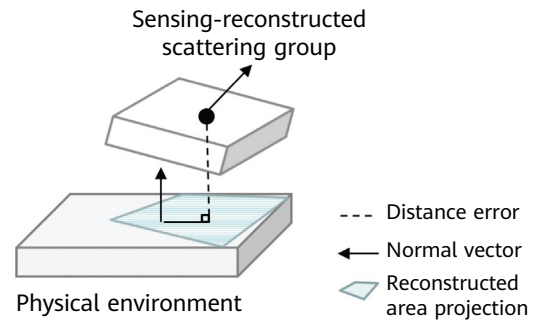
Figure 2 Scattering group for reconstruction

The double-belief algorithm is developed using the research mentioned above. It utilizes GMM to initialize semantic information extraction and environment reconstruction. This allows us to obtain the first belief1 based on the data distribution. Additionally, the initial reconstructed environment helps in easily acquiring the first belief2 based on physical propagation characteristics. To improve the accuracy of the reconstruction, we can update belief2 β using loops, as illustrated in Figure 3. The loop will end when the normal vector $|\hat{\delta}_{n+1} - \hat{\delta}_n| < \epsilon$. ϵ is a preset threshold.
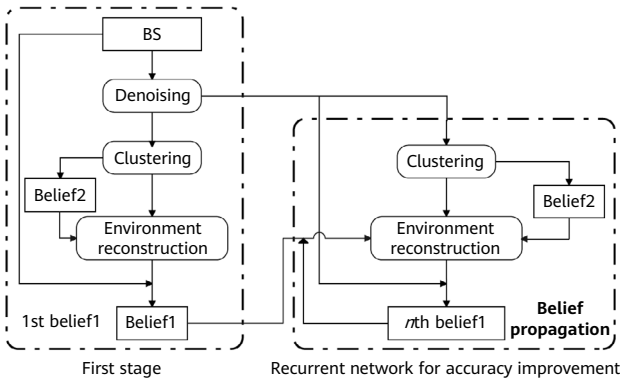
Figure 3 Double-belief algorithm for reconstruction

3 Experiment

3.1 Measurement Campaign

In order to validate the environmental reconstruction and sensing-assisted communication performance, measurement campaigns for ISAC are implemented at 10 GHz. The BS configurations are shown in Table 1.

Table 1 Configurations of BS sensing measurement

Waveform	OFDM
Frequency (GHz)	10–10.5
Bandwidth (MHz)	500
TRx antenna beam width (°)	15
Horizontal sweeping angle (°)	0:5:355
Vertical sweeping angle (°)	-30:5:30
Sweeping frequency interval (MHz)	1

Figure 4 shows a typical transceiver simulated for BS sensing measurement.
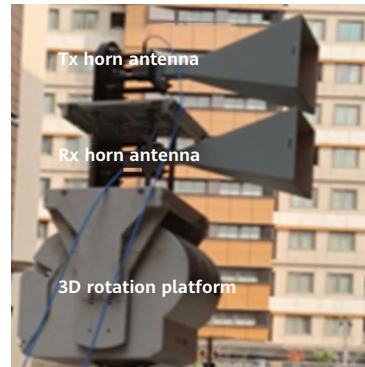
Figure 4 Transceiver simulated for BS sensing measurement

The outdoor scenario and the reconstructed environment are shown in Figure 5, where two BSs and three UEs are located.
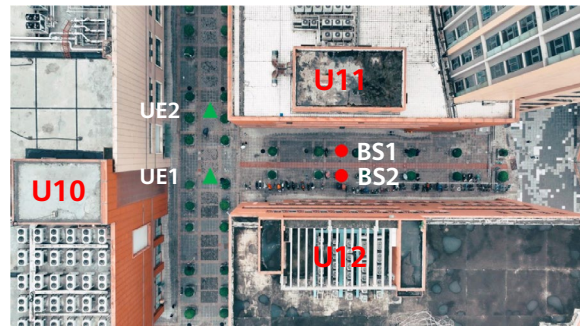
Figure 5 Sensing scenario

A corresponding communication scenario is considered as shown in Figure 6 to evaluate the sensing-assisted channel reconstruction performance. Three UEs are set in different places.
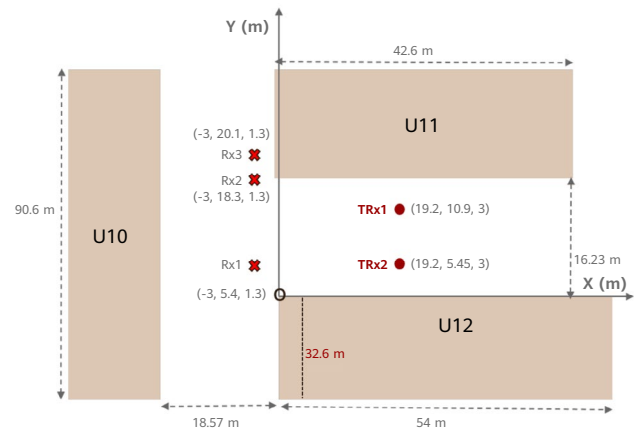
Figure 6 Communication measurement scenario

3.2 Environment and Channel Reconstruction

The environment reconstruction is shown in Figure 7.

Figure 7 Environment reconstruction result

Since the ground truth data of the foliage is hard to decide, the location and normal vector of two walls from U11 and U12 are used to evaluate the accuracy of the reconstructed environment. The environment reconstruction accuracy is shown in Table 2. It is evident that the double-belief algorithm shows better performance than GMM.

Table 2 Environment reconstruction accuracy

Method	Distance Error	Normal Vector Error
GMM	20.34 cm	4.19°
Ours (Single TRP)	12.12 cm	3.23°
Ours (Multi-TRP)	8.30 cm	2.10°

The channel reconstruction results of the three UEs are displayed in Figure 8. In the figure, the 'x' and triangle symbols represent the measurement results and the results obtained through the polygon trace method based on the reconstructed environment, respectively. It can be observed that the channel parameters, such as the angle of arrival (AoA) and delay, exhibit a strong agreement between the line-of-sight (LOS) and one-bounce non-line-of-sight (NLOS) paths.

(a)

(b)

(c)
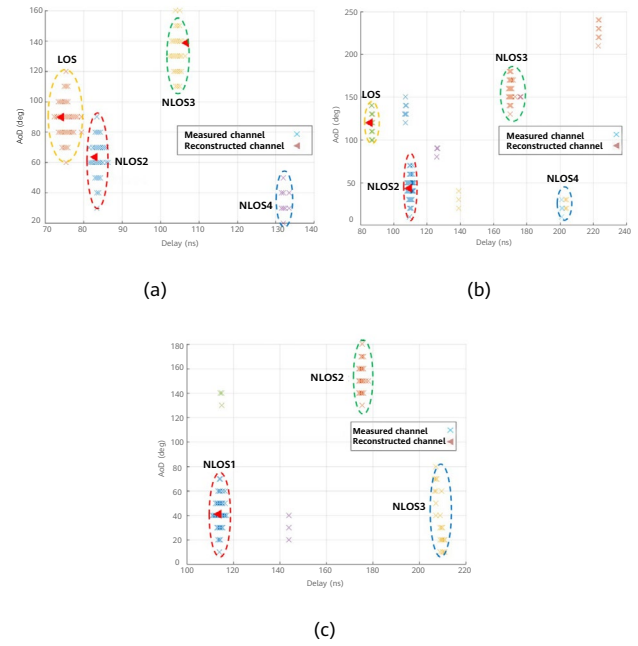
Figure 8 Channel reconstruction result (a) Rx1, (b) Rx2, (c) Rx3

The channel reconstruction errors of delay and angles are shown in Table 3. There is no LOS path between the Tx and the Rx3. The average channel reconstruction errors of delay and angles are 0.52 ns and 3.18°, respectively.

Table 3 Channel reconstruction result

Channel Reconstruction	LOS		NLOS	
	Angle Error (°)	Delay Error (ns)	Angle Error (°)	Delay Error (ns)
Rx1	0.13	1.19	4.16	1.63
Rx2	0.01	1.21	2.76	0.2
Rx3	×	×	1.08	0.14
Average	0.07	1.2	3.18	0.52

4 Conclusion

In this paper, we focused on the usefulness of sensing and environment reconstruction. We provided illustrations and verification of the implementation of environment reconstruction using a multi-TRP ISAC network. Additionally, we proposed a channel reconstruction method based on the results of sensing. To validate the feasibility of sensing-assisted communication, particularly beam alignment, we developed a 10 GHz prototype and compared the measured communication channel with the sensing-assisted reconstructed channel. Our results show that we achieved decimeter-level accuracy in outdoor scenarios for environment reconstruction. Furthermore, thanks to sensing-assisted channel reconstruction, we achieved subnanosecond-level delay and approximately 3° of beam accuracy.

References

[1] Wen Tong and Peiying Zhu, *6G, the Next Horizon: From Connected People and Things to Connected Intelligence*, Cambridge University Press, Ottawa, 2021.

[2] Danny Kai Pin Tan, Jia He, Yanchun Li, Alireza Bayesteh, Yan Chen, Peiying Zhu, and Wen Tong, "Integrated sensing and communication in 6G: Motivations, use cases, requirements, challenges and future directions," in *2021 1st IEEE International Online Symposium on Joint Communications & Sensing (JC&S)*. IEEE, 2021, pp. 1–6.

[3] Isamu Matsunami, Ryohei Nakamura, and Akihiro Kajiwara, "RCS measurements for vehicles and pedestrian at 26 and 79GHz," in *2012 6th International Conference on Signal Processing and Communication Systems*. IEEE, 2012, pp. 1–4.

[4] Oupeng Li, Jia He, Kun Zeng, Ziming Yu, Xian feng Du, Zhi Zhou, Yuan Liang, Guangjian Wang, Yan Chen, and Peiying Zhu, "Integrated sensing and communication in 6G: A prototype of high resolution multichannel THz sensing on portable device," *EURASIP Journal on Wireless Communications and Networking*, vol. 2022, no. 1, pp. 1–21, 2022.

[5] Yao Duan, Chuanchuan Yang, and Hongbin Li, "Low-complexity adaptive radius outlier removal filter based on PCA for lidar point cloud denoising," *Applied Optics*, vol. 60, no. 20, pp. E1–E7, 2021.

[6] Guorong Xuan, Wei Zhang, and Peiqi Chai, "Emalgorithms of Gaussian mixture model and hidden Markov model," *in Proceedings 2001 international conference on image processing (Cat. No. 01CH37205)*. IEEE, 2001, vol. 1, pp. 145–148.

[7] Darko Dimitrov, Christian Knauer, Klaus Kriegel, and Günter Rote, "On the bounding boxes obtained by principal component analysis," in *22nd European Workshop on Computational Geometry*, 2006, pp. 193–196.

Environment Reconstruction Based on Hybrid Aperture Sensing: Technologies and Applications

Jiajin Luo [1], Xiaohui Peng [1], Yang Yu [1], Ping Zhang [1], Baojian Zhou [1], Xiaoyan Bi [2], Yan Chen [2], Jianglei Ma [2], Peiying Zhu [2]

[1] Wireless Technology Lab

[2] Ottawa Advanced Wireless Technology Lab

Abstract

This paper proposes a hybrid aperture sensing (HAS) technology based on wireless networks, which utilizes the mobility of terminal nodes in the network to form virtual apertures and combines them with the real physical apertures of base stations to form hybrid aperture. By reusing the (uplink or downlink) air interface resources between base stations (real apertures) and terminals (virtual apertures), HAS significantly enhances the sensing resolution and enables the fusion of 3D point clouds from multiple network nodes, resulting in a digital twin environment model that corresponds to the physical world. The paper also explores the use of the reconstructed digital twin environment model for wireless communication channel prediction and clutter suppression in wireless sensing.

Keywords

hybrid aperture sensing (HAS), environment reconstruction, digital twin, sensing-assisted communication, sensing-assisted localization, 6G

Integrated sensing and communication (ISAC) as a highly promising feature for next-generation cellular networks, has been approved as one of the six major usage scenarios for IMT-2030 (6G) by ITU-R. In the scenario of ISAC, the cellular network infrastructures and devices not only serve as communication nodes for connectivity between people and things, but also as sensing nodes that help achieve environment imaging and reconstruction, high-precision target localization and tracking, gesture and motion recognition, and other sensing functions. Among them, environment reconstruction serves as an enabling technology for digital twins [1], and will play an important role in future scenarios such as digital reconstruction of a physical world, target localization, and autonomous driving. From the perspective of measurement, 3D environment reconstruction mainly has two technical routes: optical and microwave radar. Lidar-based 3D environment reconstruction solutions are known for their high precision, but are expensive, involve time-consuming deployment, and require a long data update cycle. Recently, the introduction of deep learning technology has made rapid progress in reconstructing 3D images based on 2D optical images using neural networks [2], which has lower equipment requirements but still needs to calibrate camera positions and separately training for different categories of objects. Its generalization capability and outdoor performance still need to be improved. Researchers in the radar imaging field have also conducted a lot of research on various imaging systems, such as vehicle-based and unmanned aerial vehicle-based 3D radar imaging systems [3–5], which are less vulnerable to ambient lighting and weather conditions compared to optical methods and have a larger coverage area. Traditional solutions for reconstructing environments are typically deployed and calibrated independently, which places high requirements on processing capabilities of devices. This results in high costs in terms of deployment, data update, and maintenance, failing to build ubiquitous digital twins in the future. To address this limitation, this paper proposes the hybrid aperture sensing (HAS) technology for next-generation wireless network, which fully utilizes the signal processing capabilities, data transmission capabilities, and positioning and synchronization capabilities of the network side, combines the terminal mobility, and does not require independent deployment and calibration, thus offering an on-demand, high-quality, and wide-coverage digital twin environment model. To further enhance the capabilities of wireless communication and localization, the impact of the environment on wireless channels [6] has also received more attention, and this paper further investigates channel prediction and clutter

suppression by using the reconstructed environment model and evaluates its performance preliminarily. The following sections introduce the three parts of the article: sensing-based environment reconstruction, environment sensing-assisted communication, and environment sensing-assisted localization.

1 Sensing-based Environment Reconstruction

1.1 Hybrid Aperture Sensing

The sensing mode of a typical cellular network can be divided into two categories: base stations (BSs) only sensing and BS-terminal joint sensing. This paper proposes a HAS-based scheme, as shown in Figure 1, which adopts a joint bistatic sensing mode between BSs and mobile terminal. The BS side utilizes a real apertures antenna array to ensure the vertical resolution, while the mobile terminal, such as vehicles and UAVs, can form a virtual aperture to observe the environment, thereby achieving a larger field of view and higher horizontal resolution than BSs only sensing. The HAS-based solution also features reciprocity between uplink and downlink channels, allowing both uplink and downlink signals to be used for sensing. Considering that the BSs typically have stronger signal processing and transmission capabilities, this paper takes the uplink sensing, i.e., the BS receives and the mobile terminal transmits, as an example of a specific implementation of the HAS-based solution, and analyzes and demonstrates the performance of the cellular network in 3D environment reconstruction, as well as the use of the reconstructed environment model to predict communication channels and assist in target localization.
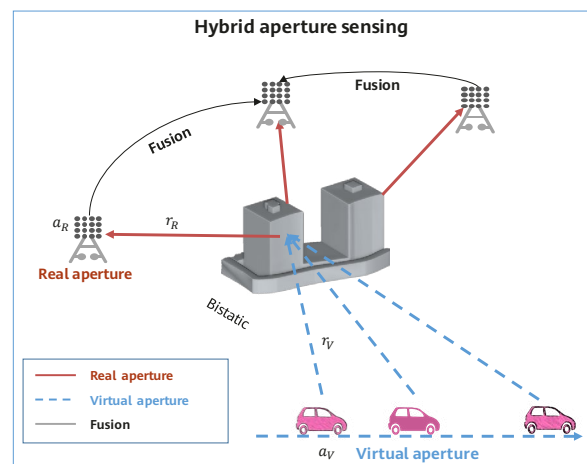
Figure 1 HAS-based solution

First, let's briefly review the basic principle of obtaining a scattered point cloud C = { $I(x, y, z)$ } by using back projection (BP) algorithm. More algorithm details are available in [7, 8]. $I(x, y, z)$ represents the scattering intensity at a spatial target point $P(x, y, z)$, and can be calculated from (1).

$$I(x, y, z) = \iint s(r_H = r_V + r_R, a_V, a_R) e^{i\frac{2\pi}{\lambda} r_H} da_V da_R \quad (1)$$

where $s(r, a_V, a_R)$ represents the echo pulse compression result at a given virtual aperture position $a_V(x_V, y_V, z_V)$ and a given real aperture position $a_R(x_R, y_R, z_R)$. λ is the wavelength of the carrier. r_V represents the slant range between imaging target point P and the mobile node at position a_V. r_R represents the slant range between point P and real-aperture array element a_R of BSs. r_V and r_R can be expressed as

$$\begin{cases} r_V = \sqrt{(x_V - x)^2 + (y_V - y)^2 + (z_V - z)^2} \\ r_R = \sqrt{(x_R - x)^2 + (y_R - y)^2 + (z_R - z)^2} \end{cases}$$

Then, the obtained point cloud $I(x, y, z)$ with a specific scattering intensity is filtered to obtain a point cloud representing the geometric shape of targets. Finally, a 3D environment model can be reconstructed by triangulating these points using a surface reconstruction algorithm, such as Poisson surface reconstruction [9], Ball Pivoting [10], Occupancy Networks [11], or DeepSDF [12].

1.2 Simulation Evaluation Result

This section describes a simulation scenario of environment sensing by using BS-terminal hybrid apertures where BSs function as receivers, and a vehicle and a UAV function as transmitters. Figure 2 shows that four BSs (BS$_1$ to BS$_4$) are deployed around the factory building (in red). A vehicle drives around the factory, while a UAV flies above the factory parallel to the ground.

In this simulation, the normal direction of the antenna arrays of BSs is parallel to the ground, and the array size is 1 m x 0.5 m. The UAV forms a horizontal virtual aperture with an equivalent length of 3 m in the air. Similarly, the moving vehicle forms five horizontal virtual apertures in different directions, each with an equivalent length of 3 m. These apertures are labeled V$_1$ to V$_5$ in Figure 2. The channel information between the transmitter and receiver can be obtained by using ray tracing method, and more channel data can be found in our open dataset, Sensiverse [13].
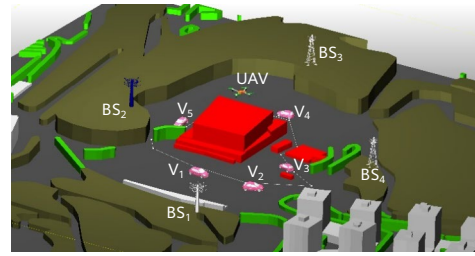
Figure 2 Simulation of HAS-based environment reconstruction

The BSs, vehicles, and UAV operate at a center frequency of 10 GHz with a bandwidth of 400 MHz. Transmitted waveforms such as OFDM, OTFS, and FMCW can be used. Echo data is processed by using imaging algorithms to obtain imaging results of the target building. Figure 3 shows the imaging results obtained through BS-mobile terminal pair sensing with the transmitted waveform set to OFDM. Due to visual angles and blocking, the coverage of each single pair of BS-mobile terminal sensing is limited, resulting in only partial imaging of the building. Therefore, the next step is to perform weighted fusion processing on the imaging results of each pair of BS-mobile terminal according to their coverage areas and corresponding theoretical resolutions.

(a) BS$_1$-UAV

(b) BS$_1$-V$_1$

(c) BS$_1$-V$_2$

(d) BS$_2$-V$_5$

(e) BS$_3$-V$_4$

(f) BS$_4$-V$_3$
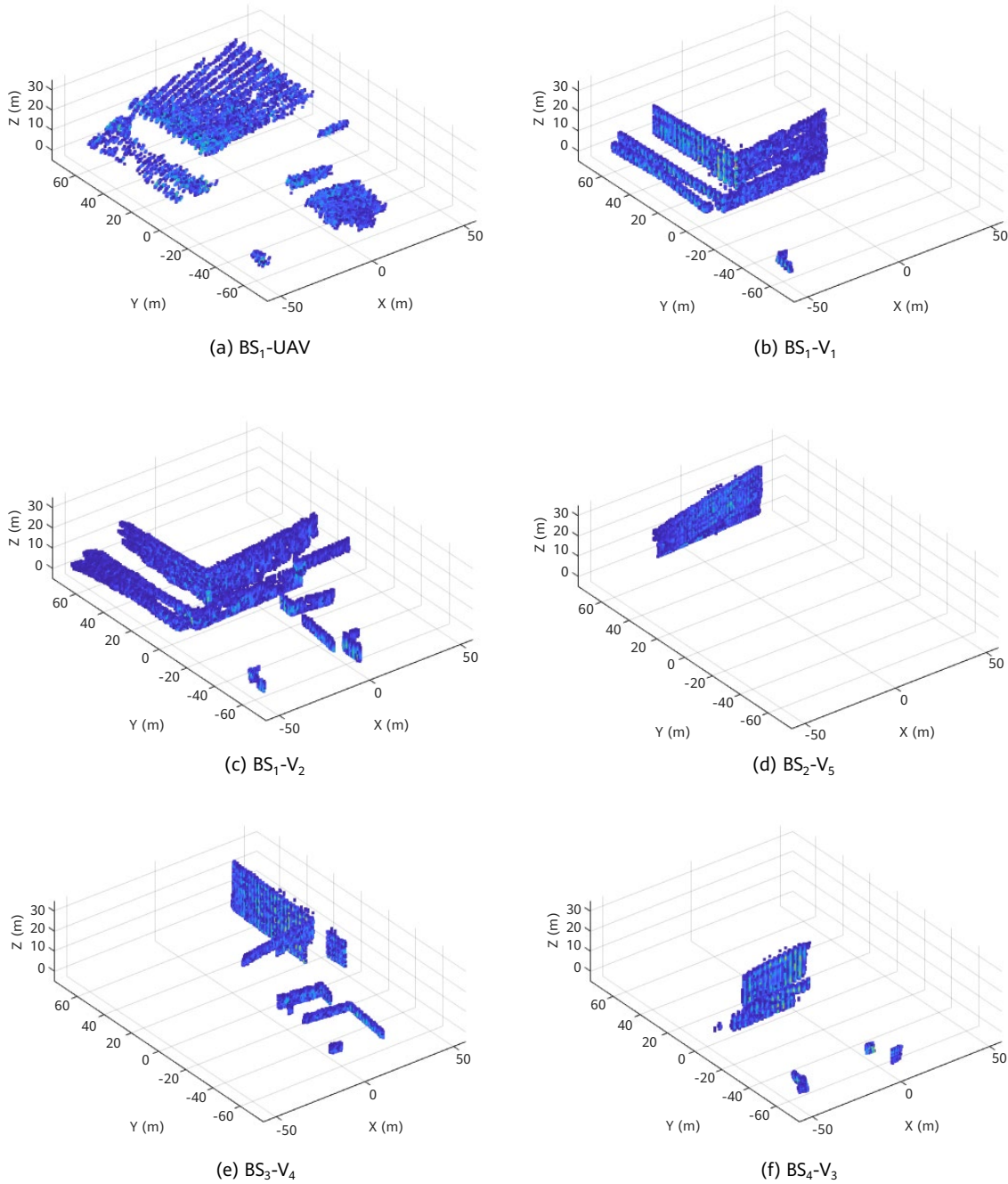
Figure 3 Imaging results based on BS-mobile terminal sensing

Figure 4 shows the fusion result of six groups of imaging point clouds. It can be seen that the fusion imaging result reflects the contour of the target building quite completely.
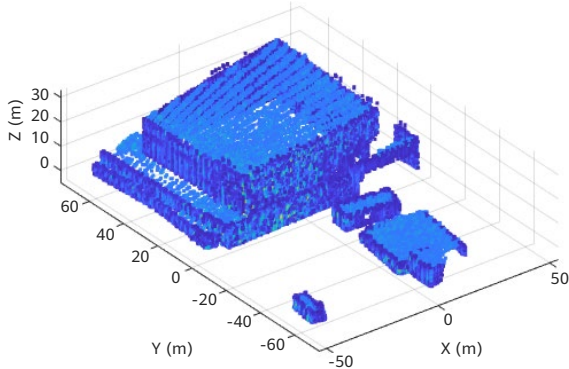
Figure 4 Fusion result of HAS-based imaging point clouds

It should be noted that while HAS can achieve higher-resolution sensing results than BS only sensing, the time and frequency synchronization errors between transmitting and receiving nodes will affect the sensing performance. This is a problem that needs to be further explored and studied in the future.

1.3 Point Cloud Modeling and Performance Evaluation

Based on the environment imaging results of the previous section, we can reconstruct a precise geometric model (shown in Figure 5) that is easy to express, process, and transmit and contains physical characteristics. This model is crucial for various applications, including three-dimensional modeling in indoor and outdoor scenarios, wireless channel estimation, mixed reality, and digital twin.

The problem of reconstructing a geometric model on a point cloud can be described as follows: given a set of unordered three-dimensional points $P = \{p_1,\ p_2,\ ...,\ p_n\}$, the goal is to reconstruct a mesh with clear geometric topology structure, $\mathrm{Mesh} = \{V, E\}$, using a surface reconstruction algorithm, where V and E represent the coordinates of the mesh vertices and the connectivity relationships between them, respectively.
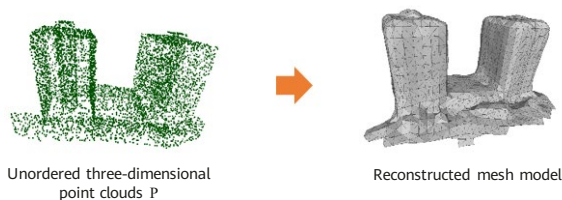
Figure 5 Reconstructed mesh model based on point clouds

Quality evaluation of the reconstructed mesh model typically involves comparing the difference between the sampling points in the reconstructed mesh model and the true value of scattering points. Common evaluation metrics include Chamfer Distance [14] and F-score [15], which are defined as follows:

$$CD\left(S_1, S_2\right) = \frac{1}{|S_1|} \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2^2 + \frac{1}{|S_2|} \sum_{x \in S_2} \min_{y \in S_1} \|x - y\|_2^2 \quad (2)$$

S_1 and S_2 respectively represent the sampling points of the reconstructed model and the true value of scattering points. To use the length dimension, this paper uses the square root of Chamfer Distance, i.e., $\sqrt{\mathrm{Chamfer\ Distance}}$ as one of the metrics for measuring the reconstruction quality. And

$$F - score = \frac{2P\left(d\right) \cdot R\left(d\right)}{P\left(d\right) + R\left(d\right)} \quad (3)$$

where $P(d)$ and $R(d)$ respectively represent the matching degree between the reconstructed model and the true value of scattering points with respect to a given distance threshold d, and are defined as follows:

$$P\left(d\right) = \frac{100}{|S_1|} \sum_{x \in S_1} \left[\min_{y \in S_2} \|x - y\|_2 < d\right]$$

$$R\left(d\right) = \frac{100}{|S_2|} \sum_{x \in S_2} \left[\min_{y \in S_1} \|x - y\|_2 < d\right]$$

By leveraging the environment imaging results from the previous section as input and using the Poisson surface reconstruction algorithm, a geometric model is reconstructed based on the input point clouds, as shown in Figure 6. The reconstructed model is evaluated quantitatively using the $\sqrt{\mathrm{Chamfer\ Distance}}$ and F-score. As presented in Table 1, the reconstructed mesh model exhibits $\sqrt{\mathrm{Chamfer\ Distance}}$ of 0.96 m, and an F-score with a distance threshold of 1.5 m is 96. In contrast, the original input point clouds exhibit $\sqrt{\mathrm{Chamfer\ Distance}}$ of 1.18 m and an F-score of 94 using the same distance threshold. By comparing the quantitative error evaluation result in Table 1, it is evident that the reconstructed geometric model is more representative of the environment than the original point cloud, and the $\sqrt{\mathrm{Chamfer\ Distance}}$ and F-score values are better than those of imaging point clouds.

Figure 6 3D environment reconstruction result

Table 1 Evaluating the environment reconstruction error

Performance	Imaging Point Cloud	Reconstructed Model
$\sqrt{\text{Chamfer Distance}}$	1.18 m	0.96 m
F-score (d = 1.5 m)	94	96

2 Environment Sensing-assisted Communication: Channel Prediction

A 3D reconstruction of the surrounding environment is obtained through 3D imaging and meshing. Let's assume that the objects are made of concrete. For a UE at a given location in the environment, multipath parameters $\{\tau_n, A_n, \text{AOD}_n, \text{ZOD}_n, \text{AOA}_n, \text{ZOA}_n\}$ between a BS and the UE can be obtained by using wireless channel prediction methods such as electromagnetic calculation and ray tracing. These multipath parameters represent the delay, amplitude, azimuth angle of departure (AOD), zenith angle of departure (ZOD), azimuth angle of arrival (AOA), and zenith angle of arrival (ZOA) of the nth multipath, respectively. Based on these parameters, the frequency domain response of the channel can be further calculated.

In channel prediction based on the reconstructed environment, errors are inevitable due to inaccuracies in environment reconstruction and UE locations. Specifically, errors in environment reconstruction are primarily geometric errors of a 3D environment and errors in the electromagnetic parameters of materials. Errors in environment reconstruction can occur due to (1) a failure in perfectly reconstructing the environment and (2) a failure in real-time imaging for dynamic targets (such as vehicles or pedestrians) in the environment. UE locations in the environment are subject to dynamic changes, which presents challenges for accurate UE localization. In practice, only estimated UE locations can be used for prediction. This paper primarily focuses on examining the impact of errors in environment reconstruction.

To quantitatively evaluate the impact of errors in environment reconstruction on channels, let's consider a frequency of 10 GHz as an example. Figure 7 depicts one BS located at a specific location and 500 UEs randomly distributed. Using the ray tracing method, we obtain multipath parameters for each BS–UE pair based on an ideal model and a reconstructed model. For statistical analysis, we only consider UEs with attenuation of the strongest path

less than 110 dB (indicated by red points). Among these UEs, we reserve the multipaths for which energy attenuation is less than 20 dB when compared with the strongest path for each UE. This enables us to use the Chamfer Distance and compare the errors in multipath parameters for the two models. Figure 8 shows the multipath parameters obtained separately based on the two models, and the Chamfer Distance is used to compare the errors in these parameters.
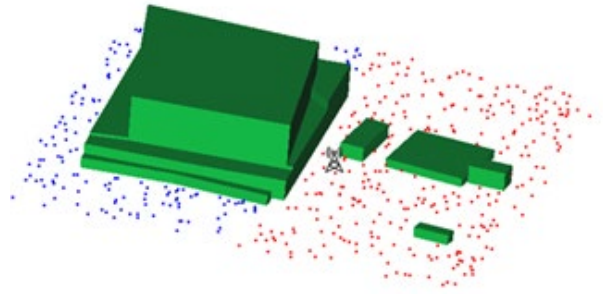
Figure 7 Scenario configuration for analysis of errors in channel multipath prediction
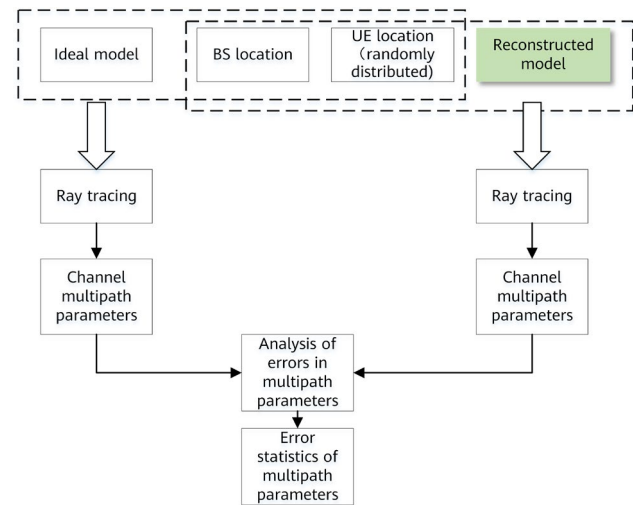
Figure 8 Process for analyzing the channel prediction errors

Figure 9 shows the statistical results of the error analysis for the multi-path parameter prediction based on the reconstructed model. From the statistical results, it can be seen that the time delay error is within 3.5 ns for 90% of the cases, the AOA error is within 3 degrees for 90% of the cases, the AOD error is within 3 degrees for 90% of the cases, the ZOA error is within 1 degree for 90% of the cases, and the ZOD error is within 1 degree for 90% of the cases. The channel prediction result can be used to assist beam sweeping in communications, reducing the beam sweeping overhead.
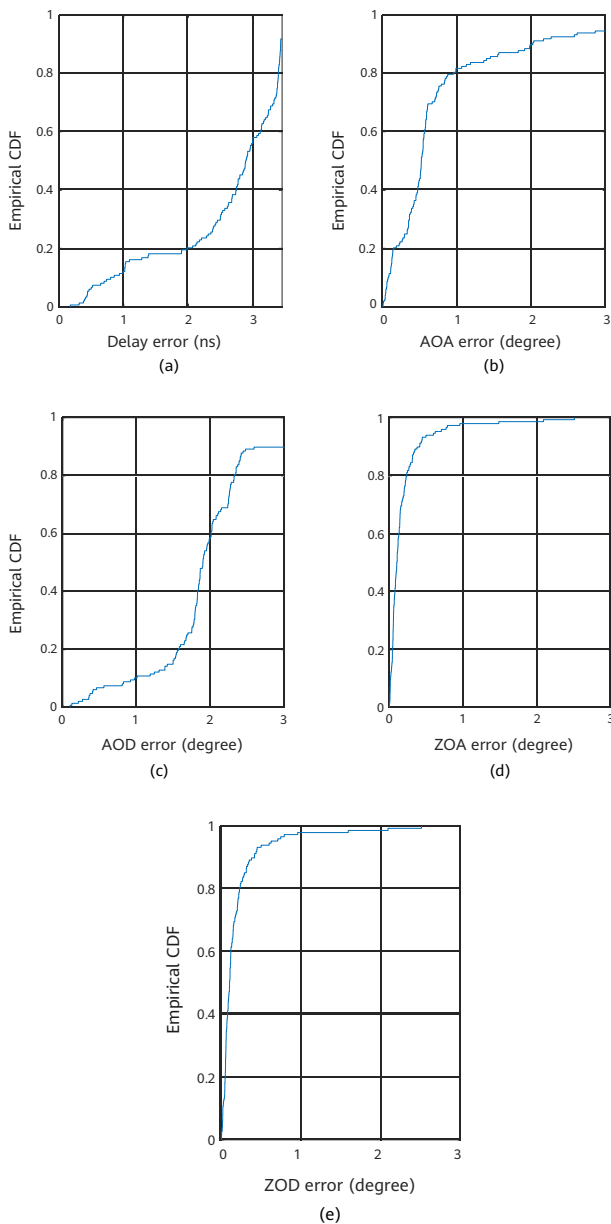
Figure 9 (a) Delay errors; (b) AOA errors;
(c) AOD errors; (d) ZOA errors; (e) ZOD errors

3 Environment Sensing-assisted Localization: Clutter Suppression

In existing cellular networks, the primary focus of positioning and tracking of targets is on active devices, such as smartphones, vehicles, and UAVs. In the future, with the introduction of sensing functions, 6G networks will be able to provide positioning and tracking not only for active devices but also for passive targets. This will open up a range of new application scenarios, including but not limited to flight tracking of UAVs, UAV intrusion detection, localization and tracking of vehicles, detection of pedestrians and animals on highways, and communications assisted by sensing results of moving targets. This section takes passive targets as an example for analysis, and the mobile targets mentioned below refer to passive targets unless specified otherwise.

Targets to be sensed and observation nodes are usually located on the ground or in a low-altitude environment, which makes ground clutter (echoes from the ground and surrounding buildings) a significant factor that affects the system sensing performance. This is especially true for detecting, localizing, and tracking small targets that move slowly in a low-altitude environment. The main problems faced include:

1. Both the reflected signals of the targets to be detected and the clutter reflected from surrounding buildings and the ground exist in the environment. If coherent accumulation is performed without any additional processing, strong side-lobe clutter may obscure targets, significantly reducing the target detection capability of BSs. Therefore, measures must be taken to suppress clutter and minimize its impact on target detection.

2. Targets that move at low speeds and have weak reflected energy. Moving targets in the environment mainly include vehicles, pedestrians, and UAVs. Among these, pedestrians move at a relatively low speed and have a small radar cross section (RCS), making them particularly susceptible to static side-lobe clutter. Consequently, detecting such targets requires a higher level of interference suppression capability.

Next, we will illustrate the performance of mobile target sensing based on a monostatic BS using a simulation example.

In the simulation, the carrier frequency is set to 10 GHz, and the bandwidth is set to 100 MHz. The BS use angle and range measurements to locate the target, with a single transmitting element and 1024 receiving elements configured in a two-dimensional uniform planar array. The receiving elements are spaced at a distance of $\lambda/2$, where λ is the wavelength of the signal, and the configuration can be equivalent to two mutually perpendicular 32-element receive uniform linear arrays using MIMO technology, without affecting the analysis results. The 32 x 32 element array is placed on the YZ plane, with its normal direction along (1,0,0) to cover the moving target. Based on monostatic BS sensing, the target's angle, range, and speed information can be obtained.

The target moves along a straight line at a constant speed, with a trajectory length of 30 m. In this example, we use a cylindrical target with a radius and height of 3 m as an extended target. The signal processing flowchart is shown in Figure 10, and it should be noted that an environment-assisted module is added compared to traditional methods (such as ECA, extensive cancellation algorithm). The reflected signal in the frequency domain can be represented as $Y(f) = S(f)*H(f)$, where $S(f)$ is the Fourier transform of the transmitted waveform $s(t)$, and $H(f)$ is the frequency-domain response of the channel. After obtaining the echo data, the target angle, range, and Doppler information can be obtained through signal processing, and finally, the target localization information can be obtained.
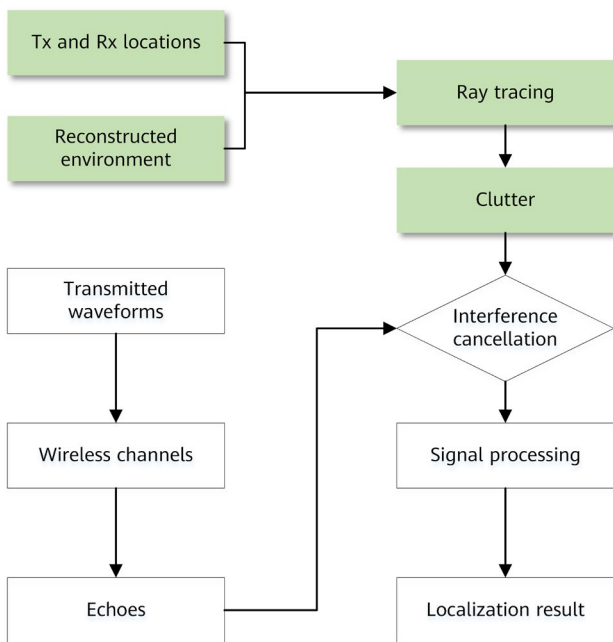
Figure 10 Flow of signal processing for detecting moving targets

As previously mentioned, wireless sensing can be used to obtain an environment reconstruction result. This result and the transceiver location can be used to obtain multipath information by using channel prediction algorithms such as electromagnetic calculation and ray tracing. By relying on multipath estimation for the static environment, background clutter can be suppressed based on the environment information.

As shown in Figure 11a and Figure 11b, we use $\sqrt{\text{Chamfer Distance}}$ to quantitatively evaluate the accuracy of target localization. Assuming that the true scattering point coordinates of the target are S1 (as shown in red line in Figure 11), and the detected target localization coordinates are S2 (as shown in blue point in Figure 11), the error of the target localization

is 1.73 m when not using the reconstructed environment prior information, and 1.07 m when using the reconstructed environment prior information. Figure 11c shows the corresponding detection results. It can be seen that using the prior information of the reconstructed environment can detect more points. Without using the prior information of the reconstructed environment, some target points will be buried by environmental echoes.
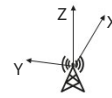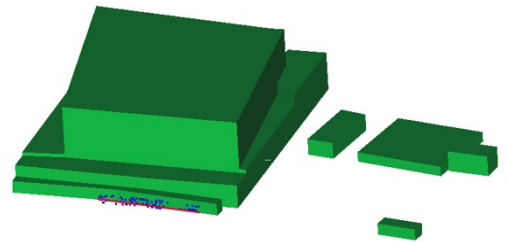
(a)
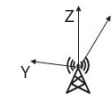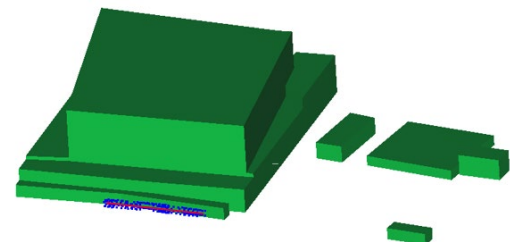
(b)
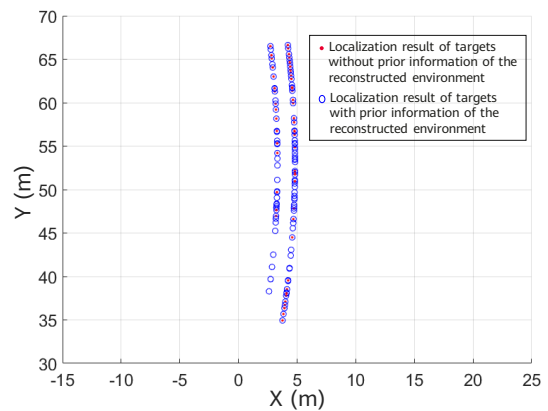
(c)

Figure 11 (a) Detection result without using prior information of the reconstructed environment; (b) Detection result using prior information of the reconstructed environment; (c) Projection of the localization results on a two-dimensional plane

4 Conclusions

This paper outlines an end-to-end solution covering HAS, digital twin environment reconstruction, and digital twin-assisted communication and localization. We evaluated the performance in environment reconstruction, parameter estimation, and target localization by simulation. The simulation results demonstrate that HAS enables submeter-level high-precision environment reconstruction in large urban scenarios on 10 GHz (centimeter waves). Based on the reconstructed environment model, ray tracing is used to predict the channels of UEs at different locations. In addition, we provided the cumulative distribution functions (CDFs) for errors in terms of delay and angles. Furthermore, the reconstructed environment model is used to enhance the detection of extended moving targets. By using environment-assisted clutter suppression, we improved the localization accuracy by approximately 40% compared to the traditional ECA approach. It should be noted that non-ideal factors, such as bistatic synchronization errors and changes in surface materials of buildings, were not considered in the simulation presented in this paper. To ensure high precision in the digital twin, several technical aspects, such as mitigating the impact of non-ideal factors and extracting the electromagnetic scattering coefficients of object surfaces, need to be further explored.

References

[1] A. Bayesteh, J. He, Y. Chen, P. Zhu, J. Ma, A. W. Shaban, Z. Yu , Y. Zhang, Z. Zhou, and G. Wang, "Integrated Sensing and Communication (ISAC) — From Concept to Practice," 11 2022. [OL]. Available: https://www.huawei.com/en/huaweitech/future-technologies/integrated-sensing-communication-concept-practice

[2] J. Gao, T. Shen, Z. Wang, W. Chen, K. Yin, D. Li, O. Litany, Z. Gojcic, and S. Fidler, "Get3D: A generative model of high quality 3D textured shapes learned from images," *Advances In Neural Information Processing Systems 35,* pp. 31841–31854, 2022.

[3] Q. Xiaolan, J. Zekun, Y. Zhenli, C. Yao, L. Bei, L. Yitong, W. Wei, D. Yongwei, Z. Liangjiang, and D. Chibiao, "Key technology and preliminary progress of microwave vision 3D SAR experimental system," *Journal of Radars,* vol. 11, p. 1–19, 2022.

[4] K. Qian, Z. He and X. Zhang, "3D point cloud generation with millimeter-wave radar," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, p. 1–23, 2020.

[5] 3GPP, "TR 22.837 Feasibility study on integrated sensing and communication," [OL]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4044

[6] Y. Yang, F. Gao, X. Tao, G. Liu, and C. Pan, "Environment Semantics Aided Wireless Communications: A Case Study of mmWave Beam Prediction and Blockage Prediction," *IEEE Journal on Selected Areas in Communications*, 2023.

[7] A. F. Yegulalp, "Fast backprojection algorithm for synthetic aperture radar," *Proceedings of the 1999 IEEE Radar Conference. Radar into the Next Millennium (Cat. No. 99CH36249)*, 1999.

[8] M. I. Duersch, "Backprojection for synthetic aperture radar," Brigham Young University, 2013.

[9] M. Kazhdan, M. Bolitho and H. Hoppe, "Poisson surface reconstruction," *Proceedings of the fourth Eurographics symposium on Geometry processing*, 2006.

[10] F. Bernardini, J. Mittleman, H. Rushmeier, C. Silva, and G. Taubin, "The ball-pivoting algorithm for surface reconstruction," *IEEE transactions on visualization and computer graphics*, graphics, vol. 5, p. 349–359, 1999.

[11] L. Mescheder, M. Oechsle, M. Niemeyer, S. Nowozin, and A. Geiger, "Occupancy networks: Learning 3D reconstruction in function space," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.

[12] J. J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove, "DeepSDF: Learning continuous signed distance functions for shape representation," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.

[13] J. Luo, B. Zhou, Y. Yu, P. Zhang, X. Peng, J. Ma, P. Zhu, J. Lu, and W. Tong, "Sensiverse: A dataset for ISAC study," *arXiv:2308.13789 [eess.SP]*, 26 August 2023.

[14] R. Sulzer, L. Landrieu, R. Marlet, and B. Vallet, "A Survey and Benchmark of Automatic Surface Reconstruction from Point Clouds," *arXiv preprint arXiv:2301.13656*, 2023.

[15] A. Knapitsch, J. Park, Q.-Y. Zhou, and V. Koltun, "Tanks and temples: Benchmarking large-scale scene reconstruction," *ACM Transactions on Graphics (ToG)*, vol. 36, p. 1–13, 2017.

Sensing-aided Communication in Integrated Sensing and Communication

Ting Zeng [1, 2], Jiebao Zhang [1, 3], Qiusha Gong [1]

[1] CICT Mobile Communication Technology Co., Ltd.

[2] Beihang University

[3] China Academy of Telecommunications Technology

Abstract

Sensing-aided communication (SAC) is a typical application in the deeper convergence phase of integrated sensing and communication (ISAC). With the sensing function, an ISAC system can obtain multi-dimensional, multi-layer user information and environmental data; when used as prior information for decision-making and management, this data can enhance communication. This paper reviews the past literature and then discusses the benefits of SAC from three perspectives: performance improvement, overhead reduction, and others. The performance improvement part touches upon sensing-aided communication channel estimation and sensing-aided radio resource management and scheduling. The overhead reduction part analyzes sensing-aided beam management overhead reduction, sensing-aided reference signal (RS) overhead reduction, and sensing-aided channel state information (CSI) overhead reduction. The "others" part centers on reconfigurable intelligent surface (RIS)-assisted sensing and sensing-aided artificial intelligence (AI) modeling and optimization. This paper concludes by discussing the prospects of SAC standardization and the challenges this technology faces, providing a reference for future research.

Keywords

1 Introduction

Integrated sensing and communication (ISAC) is an important scenario in 6G communications. In the 6G programmatic document *Framework and Overall Objectives of the Future Development of IMT for 2030 and Beyond*, released by ITU-R WP 5D at its 44th meeting held in June 2023, ISAC is listed among one of the six 6G application scenarios [1].

ISAC refers to sharing software and hardware resources in a system to provide high-quality communication and high-precision sensing, thereby reducing costs and improving system performance. Here, communication refers to conventional data transmission functions and sensing refers to range measurement, speed measurement, angle measurement, imaging, detection, and so on [2].

Through the convergence of time, frequency, and space domains [3], communication and sensing gradually evolve from coexistence and collaboration to integration, improving the overall system performance and fostering potential applications for 6G industrial upgrading. Sensing-aided communication (SAC) and communication-aided sensing (CAS) are both popular areas of research in ISAC, looking to further improve system efficiency. SAC, aiming at enhancing communication through the multi-dimensional and multi-layer information obtained by sensing, does not merely provision sensing and communication functions in the same system.

Rather, it empowers and upgrades communication for the future and is thus seen as a substantial development in ISAC.

In recent years, the academia and industry worldwide have focused on SAC research. However, much of the research is still in preliminary stages due to lack of progress in ISAC research.

The IMT2030 working group classifies SAC use cases in future communication networks into four groups: transmit end configuration, receive end algorithm selection, intelligent scheduling, and ad hoc networking [2]. The 6G Flagship in Europe suggests gathering information about obstacles and reflectors to aid beam tracking in order to save communications beam search space and reduce communication delays [4]. The research of [5] proposes to perform beam management based on radio sensing-assisted mmWave multi-input multi-output (MIMO), implement reconfigurable intelligent surface (RIS) based interference adjustment and obstacle avoidance via ambient electromagnetic sensing, and enable secure radio communications at the physical layer with assistance from an electromagnetic environment map. The research of [6] puts forward the idea of using sensing to improve communication performance, providing functions such as dynamic allocation of auxiliary radio resources, channel estimation/detection algorithm selection, dynamic adjustment of reference signal's time domain density, node selection, and more. However, most of this literature only proposes some research directions without clearly categorizing the SAC use cases or providing concrete and

Table 1 Overview of SAC use cases

Use Case	Description	Advantages	Maturity
Transmit end configuration	Adjust parameter set selection/configuration of the reference/data signal based on the sensed information [2, 6].	Reduce resource consumption, reduce latency, and improve spectrum utilization.	√√
	Perform beam training and tracking based on the sensed terminal location [2, 5, 6].	Reduce overheads and latency.	√√√√
	Perform obstacle avoidance, beam adjustment, and interference adjustment based on the sensed locations of obstacles, reflectors, and RISs [4, 5].	Enhance system performance.	√√
	Perform security communications based on environment maps built from the sensed information [5].	Improve communication security.	√
Receive end algorithm selection	Perform algorithm selection, parameter configuration, algorithm optimization, and other actions for receive end channel estimation, channel equalization, beam management, and more based on the sensed information [2, 6].	Reduce latency and improve system reliability.	√√√
Intelligent scheduling	Manage and schedule radio spectrum resources, network compute resources, network slices, and other resources based on the attributes and status of the environment, users, network, and terminals obtained from sensing and communication [2, 6].	Improve network resource utilization and save energy.	√
Ad hoc networking	Perform flexible networking for radio communication based on the attributes and status obtained from sensing [2].	Improve network performance.	√
Device-level energy saving	Adjust transmission configuration and selection based on the sensed information [2].	Save energy.	√

feasible verification solutions. To bridge this gap, this paper presents a rather comprehensive collection of SAC use cases with a categorization of them; it also puts forward corresponding technical schemes for each use case, clearly describing the current development of this technology and proposing specific directions for future research.

This paper categorizes SAC use cases into three groups: performance improvement, overhead reduction, and others. Sections 2, 3, and 4 describe the three groups in detail, covering their definitions, instances, and technical schemes. Section 5 discusses the prospects of SAC standardization and the challenges thereof. The Conclusion presents a paper summary.

2 Sensing-aided Communication Performance Improvement

Sensing-aided communication performance improvement refers to improving communication system capacity, improving spectral efficiency, and shortening communication delay using the environmental data obtained by sensing. The use cases include the widely studied sensing-aided communication channel estimation, sensing-aided radio resource management and scheduling, and more.

2.1 Sensing-aided Communication Channel Estimation

Though sensing and communication differ in channel mode — there are two modes of sensing, namely active sensing and passive sensing, while there is only one mode of communication, which is always unidirectional — they share some, if not all, scatterers in the environment. Therefore, the channel state information (CSI) obtained by sensing can reflect the status of the communication channel to some extent. The research of [6] proposes an ISAC-enabled channel model that integrates forward scattering and back scattering. Simulation has verified that the channel status features — such as delay spread, Doppler spread, and angle spread — obtained by sensing are consistent with those obtained in communication. This correlation lays the foundation for sensing-aided communication channel estimation.

Sensing-aided communication channel estimation refers to combining the information sensed about a device's environment, location, and status with the communication

information in order to estimate the channel environment more accurately, facilitating subsequent signal processing and improving the signal restoration rate.

The research of [8] presents a two-stage joint scheme for pilot optimization, target detection, and channel estimation. The scheme uses the target information obtained by sensing in conjunction with uplink channel estimation and applies the Sparse Bayesian algorithm, achieving efficient pilot optimization and channel estimation while ensuring the target detection rate.
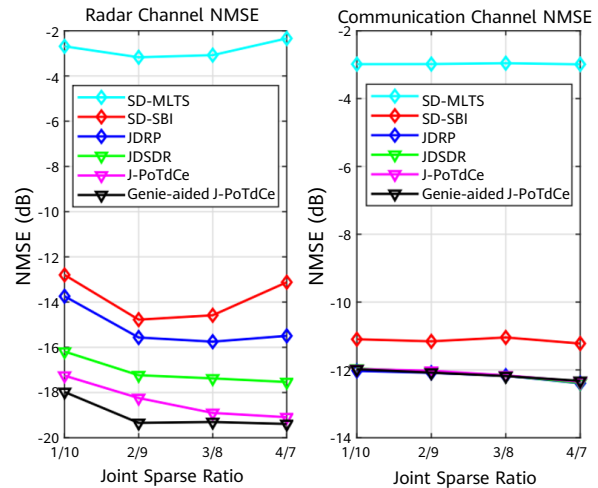
Figure 1 NMSE performance of channel estimation [8]

The research of [9] proposes an orthogonal time-frequency-space (OTFS) channel estimation method that uses the echo information sensed in the downlink. The method can save pilot resources. In addition, as the sensed signal-to-noise ratio (SNR) increases, the NMSE almost reaches the maximum likelihood estimation performance (the theoretical upper threshold) with significantly reduced complexity.
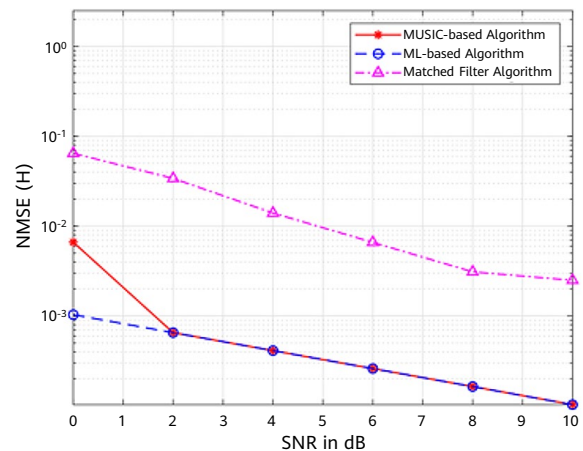
Figure 2 NMSE performance of sensing-aided OTFS channel estimation [9]

The research of [10] proposes a method that reduces the searching space for channel estimation of a sparse channel. The method uses the sensed location of scatterers to address challenges such as high computational complexity, hardware-led noise/array response errors, and off-grid instability. Eventually, it improves the estimation performance with limited measurements.

In addition, the sensed echo signal can be used to obtain the channel environment status and the number of main paths with a strong reflection in the environment. Such information can assist the communication receive end in selecting a channel estimation algorithm more accurately [2, 11]. Different from [8–10], this method focuses on adaptive selection of channel estimation algorithms in different scenarios instead of providing more information for channel estimation itself. But, like [8–10], this method can help improve the performance of communication channel estimation.

2.2 Sensing-aided Radio Resource Management and Scheduling

Sensing can also contribute to radio resource management and scheduling in communication. The location of end users is important for radio resource management and scheduling in communication; it can be used to effectively improve a communication system's performance. The research of [12] proposes a location-aided scheme of adaptive channel estimation and beamforming for millimeter waves to achieve ultra-fast initial access between nodes on the Internet of Vehicles (IoV). The scheme proves capable of speeding up the initial access process significantly. The research of [13] proves that location information allows a system to better understand the current channel status, thereby improving multi-user scheduling. This helps successfully decode the original information and, as a result, improves the throughput and downlink data rate in vehicle-to-infrastructure (V2I) systems. In addition, using Massive MIMO in ISAC systems can enable accurate sensing of the terminals' location, enhancing communication performance.

Sensing can aid in managing radio resources in communication in many ways. For example, the sensed user location and speed can be used for handover preparation to improve the handover success rate; the sensed location and/or speed of a crowd, a vehicle, an indoor device, and the like can be used to predict the requirements and consumption of time-frequency resources at communication nodes so that such resources can be preconfigured properly to accelerate time-frequency resource deployment; the sensed population density can be used to allocate radio resources at base stations.

Sensing can also aid in scheduling radio resources in communication. For example, the sensed user location can help determine a primary user and its neighboring interfering users, which helps allocate resources and suppress interference. The research of [14] suggests using the real-time sensed attributes and status of users and the environment to quickly manage, schedule, and make decisions on spectrum resources, network compute resources, and time/frequency/space domain resources. The research of [15] proposes to form dual-identity mapping
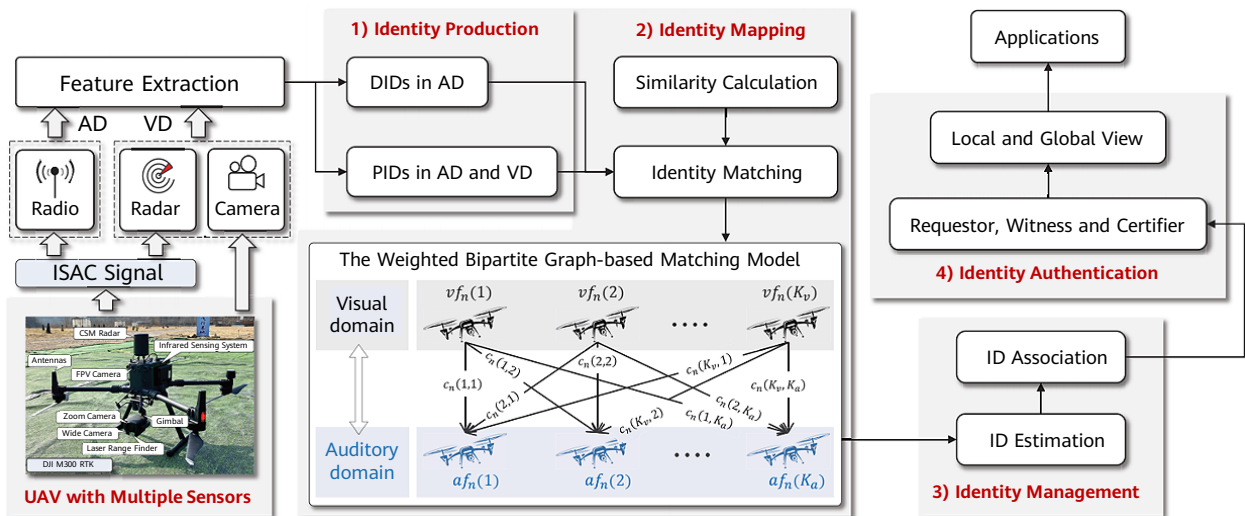
Figure 3 Framework of the ISAC-enabled dual identity solution [15]. It consists of four identity modules: production, mapping, management, and authentication.

between the physical identity, that is, the physical features obtained by sensing, such as UAV location (distance, speed, and angle) and wing type, and the digital identity such as IP/MAC address used by UAVs in communication. The solution reduces the emergency alerting latency by 66.54% on average compared to conventional feedback-based schemes. It also achieves an average of 31.55% increase in detection precision over conventional mobility-based approaches, effectively avoiding eavesdroppers' interference with the target user.

3 Sensing-aided Communication Overhead Reduction

Sensing-aided communication overhead reduction refers to reducing the radio resource overhead to be borne by a communications system by obtaining information required in communication through analysis of the sensed information, or preventing the waste of radio resources by adjusting the communication policy based on the sensed information. The use cases include the widely studied sensing-aided beam management overhead reduction, sensing-aided reference signal (RS) overhead reduction, and sensing-aided CSI overhead reduction, among others.

3.1 Sensing-aided Beam Management Overhead Reduction in Communication

Extensive research has been conducted on sensing-aided beam management in recent years. Radar sensing can obtain the location of terminals to assist in determining the direction in which the beams are pointing. This enables reliable beam alignment, effectively reducing the beam management overhead. The research of [16] suggests using "multipath fingerprints" for beam alignment, which are multipath channel characteristics highly correlated with location. When a terminal needs to determine its location, it obtains the multipath fingerprint of its current location using RF channel measurement. The terminal then compares the fingerprint against the fingerprint database before outputting its current location. The research proposes two types of fingerprints: type A and type B. Type A fingerprints, which capture the correlation between beam pairs, can be used with the beam selection method that minimizes the misalignment probability. Type B fingerprints, which store

the average received power of beams and hence enable beam ranking by power, can be used with the heuristic beam selection method. The optimal policy for reducing beam management overhead depends on the scenario.

Beam management on the IoV is an important SAC use case. In this scenario, the V2I link can be configured based on the sensed radar echo, reducing the overhead for communication beam tracking. The research of [17] proposes a radar-assisted predictive beamforming design for V2I communication, which utilizes ISAC-enabled road side units (RSUs) deployed on the road. The design uses an extended Kalman filtering (EKF) framework that builds upon measurements of echo signals and the vehicle status to track and predict the kinematic parameters of each vehicle. An optimization-based power allocation scheme is also proposed, further reducing tracking errors in the presence of multiple vehicles while ensuring communication quality. The proposed radar and communication-integrated beamforming design outperforms communication-feedback-only schemes, significantly reducing the overhead for communication beam tracking.
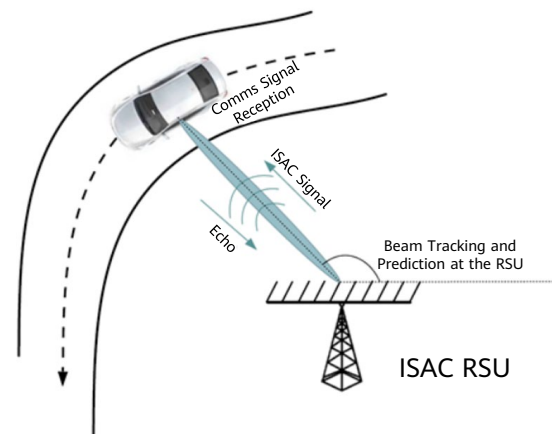
Figure 4 Radar-assisted IoV communication [18]

The research of [19] proposes a radar-aided mmWave communication scheme, which can be seen as an early form of ISAC system. In this scheme, beams can be managed based on radar sensing, reducing the communication overhead. Using different waveforms, radar and communication in this scheme can work on different frequency bands. This minimizes interference between radar and communication. Radio sensing is performed by compressed sensing-based positioning. With the sensed angle and position information, the channel estimation link and feedback link are no longer necessary for communication. The scheme illustrates that the overhead

for channel estimation, sector selection, and beam selection can all be reduced with radar sensing.
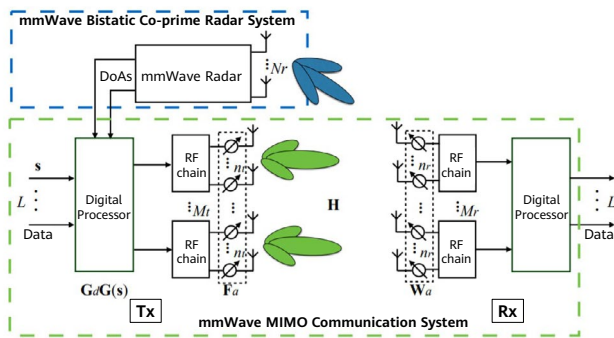
Figure 5 Diagram of the radar-aided mmWave communication system [19]

Beam sweeping can also be used for radar sensing of obstacles, eliminating the need for beam scanning in some directions and thereby reducing the beam management overhead. According to the research of [20], the angle information obtained by radar sensing can be used for beam training or beamforming so that pre-switching can be performed by detecting obstacles on the link, avoiding non-line of sight (NLOS) transmission.

Though sensing-aided beam management looks promising, there are many potential problems. The following describes some typical challenges and possible solutions.

One of the most commonly mentioned advantages of sensing-aided beam management is that communication beams can be configured based on the sensed terminal locations, which spares the need for beam scanning and thus reduces beam management overhead. However, sensing requires scanning to obtain the terminal locations, which also results in overhead. As a result, a single act of sensing-aided communication beam management will not reduce the overall beam management overhead. Even so, over the longer term, using the sensed information for prediction and tracking can save some overhead across the entire system.

Secondly, either active or passive sensing can generally obtain the optimal communication beam in only one direction. For example, active sensing at the base station can get the optimal downlink beam (from the base station to the terminals) but cannot obtain the optimal receive beam at the terminals. In such cases, the terminals can perform conventional communication beam scanning while the base station performs active sensing, or perform multi-antenna angle measurements to obtain the direction of beams from the base station and single out the optimal receive beam. In

this scenario, the base station and the terminals obtain their respective optimal beams independently, and thus feedback from the terminals is not required for beam selection at the base station. This eliminates the feedback overhead in conventional communication beam scanning.

Thirdly, current research generally assumes that only communication objects such as terminals exist in the environment, and any sensed object information is about communication objects, for example, their locations. This, however, is not true. Hence, a challenge to be addressed is distinguishing the passive objects from communication devices in the environment. An effective solution is the previously mentioned dual identity verification, that is, verification through combination of an object's physical identity such as location, and its IP address [15]. Another solution, inspired by the radio frequency identification (RFID) technology, which is commonly used in logistics and transportation, is to place a unique label on communication devices to distinguish them from other objects.

Finally, beam interference suppression is an inevitable issue in sensing-aided communication beam management. To tackle this issue, the research of [20] proposes to use phase-coded MIMO radar systems for beamforming and beam interference suppression. This can weaken the interference from communication signals to sensing signals, reducing the overhead.

3.2 Sensing-aided RS and CSI Overhead Reduction in Communication

RS is important for channel equalization and estimation. Different use cases require different RS densities, so when the demand for communication is low, there might be an oversupply of RS. This is where sensing-aided RS overhead reduction can come into play. It functions by configuring the RS density based on the sensed environmental data, specifically, increasing the RS density when the environment changes rapidly within the communication coherence time, decreasing the RS density when the environment is quasi-static, and providing no RS when the environment does not change. Thorough analysis is required to find suitable strategies for specific scenarios. For example, in smart transportation a road section such as a city-center crossing can be taken as the target area. The base stations covering this area or the RSUs within this area can be used as sensing devices to detect the traffic status. In the case of heavy traffic when the demand for communication is high,

more RS can be configured; in the case of light traffic, less RS can be configured. This reduces the RS overhead.

Like in channel estimation, the sensed information can be used with the communication information for pilot optimization. For example, the research of [8] proposes a scheme that uses the angle, radar channel, and communication channel obtained by sensing as prior information and applies the majorization-minimization method to achieve efficient pilot optimization. The scheme proves capable of improving the overall system performance while keeping the algorithm complexity essentially the same as conventional algorithms.

CSI can be classified into transmit end CSI and receive end CSI. Generally, the former plays a more significant role. This is because the transmit end can take measures such as power allocation, beamforming, and antenna selection in advance based on CSI on its end to compensate for attenuation in transmission, ensuring high-speed and reliable data transmission. In ISAC systems, CSI can be obtained by active sensing. For example, the sensed information, such as distance, speed, and angle, can be converted into delay spread, Doppler spread, and angle spread. This replaces the CSI feedback in conventional communication systems, leading to reduced overhead. The research of [18] proposes a scheme in which echo signals replace frequent channel feedbacks in communication, thereby minimizing channel quantization errors, reducing repeated overheads, and improving the accuracy of channel selection. Sensing-aided CSI overhead reduction can also work by configuring the modulation and coding scheme (MCS) level, channel quality indication (CQI) feedback quantity, and CQI feedback period based on the sensed cell information. Specifically, it means applying a lower MCS level or a longer CQI feedback period for cells with infrequent environmental changes to reduce the feedback overhead.

4 Other Aspects of SAC

Other aspects of SAC include RIS-assisted SAC, sensing-aided artificial intelligence (AI) modeling and optimization, and sensing-aided communication power control, among others. The following sections discuss the increasingly popular RIS-assisted SAC and sensing-aided AI modeling and optimization. Sensing-aided communication power control works by optimizing the multi-user power distribution strategy or allocating power to priority users based on the sensed number and status of users in the environment, meeting various requirements of communication services.

4.1 RIS-assisted SAC

RIS has unique advantages such as low cost, low power consumption, and easy deployment. It is a key technology in the massive antenna field.

With RIS, communication performance can be improved directly through joint optimization of the reflection coefficients. In addition, RIS can be used as a wireless relay sensor to flexibly control the channel environment, enabling NLOS sensing; NLOS sensing makes use of NLOS degrees of freedom (DoFs), which a radar cannot use, to provide additional sensing paths. This improves the sensing coverage and robustness [21], contributing to the acquisition of higher-accuracy channel information for facilitating decision-making in communication. Ultimately, this leads to reduced overhead.

Using RIS in ISAC systems can boost flexibility in dynamic scheduling of system resources. In beam management, RIS allows the separation of communication and sensing beams, enabling the system to better meet communication and sensing targets. When the demand for communication is high, the working mode of RIS units can switch based on the sensed information so that the ISAC systems can allocate more beams for communication, enhancing communication performance [22].
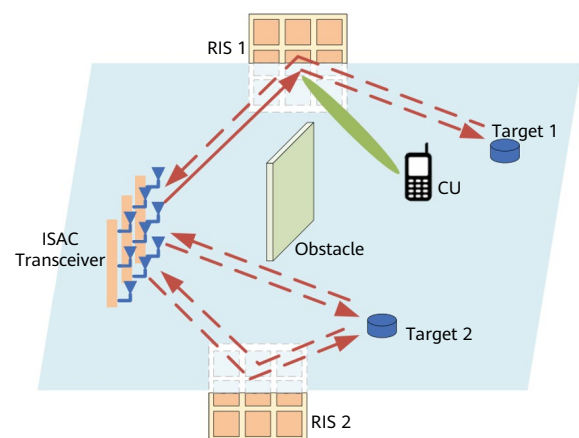
Figure 6 Diagram of an RIS-assisted ISAC system [18]

When the communication quality is unsatisfactory, the systems can either select the optimal RIS based on the sensed RIS locations to identify NLOS paths of higher communication quality or use active RISs to enhance the communication signal.

4.2 Sensing-aided AI Modeling and Optimization

Ongoing communication system upgrades have added new functions and exponentially increased the amount of data, accelerating the development of machine learning (ML)-based AI technologies. So far, AI has been applied to various communication network layers, including the physical, data link, and network layers. The sensed information in ISAC systems can be used as data input for AI modeling and optimization, improving the accuracy of constructed models and enhancing the quality of the system's decision-making. Take the IoV scenario as an example. The sensed information, properly treated by AI and edge computing technologies, can enhance communication performance, promote integration of sensing and communication, and even promote integration of sensing, communication, and computing.

In addition, the sensed information can be processed by AI algorithms to turn into high-precision information, which can be converted into data usable in communication to improve communication performance. The research of [23] proposes four optimization algorithms to decollate data sensed indoors to reconstruct the environment with high precision. Thus the communication channel can be more accurately estimated by tracking the propagation relationship between reflectors.
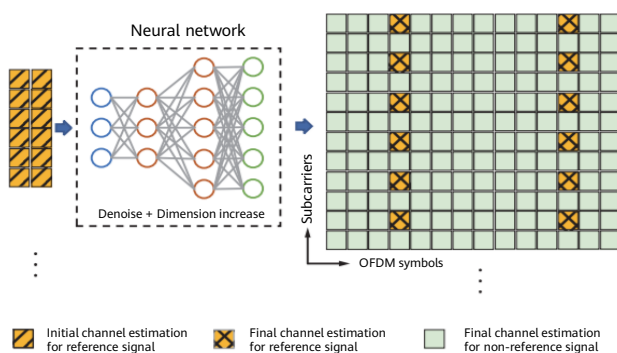
Figure 7 AI-based OFDM channel estimation [24]

Though using AI can improve the channel estimation accuracy, the generalization performance is relatively poor because the estimation accuracy relies heavily on data validity. Changes in the environment will lead to declines in estimation accuracy. The research of [24] proposes a scheme that integrates transfer learning, joint training, and model-independent meta-learning for AI-based channel estimation to improve AI models' generalization performance. To cope with the generalization issue that occurs when AI is applied to channel estimation in ISAC systems, the solution can be either to improve the sensing performance to obtain relatively accurate sensed information or to use higher-performing AI training and learning schemes to enhance the estimation accuracy.

Furthermore, AI can be applied to RIS. Specifically, AI algorithms can replace conventional optimization algorithms to optimize the RIS reflection coefficients, which are discrete values usually available in large amounts in RIS-assisted ISAC systems. This can lead to reduced system complexity [25].

5 Prospects of SAC Standardization and Technical Challenges

As an application in the deeper convergence phase of ISAC, SAC can further improve the overall system efficiency and advance "multilateral benefits" in systems. However, due to limitations in the development of ISAC, SAC standardization is not yet within sight. Apart from that, SAC itself is facing quite many technical challenges.

5.1 Prospects of SAC Standardization

SAC standardization within the 6G framework depends on how well ISAC is incorporated into 6G standards. Presently, many organizations and industries around the world are actively promoting ISAC standardization [2]. For example, the 3GPP initiated the *Study on Integrated Sensing and Communication* project in the SA1 phase in May 2022. In China, the IMT2030 working group set up a joint testing team, including representatives from the sensing and communication fields in April 2022; the engaged operators and providers are actively participating in work related to ISAC for 6G, such as channel modeling, evaluation indicators selection, and prototype testing. As ISAC standardization trends on, SAC, which falls within phase two of ISAC (three phases in total), is an important element in this initiative.

To standardize SAC, a concerted effort is required. Active work must be done to advance the standardization of ISAC. Also, it is crucial to thoroughly explore SAC use cases and their potential benefits for industries and businesses. For example, in 6G+industrial Internet, sensing can capture the movement of robots, assisting the systems in robot dispatching. Use cases like this can promote the development of SAC technologies and eventually advance SAC standardization.

5.2 Technical Challenges Faced by SAC

SAC mainly faces three technical challenges: performance tradeoff, networking, and performance measurement.

5.2.1 Performance Tradeoff

An ISAC system has only limited amount of resources, including spectra and devices. Allocating these resources in the implementation of SAC must consider the performance targets in question. Communication and sensing are meant for different purposes. Without introducing external forces such as RIS and AI, communication and sensing performance tend to have a reverse correlation [26], meaning that when one goes up, the other generally goes down. If more resources are allocated for sensing, fewer will go to communication, affecting communication performance. Therefore, the performance bound of ISAC with a tradeoff between sensing and communication must be defined so that the optimal resource scheduling can be realized in any given scenario [21]. Another way to address the performance tradeoff challenge is to converge sensing from other sources, such as laser radar, sensors, and cameras, with radio sensing. While such convergence can reduce the burden on radio sensing in ISAC, it embodies challenges such as feature sparsity, high complexity, and poor real-time performance, to name a few. In view of this, a framework more compatible with multi-source sensing convergence must be established, and higher-performing algorithms, such as ML algorithms, must be developed [25].

Another thing to mention is that communication and sensing use spatial resources differently. Communication can perform on both NLOS [28] and line of sight (LOS) paths, while sensing can perform only within the DoFs, which depend on LOS components [18]. This must be considered when implementing SAC and using RIS can somewhat mitigate this issue.

5.2.2 Networking

Networking is a crucial challenge in SAC implementation. To improve communication performance, thought must be given on how to utilize the sensed information in echo signals for traffic prediction and how to schedule resources and deploy drones, RISs, and other networking nodes based on the prediction [29]. Also, self-interference and cross-interference in communication and sensing require special consideration. Conventional radar systems do not support networking and rely on the anti-interference ability built in the devices to suppress cross-interference [30]. In this regard, coordinated beamforming and tracking can be a solution for interference suppression.

5.2.3 Performance Measurement

The channel capacity of a communication system is bounded by the Shannon limit [31], while the performance of a radar system is subject to the Cramér-Rao lower bound (CRLB) [32]. Therefore, a proper SAC performance evaluation method should evaluate communication performance on the precondition that sensing performance, such as sensing reliability, precision, and range, meets expectations. Only then will the evaluation results be meaningful. To that end, a complete ISAC performance measurement system needs to be established, in which arithmetic methods and AI technologies can be used to explore the system performance bound. Such a measurement system will provide a universal framework for ISAC performance analysis and an evaluation method for ISAC implementation. It will also help set system performance targets and boost the accuracy of performance evaluation [20].

6 Conclusion

This paper centers on SAC, a typical application in the deeper convergence phase of ISAC. It first analyzes the significance of SAC research. Then, it discusses the benefits of SAC from three perspectives: performance improvement, overhead reduction, and others, covering their definitions, instances, and technical schemes. The discussion illustrates that sensing-aided communication channel estimation and sensing-aided communication beam management currently have higher feasibility. The paper concludes by discussing the prospects of SAC standardization along with the challenges faced by SAC in performance tradeoff, networking, and performance measurement. In summary, research on SAC is still in its infancy. Continued efforts are required to explore its key technologies and use cases to enhance overall ISAC performance and foster potential applications for 6G industrial upgrading.

Acknowledgments

References

[1] ITU-R, DRAFT NEW RECOMMENDATION, "Framework and overall objectives of the future development of IMT for 2030 and beyond," June 2023.

[2] IMT-2030(6G) Promotion Group, "Research report on integrated sensing and communication[R]," Nov. 2022.

[3] J. Yu, J. Ding, J. Zhang, *et al.*, "Broadband photon-assisted terahertz communication and sensing[J]," *Journal of Lightwave Technology*, 2023, 41(11):3332–3349.

[4] "6G white paper on localization and sensing[R]," 6G-Flagship, 2020.

[5] "ICDT integrated 6G network 2.0[R]," Future Mobile Communication Forum, 2022.

[6] Shiqiang Suo, Shenghao Xu, Ting Zeng, *et al*., "Study on integrated sensing and communication for 6G RAN[J]," *Telecommunications Science*, 2022, 38(9):6.

[7] R. Yang, C. Wang, J. Huang, *et al*., "A novel 6G ISAC channel model combining forward and backward scattering[J]," *IEEE Transactions on Wireless Communications*, 2023.

[8] Z. Huang, K. Wang, A. Liu, *et al*., "Joint pilot optimization, target detection and channel estimation for integrated sensing and communication systems[J]," *IEEE Transactions on Wireless Communications*, 2022, 21(12):10351–10365.

[9] W. Wang, Z. Fei, S. Wang, *et al*., "Sensing assisted communication system design based on OTFS waveforms[J]," *IEEE Wireless Communications Letters*, 2023:1–5.

[10] L. Xie, S. Song, Y. Eldar, *et al*., "Collaborative sensing in perceptive mobile networks: Opportunities and challenges[J]," *IEEE Wireless Communications*, 2023:16–23.

[11] Jing Yang, "Butong xindao huanjing xia de OFDM zuiyou xindao guji suanfa" 不同信道环境下的 OFDM 最优信道估计算法 [D] [Optimal OFDM channel estimation algorithms in different channel environment], Harbin Institute of Technology, 2023.

[12] N. Garcia, H. Wymeersch, E. Ström, *et al.*, "Location-aided mm-Wave channel estimation for vehicular communication[C]," 2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Edinburgh, England, 3–6 July, 2016.

[13] Z. Li, S. Yang, and T. Clessienne, "Exploiting location information to enhance throughput in downlink V2I systems[C]," IEEE Global Communications Conference, 2019.

[14] Tianyu Hu, Lingxiang Li, Zhi Chen, *et al.*, "Sensing-assisted tera-hertz wireless communication[J]," *Designing Techniques of Posts and Telecommunications*, 2021(12):5.

[15] Y. Cui, Z. Feng, Q. Zhang, *et al.*, "Toward trusted and swift UAV communication: ISAC-enabled dual identity mapping[J]," *IEEE Wireless Communications,* 2023:58–66.

[16] V. Va, J. Choi, T. Shimizu, *et al.*, "Inverse multipath fingerprinting for millimeter wave V2I beam alignment[J]," *IEEE Transactions on Vehicular Technology*, 2017, 67(5):4042–4058.

[17] F. Liu, W. Yuan, C. Masouros, *et al.*, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing[J]," *IEEE Transactions on Wireless Communications*, 2020, 19(11):7704–7719.

[18] F. Liu, Y. Cui, C. Masouros, *et al.*, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond[J]," *IEEE Journal on Selected Areas in Communications*, 2022, 40(6):1728–1767.

[19] Z. Chen, Z. Cao, X. He, *et al.*, "DoA and DoD estimation and hybrid beamforming for radar-aided mmwave MIMO vehicular communication systems[J]," *Electronics*, 2018, 7(3):40.

[20] Mugen Peng, Xiqing Liu, Zile Liu, *et al.*, "Principles and techniques in communication and sensing integrated 6G systems[J]," *Control and Decision*, 2023, 38(01):22–38.

[21] Ruoyu Zhang, Weijie Yuan, Yuanhao Cui, *et al.*, "Mianxiang 6G de da guimo MIMO tongxin ganzhi

yitihua：xianzhuang yu zhanwang" 面 向 6G 的 大规模 MIMO 通信感知一体化：现状与展望 [J] [Massive MIMO-based integrated sensing and communication for 6G: The present and the future], *Mobile Communications*, 2022, 46(06):17–23.

[22] Fanghao Xia, Xinyi Wang, and Zhong Zheng, "RIS-assisted integrated sensing and communications[J]," *ZTE Technology Journal*, 2022, 28(03):58–62.

[23] Jia He, Zhi Zhou, Xianjin Li, *et al*., "6G integrated sensing and communication: Wireless sensing and sensing assisted communication[J]," *Information and Communications Technology and Policy*, 2022(09):9–17.

[24] Bule Sun, Ang Yang, Peng Sun, *et al*., "Approaches for improving generalization capability of AI based channel estimation," *Radio Communications Technology*, 2022, 48(04):652–657.

[25] Rang Liu, Honghao Luo, and Ming Li, "Applications of reconfigurable intelligent surface for integrated sensing and communication," *ZTE Technology Journal*, 2022, 28(03):53–57, 69.

[26] X. Yuan, Z. Feng, J. Zhang, *et al*., "Spatio-temporal power optimization for MIMO joint communication and radio sensing systems with training overhead[J]," *IEEE Transactions on Vehicular Technology*, 2021, 70(1):514–528.

[27] J. Xue, D. Wang, S. Du, *et al*., "A vision-centered multi-sensor fusing approach to self-localization and obstacle perception for robotic cars[J]," *Frontiers of Information Technology & Electronic Engineering*, 2017, 18(1):122–138.

[28] Chi Yuan and Bo Huang, "UWB dingwei zhong jiyu XGBoost de NLOS shibie fangfa" UWB 定 位 中 基 于 XGBoost 的 NLOS 识 别 方 法 [J] [XGBoost-based NLOS recognition in UWB positioning], *Mobile Communications*, 2023, 47(05):83–89.

[29] Chengkang Pan, "Reflections on integrated sensing and communication technology," *ZTE Techonology Journal*, 2022, 28(05):53–56.

[30] Xi Nan, Rugui Yao, Ye Fan, *et al*., "Interference cancellation algorithm of NOMA based integrated sensing and communication," *Journal of Signal Processing*, 2023, 39(06):986–995.

[31] Yawen Liu, Wenbo Ma, and Zifang Du, "Unified precision measure and its relationship with Shannon's Theorem," *Journal of Systems Science and Mathematical Sciences*, 2022, 42(01):64–71.

[32] Dazhuan Xu, Yue Chen, Yueshuai Chen, *et al*., "Multi-carrier radar system information and Cramer-Rao Bound," *Journal of Data Acquisition & Processing*, 2020, 35(06):1011–1021.

Environment Sensing in Wireless Cellular Networks: Models, Architectures, and Methods

Xin Tong, Kang Guo, Ziqing Xing, Zhaohui Yang, Zhaoyang Zhang

College of Information Science and Electronic Engineering, Zhejiang University

Abstract

Integrated sensing and communication (ISAC) is a key research area in future wireless communications technologies. ISAC can sense environmental information using the existing wireless communications architecture. This paper analyzes and discusses the key challenges in designing an environment-sensing system in wireless cellular networks. First, to accurately sense the environment, we analyze wireless channel modeling methods and establish a deterministic channel model for environment sensing that considers the impact of objects in the environment on wireless channels. Second, we propose a multi-view architecture and two design methods for the wireless environment sensing system, namely three-dimensional (3D) environment reconstruction based on the compressed sensing (CS) theory and moving/stationary object detection based on the millimeter wave (mmWave) radar. In addition, we summarize the advantages and disadvantages of the two methods. Finally, we test our environment sensing algorithm on the mmWave communications hardware development platform and verify the feasibility of the proposed ISAC system.

Keywords

1 Introduction

The rapid growth of the wireless communications industry has exponentially increased the number of connected devices and services. There are also more and more use cases for wireless mobile communications. Consequently, the propagation environment of wireless communications signals has become increasingly complex [1].

Millimeter wave (mmWave) delivers high bandwidth, reliability, and integration. In addition, because mmWave features short wavelengths, its antenna spacing is significantly shorter than in the sub-6 GHz frequency band, making it possible to arrange a large-scale antenna array.

Wireless communications base stations are densely deployed. Also, the receiver and sender have more powerful computing capabilities and physical performance than before. As a result, our information processing capability has dramatically improved, and communications devices can complete more tasks.

The integrated sensing and communication (ISAC) system can provide environment sensing services, such as hand gesture recognition, object tracking, and three-dimensional (3D) environment reconstruction, in addition to traditional communications services. Many essential use cases, such as smart homes, autonomous driving, and the positioning of uncrewed aerial vehicles, require high-precision environment sensing services. Therefore, sensing and communications systems must be jointly designed to share the same frequency band and hardware, improving spectral efficiency and reducing hardware costs. Also, environment sensing is an essential means of estimating channels and assisting communications.

Integrating the environment sensing function into a mature communications system architecture is a primary challenge. Compared with the communications process between cooperative transceivers, the radar system sends detection signals to unknown objects and parses the environmental information contained in the echo.

However, the two systems also present certain similarities. Channel state information (CSI) obtained in the communications process usually contains specific environmental knowledge. Similarly, the environment sensing result helps improve channel estimation accuracy and enhance communications performance. This suggests that the communications system could implement environment sensing in the communications process [2].

As wireless networks operate at higher frequencies and wider bandwidths, and more base stations and antennas are deployed, the long-neglected sensing capability derived from electromagnetic wave propagation in the communications system is increasingly worth exploring. The signal processing principles used in radar and communications are gradually becoming similar. This paves the way for ISAC technology, as envisaged in [3].

Two major challenges need to be addressed in addition to the system architecture design if we want to integrate communications and sensing.

First, there may be many unknown variables in the environment, including unknown objects, speed, and electromagnetic characteristics; there are even more variables to consider than in the statistical channel model. Detecting the many dimensions of environmental information requires a lot of ISAC system resources and computing power. To address this problem, we try to use the sparseness of the environment to obtain environmental information. For example, outdoors, buildings in a cellular communications network are sparsely distributed within the coverage area; and indoors, articles such as furniture are sparsely distributed in a room. Both in traditional communications and sensing, the compressed sensing (CS) [4] framework fully utilizes the sparseness of various variables, achieving a better solution with lower complexity. In computational imaging, using the inherent sparseness of objects or scatterers in the environment is the key to effective detection. According to the CS theory, such a problem is usually modeled as a sparse signal recovery problem based on pixel division, and it is solved by using widely used CS reconstruction algorithms, such as sparse Bayesian learning (SBL) [5], orthogonal matching pursuit (OMP) [6], and generalized approximate message passing (GAMP) [7].

The second challenge is the complex propagation characteristics of electromagnetic waves. In wireless communications scenarios, the blocking effect between objects in the environment blocks the propagation of electromagnetic waves. As a result, some areas are invisible to a specific signal transceiver. When sensing the environment, each transceiver can represent only a partial view of the environment. It is often difficult to distinguish which part of the environment affects the received signals. In addition, other characteristics of electromagnetic propagation, such as reflection, scattering, and diffraction, cause the electromagnetic characteristics of the observed objects to be coupled with the unknown surrounding environment and the irradiation of electromagnetic waves.

To solve these issues in the environment sensing system of a wireless cellular network, this paper discusses and analyzes the key technologies involved in the system design. First, we discuss various wireless channel modeling methods for environment sensing and determine the most suitable one. Next, we propose two design methods for the wireless environment sensing system, namely 3D environment reconstruction based on the CS theory and moving/stationary object detection based on mmWave radars. The advantages and disadvantages of these two methods are also summarized and compared. Finally, we verify the designed environment sensing algorithm on the hardware development platform.

2 Wireless Channel Modeling for Environment Sensing

To sense the environment of a wireless communications system, we must fully utilize the radio signals, such as Wi-Fi signals or cellular network signals in the environment. This will help us achieve ISAC based on the existing communications process, network, devices, and signals.

CSI offers a lot of value in the existing communications architecture because it contains rich environmental information. CSI is generated during channel estimation, an important step that must be performed periodically to overcome random channel fading. Therefore, CSI data can be obtained in a normal data communications process without significant hardware or software adjustments to the existing communications system, and algorithms can be designed to sense objects in the environment [8–10].

When creating a CSI-based environment sensing algorithm, a channel model must be designed to accurately depict the relationship between environmental information and CSI. A traditional channel model is generally based on the statistical features of channels. For example, in [11], a measured power-delay feature and a measured power-angle feature describe a 60 GHz indoor channel. Deterministic modeling methods, such as ray tracing, have been applied to communications channel simulation in research to generate a large amount of reliable channel data and serve wireless artificial intelligence (AI) tasks. However, the electromagnetic propagation model in mmWave sensing works in an ideal manner and cannot accurately reflect the impact of objects in the environment on wireless channels. Therefore, it is necessary to further explore the relationships

between the mmWave propagation process and the objects in the environment, and then design the corresponding environment-sensing algorithm.

Compared with sub-6 GHz electromagnetic waves, mmWave has larger free-space path and wall-penetration losses, and edge diffraction is more difficult. As such, mmWave has apparent blocking effects. In addition, mmWave is prone to scattering on surfaces whose roughness is proportional to their wavelengths; this part of signal energy can no longer be ignored [12]. Therefore, before designing sensing algorithms, we need to study the propagation characteristics of mmWave and establish a mathematical model of mmWave channels.

Our modeling method of mmWave reflection and scattering is inspired by the optical characteristics of objects and depicts the electromagnetic characteristics on the surface of a scatterer in the environment, as shown in Figure 1. A mmWave channel consists of a reflection path (red solid line) and multiple scattering paths (green dotted line). Based on this modeling method, an improved ray tracing method [13] is proposed, as shown in Figure 2.
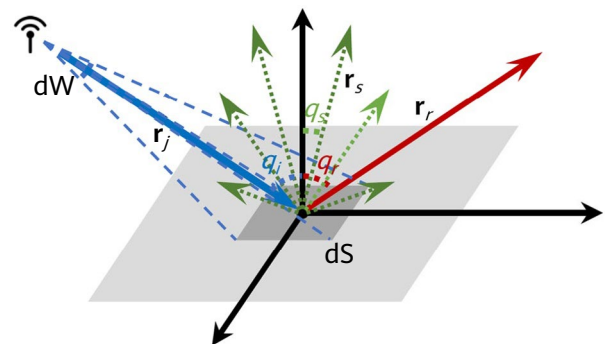
Figure 1 Reflection and scattering of mmWave at a surface element in a 3D scenario
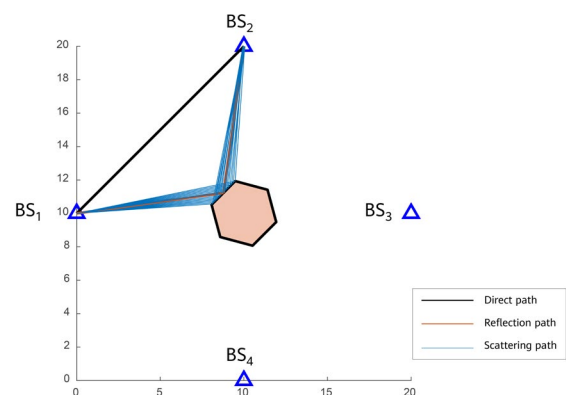
Figure 2 Improved ray tracing method

Ray tracing is a signal propagation modeling tool that can estimate path loss, angle of arrival (AoA), angle of departure (AoD), and delay. Ray tracing algorithms include the image and the shooting and bouncing ray (SBR) methods. The image method accurately calculates the reflection path using the image point of the emission source but does not consider the scattering path. In essence, the SBR method traces each ray emitted from the source to determine whether they reach the reception points. This method consists of three steps: ray emission, ray tracing, and ray reception. The method proposed in this paper simultaneously takes into account the direct path, reflection path, and scattering path. Compared with the traditional ray tracing method, our method more accurately models the environmental information in wireless channels. It uses an efficient and enhanced electromagnetic calculation tool to quickly generate the synthesis data of multiple reflection, refraction, transmission, scattering, and diffraction signals of light and mmWave. It does so based on the 3D spatial layout, transceiver array characteristics, material characteristics, object motion characteristics, and model precision constraints for any environment and object in the simulation and data acquisition stage. The results provide data for learning correlation models and conducting cooperative inversion.

3 Wireless Environment Sensing System Design

3.1 CS-based Wireless Sensing

One or more base stations and servers are deployed in a wireless communications scenario. The environment has multiple active UEs, including mobile phones, vehicles, and drones, that communicate with the base stations. The base stations receive uplink communications signals sent by each UE. These signals are reflected, scattered, and blocked by objects in the environment and transmitted to the base stations through multipath channels. Consequently, the received signals contain the environmental information that needs to be sensed. However, as UEs and base stations are located in different places, the physical impact of the environment on channels varies. As a result, signals sent by different UEs to different base stations contain environmental information from multiple views. So, these signals cannot be simply combined to achieve environment sensing.

We consider two sensing scenarios. In Figure 3, a single base

station with a central controller processes signals sent by multiple UEs. The signals of each UE are transmitted to the base station through multipath channels caused by various objects in the environment, as shown by colored arrows and beams in the figure. As a result, the received signals of the base station contain environmental information observed from multiple UE views. In a single-base-station centralized architecture, one base station processes signals from multiple UE views to sense the environment.

Figure 4 shows a scenario involving a plurality of base stations and UEs. Multiple base stations with distributed servers process signals sent by multiple UEs. Due to the physical propagation characteristics of electromagnetic waves, each base station can receive only the signals reflected by some objects in the environment and, therefore, has its unique sensing range. In a multi-base-station distributed architecture, multiple base stations coordinate to process data from multiple views and jointly sense the entire environment. Compared with the single-base-station centralized architecture, the multi-base-station distributed architecture effectively expands the sensing range. Still, it also demands a method of distributed information processing and convergence among the base stations.
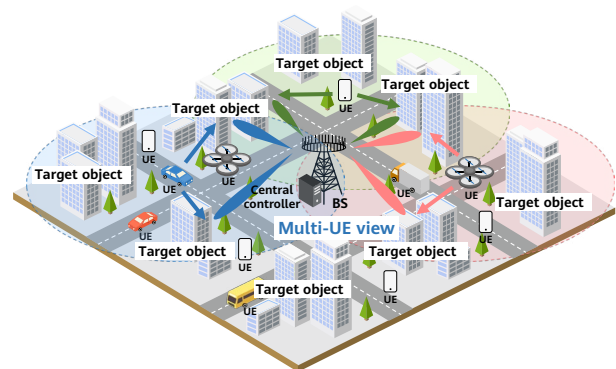
Figure 3 Architecture of multi-view sensing involving one base station and multiple UEs
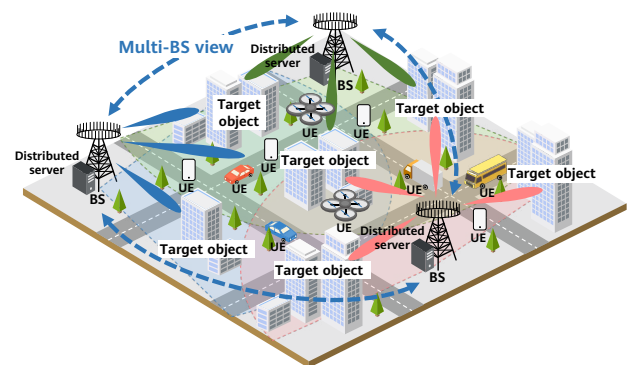
Figure 4 Architecture of multi-view sensing involving a plurality of base stations and UEs

This paper provides a centralized environment sensing method with multiple views. We divide the target environment into point clouds and then use their sparseness to convert the environment-sensing problem into a CS reconstruction problem [14, 15]. The bilinear GAMP inference method based on a factor graph (Figure 5) is a specific reconstruction algorithm. The two groups of yellow circular variable nodes represent environmental information variables and unknown views. In contrast, the three groups of blue square factor nodes represent the prior probability of sparse environmental information, received signals affected by environmental objects and views, and the physical relationship between environmental objects and views. Based on the structure of the factor graph, a sum-product algorithm (SPA) is used to obtain the environmental information iteratively. The lower part of Figure 5 briefly shows a CS-based wireless sensing result. To extend this centralized method to a distributed method, factors of different sensing nodes can be allocated based on the factor graph, and information transmitted between sensing nodes can be designed to develop multi-point distributed collaborative sensing algorithms suitable for mmWave scenarios.
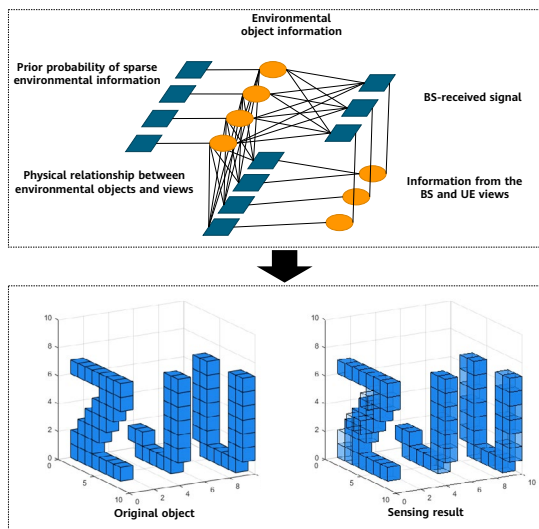
Figure 5 Multi-view sensing factor graph and simulation result

3.2 mmWave Radar–based Wireless Sensing

mmWave radar has been repeatedly studied, tested, and applied to real-world scenarios. Evolving hardware has increased the working frequency of radar. mmWave radar is playing an increasingly important role in wireless communications and intelligent vehicle fields. The band used by the current wireless communications system is also a mmWave band. Therefore, the spatial sensing capability of mmWave radar and the corresponding signal design and processing algorithm provide an important basis for ISAC. The ISAC system can implement sensing in a similar manner to radar signal processing when the computing capability permits.

Linear frequency modulated continuous wave (FMCW) radar has attracted increasing attention and has been widely used in recent years. Compared with narrow-band radar and ultra-wideband radar, FMCW radar has a simpler hardware structure along with a higher resolution and accuracy. Compared with pulsed radar, which requires a voltage-controlled oscillator to generate high-power narrow pulse signals quickly, FMCW radar does not require a quick jump of sent signals between high and low frequencies. So, it requires less advanced hardware and costs less. FMCW radar measures objects by using the constantly changing modulation signals to control the generation of the constantly changing sent signals. FMCW radar linearly modulates the frequency of continuous waves, thus reducing the system's sampling rate. With this structural feature, information such as frequency and phase can be obtained quickly, and the Fourier transform can be integrated into the hardware to extract information such as the distance, speed, and angle of objects. The process is low-cost, simple, and yields high precision.

All this said, there are still some challenges to overcome in today's radar technology. First, an offline framework is used to process FMCW radar data in real time and release the burden on firmware memory; this is expensive and time-consuming. Second, when a moving object blocks some part of the static background, signals reflected by the blocked background are affected by the moving object. As a result, the blocked static background will be incorrectly identified as moving objects in the Doppler domain. Third, due to the spectral leakage of static background in Doppler fast Fourier transform (FFT), an obviously static background with large radar cross sections can also lead to false object detection. Currently, the algorithms that process only single-frame data cannot solve these problems. It is still a challenge to use the correlation of multi-frame data instead of extracting information within a single frame to achieve more accurate and robust radar sensing.

To solve these difficulties, we propose an online tensor-based robust principal component analysis (RPCA) algorithm to separate the preprocessed FMCW radar data tensor into the static background and moving objects [16]. An online framework helps us efficiently process new data frames collected by radar, reducing the burden

on storage and processing. The tensor-based algorithm maintains the internal structure of the data tensor. The low-rank regularization of static components utilizes time correlation. The 3D total variation regularization of dynamic components utilizes spatiotemporal continuity.

As shown in Figure 6, real measurement data is collected in our experiment to verify the algorithm's effectiveness. The TI AWR 2243 mmWave FMCW radar we use can operate at the 76–81 GHz band in TDM mode. Three corner reflectors (CRs) and a metal cuboid are placed indoors as a static background. There is a path in the room, and a pedestrian walks along the path as a moving object. The pedestrian will block CR 1 when walking forward.

Figure 7 shows the comparison result of the Doppler FFT and the online tensor-based RPCA. The red boxes indicate the positions of CR 1. In Figure 7a, CR 1 is not found by Doppler FFT through static background detection. However, in Figure 7c, the RPCA algorithm can detect CR 1 and identify it as a static background, even if the pedestrian blocks it in the frame. In Figure 7b, Doppler FFT can also detect CR 1, but it incorrectly identifies CR 1 as a moving object. In Figure 7d, the RPCA algorithm correctly determines that CR 1 is not a moving object.
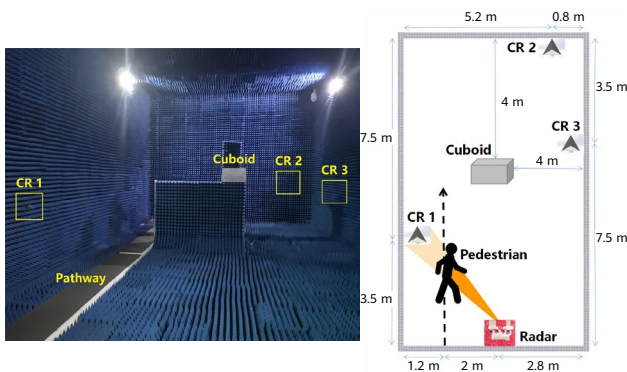
Figure 6 Experiment scenario settings
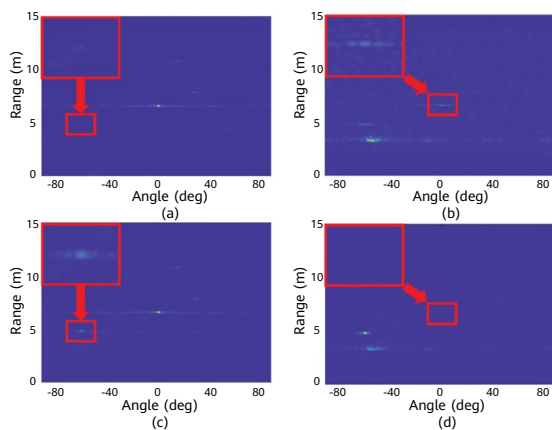
Figure 7 Comparison between Doppler FFT and online tensor-based RPCA

3.3 Summary of Technical Characteristics

CS-based wireless sensing is more convenient to deploy in wireless cellular networks, and its architecture is highly compatible with the existing wireless communications system. In addition, it has a relatively small interference and impact on the existing communications process and can implement precise 3D point cloud reconstruction for objects. That said, CS-based wireless sensing requires a lot of computing and storage resources for point cloud processing and imaging. Also, the performance of CS sparse reconstruction deteriorates as the number of objects in the environment increases.

There are several advantages to introducing radar signal processing into wireless sensing: (1) Radar has an extensive sensing range; (2) The radar system structure and signal processing are simple; (3) Radar is especially good at detecting and tracking moving objects. However, the following disadvantages also need to be addressed: (1) The combination of radar and communications signals requires a new design solution; (2) The ISAC system requires additional hardware for radar systems; (3) An extra collaborative processing solution for received signals needs to be designed for the ISAC system.

4 ISAC Platform Verification

As shown in Figure 8, the ISAC verification platform utilizes the LabVIEW language for system development. It receives and sends standard OFDM signals in compliance with IEEE 802.11ac based on the NI mmWave system. Transceiver antennas are deployed on both sides of the room. A 64-antenna uniform planar array is used to irradiate the environment. The environment is sensed by adjusting the reconfigurable intelligent surface (RIS) consisting of 1024 units and the received signals. A metal object is placed between transceiver antennas as an object to be sensed.
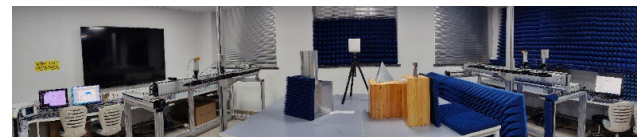
Figure 8 ISAC verification platform and its application scenarios

4.1 Signal Generation and Processing

A host computer utilizes the LabVIEW language to program and generate self-defined baseband signals. The computer

transmits control signals and waveform data to each field programmable gate array (FPGA) through the PXIe bus. The PXI-7902 FPGA encodes and modulates baseband signals, processes complex data, generates complete baseband digital signals, and transmits the data to the PXI-3610 FPGA with a sampling rate of 3.072 GS/s for digital-to-analog conversion. The generated analog IQ signals are up-converted by the PXI-3620 FPGA to 11 GHz IF signals and then up-converted by the RF front end again to 28 GHz RF signals. Similarly, after receiving the signals, the receiver shifts the signals from 28 GHz to 11 GHz by using the RF front end. The receiver then injects the signals to the PXI-3620 FPGA for down-conversion to generate baseband analog signals. The PXI-3630 FPGA converts the analog signals into the digital signals required for processing at the same sampling interval of 3.072 GS/s. The PXI-7902 FPGA and the host computer at the receive end complete subsequent processing such as synchronization, channel estimation, and demodulation.

In the preceding signal generation process, UE data can be mapped in modulation modes such as BPSK, QPSK, 16QAM, or 64QAM; and 64, 256, 512, or 1024 subcarriers can be used. The receiver performs complete timing synchronization, frequency offset estimation and compensation, and channel estimation and equalization; it also demodulates the sent UE data and estimates the environment information using the sensing algorithm.

4.2 Environment Sensing Algorithm Based on Sidelobe Beam Sweeping

In mmWave communications scenarios, beamforming technology is widely applied to devices, such as vehicle-mounted phased array and large-scale base station array, to improve signal energy concentration. In the formed beam, in addition to the energy concentrated in the beam direction, certain signals are radiated in other directions; these signals have unique amplitude and phase characteristics. Because beamforming parameters are known, complete characteristics of the beam in all directions can be obtained in one signal transmission process, and it's possible to observe the environment from all views. Under the combined effect of multipath transmission and additional sidelobe characteristics, the received signals contain extremely rich beams pointing to the external environment. Based on the known complete beam characteristics, information that was previously considered interference can be subtly extracted. For the beam sweeping policy, we

can quickly obtain all of the environment information while incurring low hardware costs and achieving accurate sensing [17].

We divide the space into several uniform voxels. By considering all beam characteristics, including the main lobe and parasitic side lobe with a certain width, signals reach all cubes regardless of the beam direction. So, there is a plurality of transmitter-cube-receiver paths in each signal transmission process. In addition, the way to obtain the amplitude and phase shift of the channel coefficient based on beams is unlike conventional channel modeling. The amplitude and phase shift are determined by the path itself and the beam characteristics in the path direction. This uses all of the beam characteristics, efficiently sensing the environment.

Because the scatterers in the environment are sparse, they can be accurately reconstructed based on multipath transmission and CS. As shown in Figure 9, in a 3D simulation scenario, standard OFDM signals with a carrier frequency of 28 GHz are sent through an 8 x 8 uniform planar array. The receive end processes the received signals using the above method so that scatterers, such as buildings outdoors, can be accurately sensed with a pixel precision of 1 meter. In Figure 11, horizontal and vertical coordinates represent the beam numbers of the receive end and the transmit end, respectively, where the beam spacing is 2°. The algorithm can also sense a scatterer at a corresponding angle location in the 2D scenario shown in Figure 10.

(a) Original scene diagram

(b) SNR = 10 dB

(c) SNR = 15 dB

(d) SNR = 20 dB
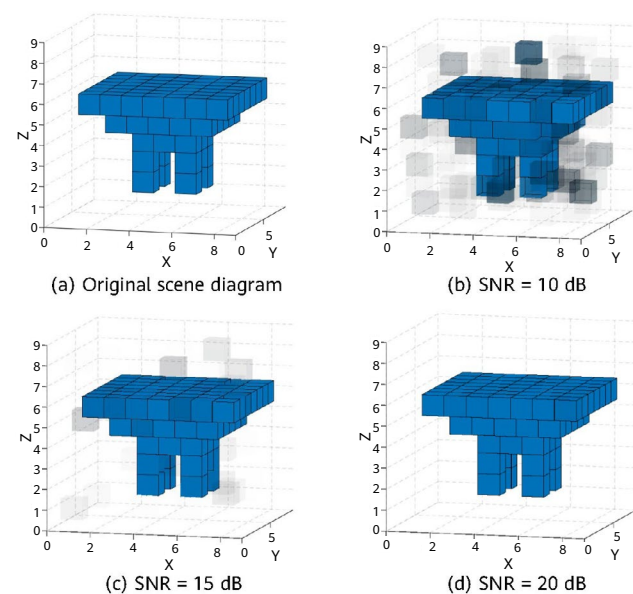
Figure 9 Sensing result in a 3D simulation scenario
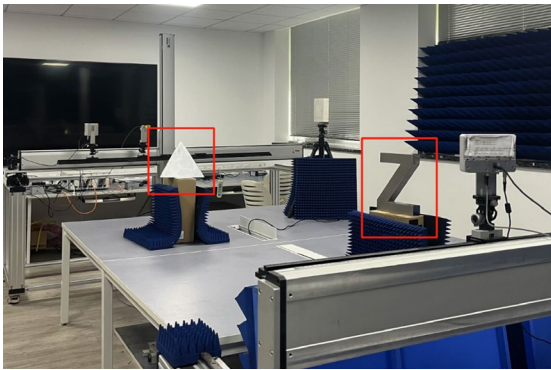
Figure 10 Test scenario settings
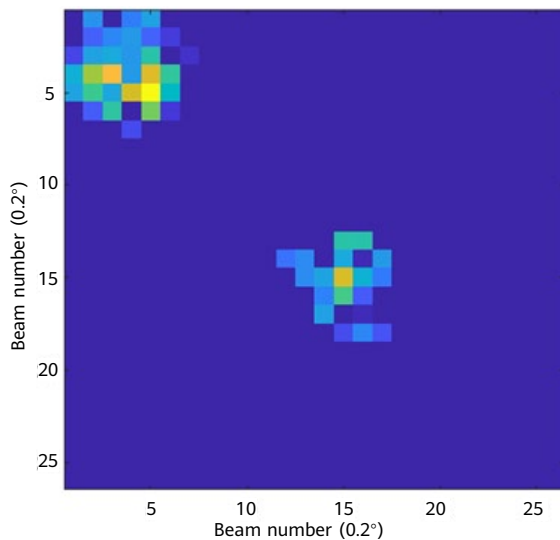
Figure 11 Sensing result in a 2D test scenario

5 Summary and Prospects

This paper first analyzes the wireless channel modeling method for environment sensing. To achieve more accurate environment sensing, we should thoroughly analyze the impact that objects in the environment have on channels. Then, we propose two design methods for the wireless environment sensing system, namely 3D environment reconstruction based on CS theory and moving/stationary object detection based on mmWave radar. The advantages and disadvantages of these two methods are also summarized and compared. Finally, we develop an ISAC verification platform based on the mmWave communications system and prove the algorithm's feasibility in real-world hardware systems.

References

[1] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.

[2] C. Jiao, Z. Zhang, C. Zhong, *et al.*, "Millimeter wave communication with active ambient perception," *IEEE Transactions on Wireless Communications*, vol. 18, no. 5, pp. 2751–2764, 2019.

[3] F. Liu, Y. Cui, C. Masouros, *et al.*, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, 2022.

[4] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 912–926, 2011.

[6] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4680–4688, 2011.

[7] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *2011 IEEE International Symposium on Information Theory Proceedings*, 2011, pp. 2168–2172.

[8] X. Tong, Z. Zhang, J. Wang, *et al.*, "Joint multi-user communication and sensing exploiting both signal and environment sparsity," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 6, pp. 1409–1422, 2021.

[9] Y. Zhang, Z. Zhang, X. Tong and C. Huang, "3D environment sensing with channel state information based on computational

imaging," *2022 IEEE International Conference on Communications Workshops (ICC Workshops)*, Seoul, Republic of Korea, 2022, pp. 842–847.

[10] X. Tong, Z. Zhang, Z. Yang and J. Che, "Environment sensing with beam sweeping and non-uniform pixelation in wireless communication systems," *2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, accepted.

[11] H. Xu, V. Kukshya and T. S. Rappaport, "Spatial and temporal characteristics of 60-GHz indoor channels," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 3, pp. 620–630, 2002.

[12] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, *et al.*, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 870–913, 2018.

[13] Z. Xing, Z. Zhang, X. Tong, *et al.*, "Physics-inspired target shape detection and reconstruction in mmWave communication systems," *GLOBECOM 2023 – 2023 IEEE Global Communications Conference,* accepted.

[14] X. Tong, Z. Zhang, Y. Zhang, *et al.*, "Environment sensing considering the occlusion effect: A multi-view approach," *IEEE Transactions on Signal Processing*, vol. 70, pp. 3598–3615, 2022.

[15] X. Tong, Z. Zhang and Z. Yang, "Multi-view sensing for wireless communications: Architectures, designs, and opportunities," in *IEEE Communications Magazine*, vol. 61, no. 5, pp. 40–46, May 2023.

[16] Y. Lu, Z. Zhang, X. Tong and Z. Yang, "Online tensor method for moving objective detection with FMCW radar," *2023 IEEE 97th Vehicular Technology Conference (VTC-Spring)*, accepted.

[17] K. Guo, Z. Zhang, X. Tong , *et al.*, "Sidelobe-enhanced beam sweeping for environment sensing in vehicular communication," submitted to *2023 IEEE 98th Vehicular Technology Conference (VTC-Fall)*.

Bistatic Channel Estimation in Integrated Sensing and Communication by Monostatic Sensing

Narcis Cardona [1], Wenfei Yang [2], Alejandro Castilla [1], Alejandro Lopez-Escudero [1], Jian Li [2]

[1] iTEAM Research Institute, Universitat Politècnica de València

[2] Wireless Technology Laboratory, Huawei Technologies Co., Ltd.

[1] {ncardona, a.castilla, alloes3}@iteam.upv.es

Abstract

The current growing interest in integrated sensing and communication (ISAC) for the next generation of radio access networks towards 6G is opening new challenges in channel estimation and modeling. This paper discusses the challenges in channel estimation and modeling for ISAC in the context of the next generation of radio access networks towards 6G. The authors propose a dual-channel model description for ISAC and a technique to relate the sensing and communication channels, which are described analytically and verified with measurements. The technique is based on exploiting the mutual interference between frequency modulated continuous wave (FMCW) sensing units to estimate the multipath communication channel between them.

Keywords

ISAC, JSAC, 6G, FMCW, channel modeling

1 Introduction

Anticipated to operate at extremely high frequencies, the next-generation communication networks (6G) are expected to merge communication and sensing [1]. To cope with the reduced channel coherence time and facilitate faster and more frequent updates to capture radio variations, future 6G systems will likely demand the creation of new dynamic radio channel models capable of capturing the channel impulse response with higher resolution [2]. Consequently, there is a need to develop innovative and efficient approaches to estimating radio channel properties. The integration of sensing and communication functions within a single wireless system combines the ability to sense and gather information about the environment with communication capabilities, allowing the system to adapt and optimize its performance based on the sensed data. Integrated sensing and communication (ISAC) has become a design paradigm for 6G, for which understanding and modeling the behavior of the sensing and communication channels simultaneously is crucial.

Sensing has been increasingly integrated with communication, particularly in automotive applications [3–8]. In [8], sensing information is used to predict the communication channel, defining perceptive mobile networks (PMN) with sensing-aided functions in radio access. Authors in [9] and [10] suggest modified OTFS/OFDM schemes for ISAC, while [11] explores parallel sensing and communication waveforms with shared bandwidth. In this paper, ISAC refers to communication-centric implementation using communication signals for sensing, while joint sensing and communication (JSAC) denotes systems using both sensing and communication together for mutual benefit.

Monostatic vs Bistatic sensing. The main ISAC applications in 6G mobile networks are vehicular communication (V2X), sensing as a service, remote sensing, and environmental monitoring, while those applying to short-distance wireless channels are, among others, in-cabin sensing, smart home, and human-computer interaction [12, 13]. This confers different ISAC channel scenarios, which can be summarized into three types, namely, monostatic, bistatic, and distributed channels, depending on the deployment of the transmitter (Tx) and receivers (Rx).

ISAC scenarios are monostatic when the sensing transmitter and receiver are co-located; the bistatic case refers to the sensing transmitter and receiver being separated, while distributed sensing accounts for multiple sensing paths available for a target, which means multiple sites can be employed to sense the same target. Figure 1 summarizes the three cases. Sensing channels can be modeled from multipath propagation when the transmitter and receiver are at the same position, with most of the sensing energy coming from line-of-sight (LOS) components. Reconstructing the environment is vital for monostatic ISAC, and understanding electromagnetic scattering from targets, including angle, roughness, and frequency, is crucial.

The estimation of the communication channel from the sensing one is investigated in [10], where the communication receivers are the targets of a monostatic OTFS-based ISAC system. Thus, the two channels in this case are highly correlated and the delay and Doppler parameters can be predicted from the sensing echoes. In [14], the authors use OTFS modulation for joint sensing parameter estimation and communication, exploiting the channel sparsity in the delay-Doppler domain for efficiently separate sensing parameter estimation and communication.

The implementation of monostatic ISAC sensing requires the full-duplex mode of the radio interface, which is not
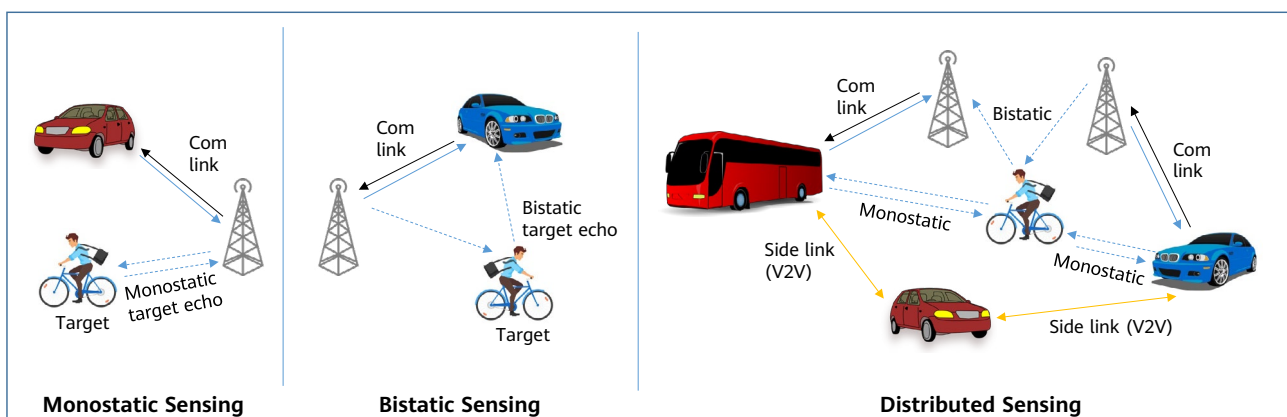
Figure 1 Monostatic, bistatic and distributed ISAC scenarios

implemented so far in current 3GPP standards. This is why some proposed systems for ISAC refer to a JSAC strategy, in which a sensing unit is used to improve communication links.

When two terminals are active, like in any communication link, the bistatic case for sensing applies, so the transmitter sends communication and sensing signals, and the receiver captures the echoes from targets and clutter, together with the multipath components (scattering) that compose the communication channel. Consequently, any model that describes simultaneously the sensing and the communication channels has to be geometrically accurate for simultaneous multiple links and reproduce moving targets along different tracks with spatial consistency, including phase continuity [15].

Some authors call a generalization of a distributed sensing system along a radio access network "MIMO radar." In [16] authors analyze the performance of an integrated radar-communication system based on MIMO. The MIMO-radar-based communication system shares the hardware and the frequency spectrum with the MIMO-radar and modulates the radar baseband waveform with the communication symbols.

ISAC channel models. Some channel models for ISAC have been proposed and published under different perspectives, corresponding to every application scenario in which the system is used, and can be categorized as deterministic, stochastic, and hybrid methods. While deterministic channel models are more suitable for deployment planning, statistical models are useful for developing and testing wireless systems. Hybrid channel models can achieve a good trade-off between accuracy and complexity by combining the benefits of different modeling methods.

Recent approaches so far in published papers about ISAC channel modeling propose geometry-based stochastic model (GSCM) extensions of the 3GPP channel model [17], like in [18–20]. The stochastic channel modeling methods dominate the evaluation of wireless communications in 5G due to their low computational complexity and easy standardization, which is sufficient to evaluate communication performance [17].

To apply GSCM models to ISAC, [21] uses stochastic and deterministic methods with spatial consistency based on real measurements to determine scatterer distribution. [20] proposes a shared cluster-based stochastic JSAC model using

28 GHz measurements in indoor scenarios. Both papers assume a good correlation between channels, more evident in monostatic scenarios. Ensuring effective stochastic modeling of ISAC channels needs new measurements and analysis to jointly model and maintain spatial consistency.

For accurate ISAC modeling, deterministic models with physical descriptors are recommended, contrary to the statistical GSCM approach. Until now, deterministic channel modeling has modeled scatterers for communications and considered the deflection from them in the form of specular or diffuse scattering. From a sensing perspective, any target in the scene can be modeled according to its radar cross section (RCS), but for ISAC, a more detailed parameter has to be defined, e.g., the bistatic target reflectivity (BTR) [15]. BTR includes at least the time-variant aspects of the target, its displacement, orientation, distance to the sensing units, and carrier frequency.

Combining the stochastic and deterministic approaches has the potential to benefit from the efficiency of stochastic and realism of deterministic approaches. In an ISAC context, the hybrid models would use a deterministic method to identify primary signal propagation paths and a stochastic method to generate additional objects and clusters. The level of determinism or stochasticity in the channel model can be adjusted based on the scenario and goal. For example, 3GPP TR38.901 already includes a hybrid channel model where ray tracing is used to find propagation paths, and stochastic clusters are generated afterward [17].

This paper demonstrates the possibilities of extracting the bistatic channel from monostatic measurements. In particular, this applies to two FMCW sensing units that interfere with each other, being the individual sensing channels obtained and separated from the mutual interference, which is used to estimate the bistatic (communication) channel between them. The concept is also proved with measurements at the E-band, as an extension of those presented in [22], and might be of application to channel sounding and dual ISAC channel modeling. The rest of the paper is organized as follows: Section 2 describes the nomenclature of the ISAC dual channel model employed, Section 3 provides an analytical description of the effects that are characterized, Section 4 shows some measurement results that verify the proposed concepts, and finally, Section 6 wraps up some conclusions and open challenges.

2 Dual-Channel Model for ISAC

In general terms, the ISAC channel models are a combination of two channels in one single geometry, i.e., a common framework that develops two approaches, either simulated with more or less precision, or stochastically created. The two channels in ISAC interact and interfere with each other, and, in some cases, such interaction may be used to understand better their common components. In [22], the ISAC channel is defined as a two-port system, with sensing and communication channels, and common interference between them, modeled using reflectivity and transmission functions. This schema is adaptable to any current model in the literature, either deterministically simulated by precise 3D ray-tracing, stochastically generated from a random distribution of scatterers, targets, and clutter, or a hybrid of the two.

The "dual-channel model" notation proposed in [22] to describe the ISAC channel between two terminals has the form of a two-port system (Figure 2), with R11 and R22 being the monostatic sensing channels from every terminal position, and T21 and T12 the communication channels in one and the other direction. In the case that sensing and communication channels are not sent within the same waveform, i.e., for generic ISAC systems, a more generalist approach accounts for separating the communication channel from point 1 to point 2 (T21), and the cross "interference" effect of a sensing signal of transmitter 1 being received by sensing receiver 2 (R21). So, 6 channel functions are defined as depicted in Figure 2, where R identifies the sensing channels, while T stands for transmission — or communication — channels [22].

This scheme corresponds to a scenario in which two terminals establish a dual communication link and are monostatically sensing their environment, and their respective sensing signals can be detected as mutual interference between them.

[22] analyzed the possibility of estimating T21 from R22 — since R22 will contain the R21 "interference" — and whether R21 is extractable in a fully operative sensing unit. From a
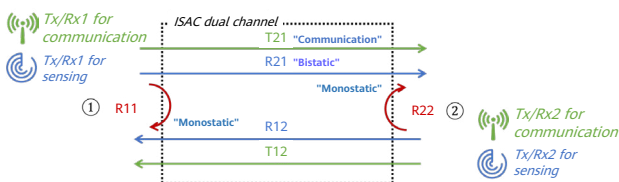
sensing-only perspective, if the former two assumptions can be verified, R21 can give the channel impulse response, so it could become an alternative low-cost channel sounder or an effective implementation of sensing and channel estimation in future systems. In summary, the T21 communication channel could be extracted from R22 monostatic sensing signals only.

3 Mutual Interference Between FMCW Sensing Units

The work described in this paper is based on using the mutual interference between sensing signals of any pair of terminals. In particular, it is considered the frequency modulated continuous wave (FMCW) sensing units, in which the sensing signal is a chirp (frequency ramp) with chirp time T_c, chirp repetition time T_p and bandwidth B, and that leads to frequency chirp slopes of $K_p = B/T_c$. This chirp signal is transmitted and combined with the received echo signal at the receiver. The result is an intermediate frequency (IF) signal that contains a carrier for every reflected signal at surrounding targets (Figure 3). From this set of IF carriers the clutter and moving targets of the environment are detected.

The effects of interference between FMCW sensing units are commonly found in the literature for purposes related to the improvement of sensing systems, i.e., for interference cancellation. Mitigation techniques to cancel interference are proposed in many literature papers [23–27]. In [25], the authors discuss interference between two FMCW sensing units using FMCW modulation. Some authors describe such effects of mutual interference as "ghost targets" [26, 27], since the consequence of the interference signal creating constant IFs for each ramp of the local oscillator signal, which lies in the passband of the anti-aliasing low-pass filter.

Figure 2 Definition of ISAC dual-channel model parameters
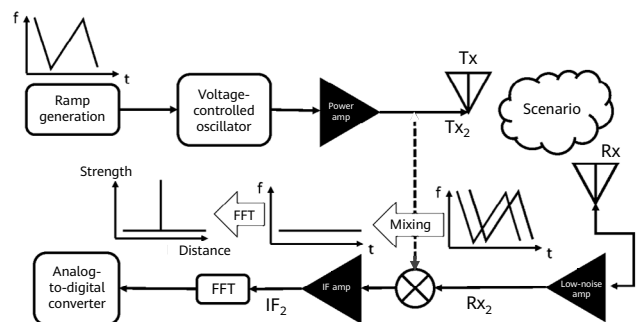
Figure 3 Block diagram of FMCW sensing units

Unlike the publications mentioned earlier, the focus of the work described in this paper is to extract the interference from received signals and examine the relationship between this interference and the communication channel.

The analytical expressions of the effect of mutual interference between two non-synchronized FMCW monostatic sensing units using the same carrier frequency are described in a previous publication [22]. The two sensing units generate their frequency ramps with not the same chirp slopes, called K_1 and K_2 respectively. Unless K_1 and K_2 were exactly the same value, which is not achieved in practice, the ramp slope difference creates a "ghost target" interference signal at each sensing receiver. This interference appears at the output of the sensing unit as a very high Doppler (speed) target, which in the time-delay domain corresponds to a linearly increasing tap delay, as depicted in Figure 6 of the next section.

The analog output signal (IF2 in Figure 3) of a monostatic FMCW sensing unit, when an interfering sensing unit is captured, is

$$IF_2 = Rx_2 \cdot Tx_2^* \tag{1}$$

being Rx_2 the combination of the sensing unit's clutter signals plus the effect of the interference. Now, the IF signal can be expressed as

$$IF_2 = Rx_2^{(2)} \cdot Tx_2^* + Rx_2^{(1)} \cdot Tx_2^* \tag{2}$$

where $Rx_2^{(2)}$ contains the terms of the received signal at Rx_2 caused by the clutter of sensing unit 2, so it is the regular output of the sensing unit or, as described in Section 2, the "sensing channel" R22. $Rx_2^{(1)}$ is the term added to the received signal at Rx_2 caused by the interference from sensing unit 1, so $Rx_2^{(1)}$ contains the R21 channel, as described in the model of Figure 2.

$Rx_2^{(1)}$ arrives at the victim sensing unit through the propagation channel from sensing unit 1 to sensing unit 2, assumed to include M scatterers or "paths" of a multipath model, with delays τ_k and path gains b_k (Figure 4). The initial phase of sensing unit 1 is non-coherent with sensing unit 2 and will be written as θ_1

$$Rx_2^{(1)} = \sum_k A_1 b_k e^{-j[(2\pi f_0 + \pi K_1(t-\tau_k))(t-\tau_k) + \theta_1]} \tag{3}$$

where, as described in detail in [22], A_1 is the amplitude of the transmitted signal at the output of sensing unit 1, and K_1 is the chirp slope of sensing unit 1's FMCW generator.

Then, the term in the received IF signal at sensing unit 2 (IF2) containing the effect of the interference from channel 1 is

$$Rx_2^{(1)} \cdot Tx_2^* = A_2 e^{j[(2\pi f_0 + \pi K_2 t)t + \theta_2]}$$
$$\times \sum_k A_1 b_k e^{-j[(2\pi f_0 + \pi K_1(t-\tau_k))(t-\tau_k) + \theta_1]} \tag{4}$$
$$= \sum_k A_2 A_1 b_k e^{j[2\pi K_1 \tau_k t + \pi(K_2 - K_1)t^2 + 2\pi f_0 \tau_k - \pi K_1 \tau_k^2 + \theta_2 - \theta_1]}$$

being A_2 the amplitude of the transmitted signal at the output of sensing unit 2, θ_2 the initial phase of sensing unit 2's output, and K_2 the chirp slope of sensing unit 2's FMCW generator.
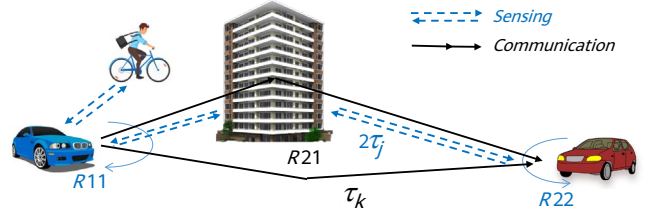
Figure 4 Schematic of sensing (R22, R11) and communication (R21) channel paths

From Equation (4), it can be noted that a term containing the multipath components appear in the output of the mixer (IF), identified by IF frequencies of $K_1 \tau_k$ and phases shifted by $2\pi f_0 \tau_k - \pi K_1 \tau_k^2$ plus the phase difference between the two sensing unit outputs $\Delta\theta = \theta_2 - \theta_1$. Then, the additional term $\pi(K_2 - K_1)t^2$ due to the difference between the chirp slopes K_2 and K_1, produces the instant IF frequency to be not constant, but to linearly increase with time, according to

$$IF_2 = K_1 \tau_k + (K_2 - K_1)t \tag{5}$$

So, as a consequence, the multipath components of the channel between sensing unit 1 and sensing unit 2 are received at the victim sensing unit (sensing unit 2) as any other clutter signal but with a linearly increasing frequency in the form of an IF chirp, being equivalent to a constant Doppler shift in the delay-Doppler domain. This effect is observed in the measurements as described in [22] for the mmWave band (24 GHz) and in this paper for the E-band.

4 Estimation of Communication Channels from Sensing Channels at the E-Band

The theoretical analysis of the previous section was confirmed by a set of measurements performed with commercial sensing units at the band of 24 GHz, and published in [22]. The limitation of those measurements was the bandwidth of 250 MHz, which reduces the delay resolution to 4 ns or the

Figure 5 Measurement scenario

range resolution to 60 cm for monostatic sensing. This paper presents measurements at sensing unit frequencies in the low E-band, from 60 to 64 GHz, so with a bandwidth of 4 GHz and a delay resolution of 0.25 ns, or a range resolution of 3.75 cm for monostatic sensing.

The measurements have been taken in several outdoor and indoor environments. In the following, results for an outdoor static scenario with antennas at the street level (Figure 5) are used to illustrate the concepts proposed in the paper. The two sensing units are active and sensing in a monostatic configuration, use the same bandwidth, and are not synchronized.

Sensing parameters correspond to those of Table 1, being the two sensing terminals (sensing units) of the same kind and configuration. As mentioned, the theoretical chirp slope of 19.457 MHz/μs will not be the same in practice, and the chirp slopes of one and the other sensing units differ by about 2 ppm, i.e., $K_2 - K_1 = 38.9$ Hz/μs.

The measurements performed are captured in frames of 128 chirps each, along a 4 GHz bandwidth, for which the receiver unit applies a 2048-points fast Fourier transform (FFT) to obtain the IF values from the mixer output, limiting the maximum range to 38.4 meters and the delay resolution to 0.25 ns.

Figure 6 shows one measured frame of 128 chirps at the IF output of the sensing unit. Every row in the graph corresponds to the sensing channel captured for one chirp. The sensing unit determines the distance to targets, which can be converted to delay, to have the equivalent impulse response, or the power delay profile (PDP), of the sensing channel at every chirp, i.e., every 205.32 μs in this case.

Then, the captured data can be used as the time-varying impulse response of the sensing channel, e.g., $h_{22}(t,\tau)$. Since h_{22} contains interference from sensing unit 1, which is observed in Figure 6 as a phantom target with linearly swept delay along time, some post-processing is applied to separate the sensing channel (R22) and such interference

Table 1 FMCW sensing units' main parameters

Parameter	Value
Start frequency (GHz)	60
Frequency slope (MHz/μs)	19.457
End frequency (GHz)	63.995
Number of ADC samples	2048
Sample rate (ksps)	10000
Chirp time (μs)	205.32
Rx gain (dB)	30
Number of chirp loops	128
Frame duration (ms)	27
Max. sensing range (m)	38.6
Tx antenna gain (dB)	7.7
Rx antenna gain (dB)	9.4
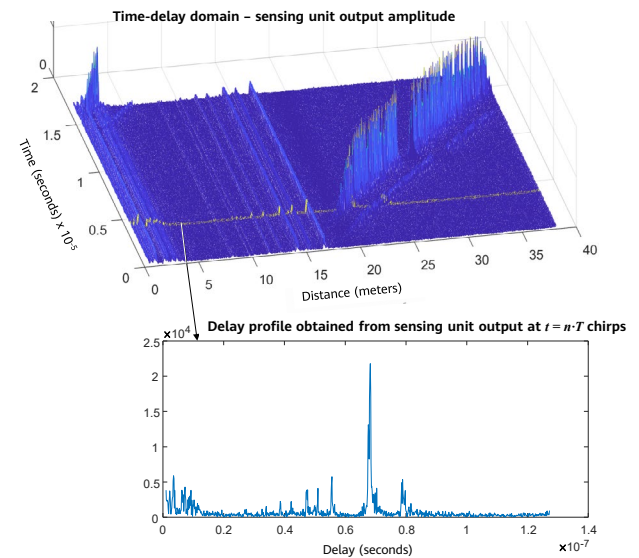
Figure 6 Measured time-delay domain for one frame (128 consecutive chirps), and delay profile at one of the chirps

(R21), i.e., the two ISAC channels are obtained from one single sensing measurement [22].

It is worth noting that this is not a bistatic sensing scheme, but the analysis of the sensing signal at only one of the terminals, from which the communication channel is estimated, although not synchronized. Time synchronization of the estimated communication channel is applied for the measured data by knowing the distance between the terminals. In a practical application of this technique to channel sounding or communication systems, some common synchronization strategies could be used, depending on the final application.

Figure 7 shows the result of estimating the communication channel R21 and the sensing channel R22. Note that the
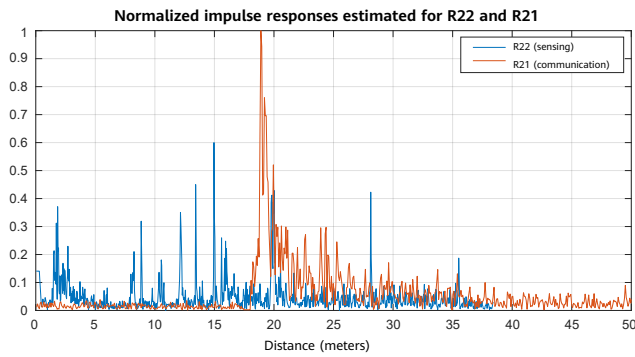
Figure 7 Extracted sensing (R22) and communication channels (R21) from the monostatic sensing unit's output

Table 2 Correlation between sensing and communication channels

Environment	Distance	Correlation Coefficient		
		R21 vs R12	R21 vs R22	R22 vs R11
Indoor	8 m	0.91	0.24	0.04
Outdoor	18 m	0.94	0.08	0.06
	23 m	0.87	0.11	0.05
	35 m	0.85	0.05	0.53

resolution and range for the communication channel are twice the values of those of sensing, since sensing signals correspond to backscattering from targets, while the communication channel occurs along propagation paths between the two terminals.

As shown in Table 2, in all the measurements performed in this work, and as for the examples of Figure 5 to Figure 7, the correlation between the PDPs of monostatic sensing (R22) and the communication channel (R21) is very low for static scenarios. Although the monostatic sensing and communication channels share the same propagation environment, delays and powers of the paths regarding the same scatterer can be much different between the monostatic sensing and communication channels due to different propagation distances and RCSs that vary with incident angles. However, in the angular domain, the paths regarding the same scatterer have the same angle of departure (AOD). Moreover, in the case of dynamic scenarios, i.e., when one of the terminals is moving, the relative speed of targets will be aligned with the relative speed of communication scatterers. Therefore, the correlation between sensing and communication channels is expected to be further studied in the angular and Doppler domains. On the other hand, the measured correlation between R21 and R12 is high, due to the reciprocity of the radio channel in these environments.

Table 2 summarizes the values obtained for several scenarios measured.

5 Conclusion

The paper proposes a technique for dual-channel model description for ISAC that is promising for channel sounding and channel estimation from only sensing measurements. The technique is based on exploiting the mutual interference between sensing units to estimate the multipath communication channel between them. The effect that causes the mutual interference is a consequence of the local oscillator uncertainty, which could be controlled if the differences between the K factors for every sensing unit are designed in future applications.

The analytical assumptions are confirmed by measurements performed with commercial FMCW sensing units at the E-band. The proposed technique has several advantages. First, it allows for the estimation of the communication channel from the sensing channel only, which reduces the need for additional hardware and complexity in the system. Second, it provides a generalist dual-channel model description for ISAC, which can be applied to different frequency bands and novel techniques for joining the sensing and the transmission of information over the same waveform. Third, it is based on exploiting the mutual interference between sensing units, which is a controllable phenomenon.

The correlation between sensing and communication channels in the scenarios of this work is confirmed to be low for all cases, as well as for the pairs of monostatic sensing channels measured. Correlation in bistatic and distributed scenarios is not properly studied and modelled so far.

References

[1] Carlos De-Lima *et al.*, "Convergent communication, sensing and localization in 6G systems: An overview of technologies, opportunities and challenges," IEEE Access, Jan 2021.

[2] P. Pärssinen, A., Alouini, M., Berg, M., Kuerner, T., Kyösti, P., Leinonen, M. E., Matinmikko-Blue, M., McCune, E., Pfeiffer, U., and Wambacq, "White paper on RF enabling 6G – Opportunities and challenges from technology to spectrum," 6G Research Visions, vol. 13, pp. 1–68, 2020.

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

[4] S. H. Dokhanchi, B. S. Mysore, K. V. Mishra, and B. Ottersten, "A mmWave automotive joint radar-communications system," IEEE Trans Aerosp Electron Syst, vol. 55, no. 3, pp. 1241–1260, Jun. 2019.

[5] P. Kumari, J. Choi, N. Gonzalez-Prelcic, and R. W. Heath, "IEEE 802.11ad-based radar: An approach to joint vehicular communication-radar system," IEEE Trans Veh Technol, vol. 67, no. 4, pp. 3012–3027, Apr. 2018.

[6] J. A. Zhang, X. Huang, Y. J. Guo, J. Yuan, and R. W. Heath, "Multibeam for joint communication and radar sensing using steerable analog antenna arrays," IEEE Trans Veh Technol, vol. 68, no. 1, pp. 671–685, Jan. 2019.

[7] H. Hua, J. Xu, and T. X. Han, "Transmit beamforming optimization for integrated sensing and communication," 2021 IEEE Global Communications Conference, GLOBECOM 2021 - Proceedings, 2021.

[8] Z. Ni, J. A. Zhang, X. Huang, K. Yang, and J. Yuan, "Uplink sensing in perceptive mobile networks with asynchronous transceivers," IEEE Transactions on Signal Processing, vol. 69, pp. 1287–1300, 2021.

[9] P. Raviteja, K. T. Phan, and Y. Hong, "Embedded pilot-aided channel estimation for OTFS in delay-Doppler channels," Aug. 2018, [Online]. Available: http://arxiv.org/abs/1808.08360

[10] S. Li *et al.*, "A novel ISAC transmission framework based on spatially-spread orthogonal time frequency space modulation," IEEE Journal on Selected Areas in Communications, vol. 40, no. 6, pp. 1854–1872, Jun. 2022.

[11] Dario Tagliaferri *et al.*, "Integrated sensing and communication system via dual-domain waveform superposition," preprint arXiv:2212.07950, Dec. 2022.

[12] D. K. Pin Tan *et al.*, "Integrated sensing and communication in 6G: Motivations, use cases, requirements, challenges, and future directions," 2021 1st IEEE International Online Symposium on Joint Communications and Sensing, JC and S 2021, Feb. 2021.

[13] C. Han, Y. Wu, Z. Chen, Y. Chen, and G. Wang, "THz ISAC: A physical-layer perspective of terahertz integrated sensing and communication," Sep. 2022. doi 10.48550/arxiv.2209.03145.

[14] L. Gaudio, M. Kobayashi, G. Caire, and G. Colavolpe, "On the effectiveness of OTFS for joint radar parameter estimation and communication," IEEE Trans Wirel Commun, vol. 19, no. 9, pp. 5951–5965, Sep. 2020.

[15] R. S. Thomä, C. Andrich, S. J. Myint, C. Schneider, and G. Sommerkorn, "Characterization of multi-link propagation and bistatic target reflectivity for distributed ISAC," Oct. 2022 arXiv preprint arXiv:2210.11840, 2022.

[16] R. Xu, L. Peng, W. Zhao, and Z. Mi, "Radar mutual information and communication channel capacity of integrated radar-communication system using MIMO," ICT Express, vol. 1, no. 3, pp. 102–105, Dec. 2015.

[17] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," TR38.901, 2018.

[18] Y. Liu, J. Zhang, Y. Zhang, Z. Yuan, and G. Liu, "A shared cluster-based stochastic channel model for joint communication and sensing systems," doi 10.1109/GCWKSHPS56602. 2022.10008672.

[19] A. Lopez-Reche *et al.,* "Considering correlation between sensed and communication channels in GBSM for 6G ISAC applications," pp. 1317–1322, Jan. 2023. doi 10.1109/GCWKSHPS56602.2022.10008630.

[20] L. Huang, Y. Zhang, Q. Li, and J. Song, "Phased array radar-based channel modeling and sparse channel estimation for an integrated radar and communication system," IEEE Access, vol. 5, pp. 15468–15477, Aug. 2017.

[21] Andrea Ramos *et al.*, "Evaluation methodology for 6G sensing-assisted communication system performance," IEEE Access, vol. 4, 2016.

[22] N. Cardona, J. S. Romero, W. Yang, and J. Li, "Integrating the sensing and radio communications channel modelling from radar mutual interference," ICASSP 2023 Rhodes Island, Greece, June 2023, pp. 1–5.

[23] S. Alland, W. Stark, M. Ali, and M. Hegde, "Interference in automotive radar systems: Characteristics, mitigation techniques, and current and future research," IEEE Signal Process Mag, vol. 36, no. 5, pp. 45–59, Sep. 2019.

[24] R. Amar, M. Alaee-Kerahroodi, and M. R. Bhavani Shankar, "FMCW-FMCW interference analysis in mm-wave radars; An indoor case study and validation by measurements," 21st International Radar Symposium (IRS), Jun. 2021, pp. 1–11.

[25] Y. Makino, T. Nozawa, M. Umehira, X. Wang, S. Takeda, and H. Kuroda, "Inter-radar interference analysis of FMCW radars with different chirp rates," The Journal of Engineering, vol. 2019, no. 19, pp. 5634–5638, Oct. 2019.

[26] M. Toth, P. Meissner, A. Melzer, and K. Witrisal, "Analytical investigation of non-coherent mutual FMCW radar interference," 2018 15th European Radar Conference, EuRAD 2018, pp. 71–74, Nov. 2018.

[27] S. Rao and A. V. Mani, "Interference characterization in FMCW radars," IEEE National Radar Conference - Proceedings, vol. 2020-September, Sep. 2020.

Interference Analysis for Collaborative Integrated Sensing and Communication Networks

Xiaozhou Zhang, Lincong Han, Zixiang Han, Liang Ma, Yajuan Wang, Mengting Lou, Jing Jin, Qixing Wang
China Mobile Research Institute
{zhangxiaozhou, hanlincong, hanzixiang, maliangyjy, wangyajuan, loumengting, jinjing, wangqixing}@chinamobile.com

Abstract

This paper analyzes the interference characteristics of collaborative integrated sensing and communication (ISAC) networks in the sixth-generation (6G) mobile communication system. It starts by presenting a collaborative ISAC cellular system model. Next, it analyzes and evaluates the inter-base station interference for typical high-frequency and low-frequency ISAC cellular systems. The paper then proposes principles for inter-base station interference management, based on the evaluation results, to ensure successful detection of sensing signals and improve sensing performance. Finally, the paper verifies the sensing accuracy and communication efficiency of the proposed interference management schemes through simulation.

Keywords

collaborative sensing, ISAC, interference characteristic, 6G

1 Introduction

In the sixth-generation (6G) mobile communication system, the core functions of communication networks will extend from wireless transmission to sensing driven by a large number of new services and scenarios. Integrated sensing and communication (ISAC), as one of the key enabling technologies of 6G, leverages integrated signal waveforms, multi-point collaboration, intelligent information interaction, and other cutting-edge technologies to make 6G endogenously capable of sensing. As shown in Figure 1, ISAC has two working modes: independent sensing and collaborative sensing.
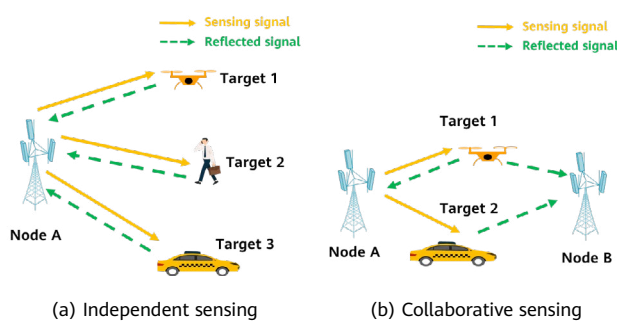
Figure 1 Working modes of ISAC networks

In the independent sensing mode, a single sensing node sends and receives sensing signals. These signals are processed to obtain the information of the sensing targets, including the angle, distance, and speed. At the signal processing level, [1] and [2] studied the impact of pilot patterns on the accuracy of measuring a target's location and speed. And at the system level, [3] studied the signal waveforms and beamforming algorithms of the sensing node for independent sensing to improve sensing accuracy. Although target sensing can be completed without assistant nodes, the sensing node must support high self-interference cancellation capabilities, and base station hardware must be adjusted significantly. Consequently, the independent sensing mode is difficult to deploy. To address this challenge, collaborative sensing is proposed. Collaborative sensing uses one node to send sensing signals and another node to receive signals reflected by the object. Nodes do not need to eliminate the self-interference as they do in the independent sensing mode. Collaborative sensing can be implemented through software upgrade without adjusting any base station hardware, significantly reducing the deployment cost. [4] proposed a dual-base-station collaborative sensing method to obtain the object location by estimating the angle and delay in two stages. And [5] analyzed the performance of the dual-base-station

collaborative sensing system. In collaborative sensing, the clocks of different nodes must be fully synchronized to accurately estimate the path propagation delay.

Collaborative sensing technology can be introduced to large-scale cellular networks to implement multi-node joint reception based on widely distributed base stations, thereby improving sensing accuracy and coverage. The formula of determining the accuracy of object locating in multi-node collaborative sensing was derived in [6]. [7] proposed the beamforming design for multi-node collaborative sensing on cellular networks. However, the inter-base station interference, and the interference between base stations and users on cellular networks, makes it very difficult to extract sensing signals and calculate sensing parameters. [8] provided the general principle of inter-base station interference management. To the best of the author's knowledge, little research has been conducted on the quantitative analysis of interference between base stations or between base stations and users.

This paper focuses on analyzing the interference characteristics of collaborative ISAC networks in 6G. It starts by presenting a collaborative ISAC network model in order to quantitatively analyze the interference characteristics between base stations in the systems. Then, it analyzes the characteristics of the inter-base station interference. Taking account of the analysis results, the paper proposes two feasible interference management schemes aimed at ensuring successful detection of echo signals and improving sensing accuracy. Finally, the sensing accuracy and communication performance of the proposed schemes are verified through simulation in different scenarios.

The contribution of this work is concluded as follows:

1. Focusing on collaborative sensing and the communications system, this work analyzes the characteristics of inter-base station interference over different frequency bands and presents the interference management strategies based on the analysis results.

2. A system-level simulation platform is deployed to evaluate sensing accuracy and the loss of communication efficiency under different inter-base station interference management strategies.

2 System Model

As shown in Figure 2, the proposed collaborative ISAC network includes a sensing transmitting point (STP) with

N_T antennas, a sensing receiving point (SRP) with N_T antennas, and M downlink communication transmitting points (CTPs) with N_R antennas. Each CTP serves N pieces of single-antenna user equipment (UE). In the sensing process, the STP transmits a sensing reference signal $\mathbf{x}_s \in \mathbb{C}^{N_T \times 1}$ with the transmit power of P_s. The signal is then reflected by the object and received by the SRP as an echo signal. The object can be a connected user or a non-connected object. This paper focuses on device-free sensing where the object cannot receive, transmit, or process sensing reference signals. In the communication process, each downlink CTP_m $(m = 1, \ldots, M)$ sends a communication signal $\mathbf{x}_{c,m} \in \mathbb{C}^{N_T \times 1}$ with the transmit power of $P_{c,m}$ to a communication user in the serving cell to transmit communication data. Each downlink CTP serves one communication user UE_m each time.
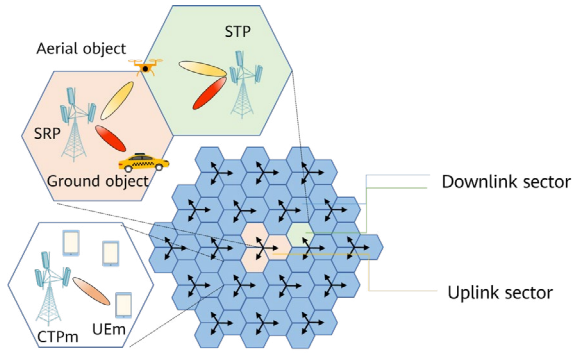
Figure 2 System model

For the sensing system, the signal $\mathbf{y}_s$ received by the SRP is presented as follows:

$$\mathbf{y}_s = \sqrt{P_s/\rho_s}\mathbf{H}_s\mathbf{x}_s + \sqrt{P_s/\rho_{ss}}\mathbf{H}_{ss}\mathbf{x}_s$$
$$+ \sum_{m=1}^{M} \sqrt{\alpha_m P_{c,m}/\rho_{cs,m}}\mathbf{H}_{cs,m}\mathbf{x}_{c,m} + \mathbf{n}, \quad (1)$$

where the first term on the right of Equation (1) is the sensing signal received by the SRP, the second term on the right of Equation (1) is the interference of the sensing signal generated by the STP on the SRP, the third term on the right of Equation (1) is the communication interference generated by M downlink CTPs on the SRP, and the fourth term on the right of Equation (1) is the noise signal. $\mathbf{H}_s, \mathbf{H}_{ss},$ and $\mathbf{H}_{cs,m} \in \mathbb{C}^{N_R \times N_T}$ are channel state matrixes for the link from the STP to the SRP reflected by the object, the link from the STP to the SRP, and the link from the CTP_m to the SRP, respectively. ρ_s is the free space path loss (FSPL) of the sensing path (unit: dB) and is presented as follows [9]:

$$PL_s(d_{s,1}, d_{s,2}) = PL(d_{s,1}) + PL(d_{s,2})$$
$$+ 10\log_{10}\left(\frac{\lambda^2}{4\pi}\right) - 10\log_{10}(\sigma_{RCS}), \quad (2)$$

where $PL(\cdot)$ is the path loss of a conventional communication link. $d_{s,1}$ and $d_{s,2}$ are the distances from the STP to the object and from the object to the SRP, respectively. λ is the wavelength, and σ_{RCS} is the radar cross section (RCS) of the object. $\rho_{cs,m} = 10^{0.1PL(d_{cs,m})}$ and $\rho_{ss} = 10^{0.1PL(d_{ss})}$ are the FSPLs from the CTP_m to the SRP and from the STP to the SR0050, respectively. $d_{cs,m}$ and d_{ss} are the distances from the CTP_m to the SRP and from the STP to the SRP, respectively. The formula of the path loss $(PL(\cdot))$ is provided in 3GPP TS 38.901 [10]. Performing interference coordination on CTPs can reduce their interference to the received sensing signals. $\alpha_m \in [0,1]$ is the interference coordination factor at CTP_m.

According to Equation (1), the signal to interference plus noise ratio (SINR) of the sensing signal received by the SRP is calculated as follows:

$$\mathrm{SINR}_s = \frac{P_s\|\mathbf{H}_s\|_F^2}{\rho_s(I_c + I_{ss} + N)}, \quad (3)$$

where N is the environment noise power, I_c and I_{ss} are the interference signal power of the CTPs and the STP on the SRP, respectively.

$$I_c = \sum_{m=1}^{M} \alpha_m P_{c,m}\|\mathbf{H}_{cs,m}\|_2^2/\rho_{cs,m},$$
$$I_{ss} = P_s\|\mathbf{H}_{ss}\|_F^2/\rho_{ss}.$$

For the communication system, the signal $y_{c,m}$ received by UE_m is calculated as follows:

$$y_{c,m} = \sqrt{P_{c,m}/\rho_{c,m}}\mathbf{h}_{c,m}^T\mathbf{x}_{c,m} + \sqrt{\beta_s P_s/\rho_{sc,m}}\mathbf{h}_{sc,m}^T\mathbf{x}_s$$
$$+ \sum_{q=1,q\neq m}^{M} \sqrt{\beta_{c,q,m}P_{c,m}/\rho_{cc,q,m}}\mathbf{h}_{cc,q,m}^T\mathbf{x}_{c,q} + n, \quad (4)$$

where, the first term on the right of Equation (4) is the data received by UE_m from CTP_m, the second term on the right of Equation (4) is the interference received by UE_m from the STP, the third term on the right of Equation (4) is the interference caused by other CTPs to UE_m, and the fourth term on the right of Equation (4) is the noise signal. $\mathbf{h}_{c,m}, \mathbf{h}_{cc,q,m},$ and $\mathbf{h}_{sc,m} \in \mathbb{C}^{N_T \times 1}$ are the channel state vectors of the link from CTP_m to UE_m, from CTP_q to UE_m, and from the STP to UE_m, respectively. $\rho_{c,m}, \rho_{cc,q,m},$ and $\rho_{sc,m}$ are the FSPLs from CTP_m to UE_m, from CTP_q to UE_m, and from the STP to UE_m, respectively. Similar to the sensing signal, $\beta_{c,q,m}$ and $\beta_s \in [0,1]$ are the interference coordination factors at CTP_q and the STP, respectively.

According to Equation (4), the SINR of the communication signal received by UE_m is calculated as follows:

$$\mathrm{SINR}_{c,m} = \frac{P_{c,m}\|\mathbf{h}_{c,m}\|_2^2}{\rho_{c,m}(I_{c,m} + I_{s,m} + N)}, \quad (5)$$

where $I_{c,m}$ and $I_{s,m}$ are the power of the interference signals of other CTPs and the STP on UE_m, respectively:

$$I_{c,m} = \sum_{q=1,q\neq m}^{M} \beta_{c,q,m} P_{c,q} \left\| \mathbf{h}_{cc,q,m} \right\|_2^2 / \rho_{cc,q,m},$$

$$I_{s,m} = \beta_s P_s \left\| \mathbf{h}_{sc,m} \right\|_2^2 / \rho_{sc,m}.$$

3 Interference Characteristic Analysis

In a collaborative ISAC network, the downlink communication interference received by a UE is mainly caused by the links between the downlink CTPs and UE, similar to that in a conventional communication system. Because the SRP works in the uplink mode, the interference to the received sensing signal is caused by the links between the STP and SRP, and between the downlink CTPs and the SRP. Such sensing interference is much stronger than communication interference.

This section analyzes the characteristics of sensing signal interference between base stations in typical high- and low-frequency collaborative ISAC networks. The considered ISAC system has 19 cells, 57 sectors, and channels that meet urban macrocell (UMa) requirements. Table 1 lists the other parameters involved in these scenarios. The sensing transmitter and receiver are located in sectors with optimal channel conditions at different sites. Two types of interference characterization are considered: small-sized aerial object sensing and medium-sized ground object sensing, where the RCS of the object is 0.01 m^2 for the first type and 10 m^2 for the second. The links between the object and STP and between the object and SRP are line-of-sight (LOS) paths. For the high-frequency system, L beams are preconfigured, among which the beam corresponding to the maximum reference signal received power (RSRP) measured by the receiving node is matched to UEs and the object through beam scanning. The CTPs randomly select a UE they serves to transmit communication signals.

Figure 3 shows the cumulative distribution function (CDF) of the inter-base station interference in typical high- and low-

frequency systems under different inter-site distances (ISDs). In this figure, the three blue curves on the right indicate the sensing interference intensity when the interference from two neighboring sectors at the same site occurs. The three red curves on the left indicate the sensing interference intensity when the neighboring sectors are disabled. The inter-base station interference in aerial sensing scenarios is lower than that in ground sensing scenarios. This is because the direction of the beam for receiving sensing signals and the interference beam in the aerial sensing scenario are different. Furthermore, in both high- and low-frequency systems, the interference from the two neighboring sectors at the same site is much higher than that from other sectors. As a result, when interference from two neighboring sectors at the same site occurs, the intensity of sensing interference is almost impervious to ISD changes.

To ensure that sensing signals can be successfully detected, the SINR of received sensing signals needs to be greater than a specific threshold, which depends on hardware parameters. For example, a 10-bit analog-to-digital converter (ADC) at the sensing receiver has an ADC dynamic range of about 60 dB, meaning that SINR of sensing signals must be greater than 60 dB. In other words, the upper

| (a) High-frequency, aerial sensing | (b) High-frequency, ground sensing |
| (c) Low-frequency, aerial sensing | (d) Low-frequency, ground sensing |
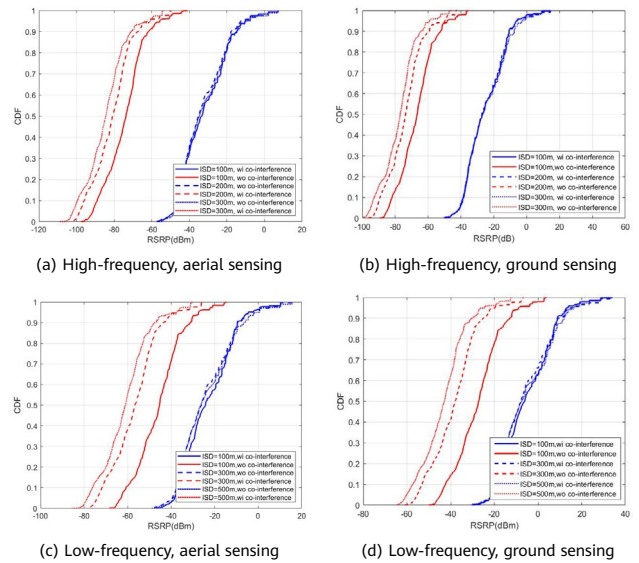
Figure 3 CDF of the sensing interference in collaborative ISAC networks

Table 1 Parameter settings for typical high- and low-frequency collaborative ISAC networks

System Settings	Carrier Frequency	Bandwidth	Transmit Power	Antenna Element	Gain per Element
High-frequency	26 GHz	400 MHz	44 dBm	16 x 16 x 2	8 dBi
Low-frequency	2.6 GHz	100 MHz	52 dBm	12 x 8 x 2	8 dBi

interference threshold is calculated as follows:

$$\text{Interference upper bound} =$$
$$\text{Echo signal strength} + \text{ADC dynamic range.} \quad (6)$$

When there is strong collaborative sensing interference, interference management must be performed to ensure that sensing signals can be detected. Assume that an object is located at the cell edge and that the ADC dynamic range of the sensing receiver is 60 dB. Tables 2 and 3 describe the strength of the echo signal, upper interference bound, and intensity of sensing interference at different ISDs in typical high- and low-frequency scenarios, respectively. Equations (3) and (6) are used to obtain the echo signal strength and the upper interference bound, respectively. The interference intensity is obtained from the system-level simulation, as shown in Figure 3. For small-sized aerial object sensing, in both typical high- and low-frequency systems, when the ISD is 300 m or more, two neighboring sectors working at the same site as the sensing receiver will cause the intensity of sensing interference to exceed the interference upper bound (≥ 8 dB). As a result, detection of sensing signals fails. To ensure that such detection is successful, interference management must be performed for two neighboring sectors in long-distance collaborative sensing. Such management is not necessary in medium-sized ground object sensing. Note that interference management on the STP, SRP, and CTPs can be implemented to further improve the SINR of sensing signals and sensing accuracy.

To ensure that sensing signals can be successfully detected and meet sensing performance requirements, interference

Table 2 Echo signal and interference intensity in a typical high-frequency collaborative ISAC network

High-Frequency Aerial Sensing, RCS = 0.01 m²				
ISD (m)	Echo Signal (dBm)	Upper Interference Limit (dBm)	Intensity of Sensing Interference (dBm)	
			Neighboring Sector Enabled	Neighboring Sector Disabled
100	−57.44	2.56	−14.72	−61.60
200	−70.68	−10.68	−17.64	−66.99
300	−78.43	−18.43	−14.10	70.84
High-Frequency Ground Sensing, RCS = 10 m²				
ISD (m)	Echo Signal (dBm)	Upper Interference Limit (dBm)	Intensity of Sensing Interference (dBm)	
			Neighboring Sector Enabled	Neighboring Sector Disabled
100	−27.44	32.56	−10.70	−52.49
200	−40.68	19.32	−10.72	−60.00
300	−48.43	11.57	−7.58	−64.50

Table 3 Echo signal and interference intensity in a typical low-frequency collaborative sensing system

Low-Frequency Aerial Sensing, RCS = 0.01 m²				
ISD (m)	Echo Signal (dBm)	Upper Interference Limit (dBm)	Intensity of Sensing Interference (dBm)	
			Neighboring Sector Enabled	Neighboring Sector Disabled
100	−37.96	22.04	−9.61	−31.22
200	−58.95	1.05	−9.62	−43.29
300	−68.71	−8.71	−6.45	−47.79
Low-Frequency Ground Sensing, RCS = 10 m²				
ISD (m)	Echo Signal (dBm)	Upper Interference Limit (dBm)	Intensity of Sensing Interference (dBm)	
			Neighboring Sector Enabled	Neighboring Sector Disabled
100	−7.96	52.04	8.88	−12.93
200	−28.95	31.05	9.53	−25.21
300	−38.71	21.29	11.98	−29.49

coordination can be performed from the perspectives of the time domain, frequency domain, space domain, or power domain. In the time domain, we can restrict downlink scheduling of the interfering communication sectors, adjust the frame structure, or use a frame structure with a large guard period (GP) for both the interfering communication sectors and the coordinated receiving sector. In the frequency domain, we can configure orthogonal uplink and downlink frequency domain resources for the interfering communication sectors and the coordinated receiving sector, or restrict scheduling on the uplink sub-band for the coordinated receiving sector. In the space domain, we can limit the transmitting beam direction of the interfering communication sectors and the receiving beam direction of the sensing signal, adjust the downtilt, or use advanced interference cancellation algorithms. In the power domain, we can decrease the transmit power of the interfering communication sectors or increase the transmit power of the coordinated sector.

4 Simulation Result

This section provides the system-level evaluation result of how different interference management schemes impact on the performance of collaborative ISAC in cellular networks. The simulation is performed in a time division duplex mobile communication system with macro cells. The network topology has many hexagonal cellular structures. At the center of each cellular structure, one site is deployed. Each site includes three base stations, each of which covers a different sector in the cellular structure. At each site, the three base stations have the same uplink and downlink timeslots. In this simulation, one sensing transmitter, one sensing receiver, and 53 communication transmitters are configured. The sensing receiving sector is located at the center of the network. The sensing transmitter and receiver are deployed at two different sites that are adjacent and have optimal channel conditions. In the simulated collaborative ISAC networks, one symbol among 14 orthogonal frequency division multiplexing (OFDM) symbols in each timeslot is used for sensing, and the other 13 are used for communication.

The impacts of the proposed interference management schemes on the ISAC networks are evaluated on 26 GHz and 2.6 GHz. Table 1 lists the typical base station configurations, and Table 4 lists other system parameters. Different interference coordination factors are configured for the proposed interference management schemes. Interference

coordination factor 1 indicates no interference coordination, and 0 indicates perfect interference coordination. The same interference management scheme is applied to all interference coordination sectors. In this simulation, sensing accuracy is evaluated by the root mean squared error (RMSE) of distance estimation, which is calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \left(\hat{d}_n - d^\star \right)^2},$$

where N is the number of experiments. $\hat{d}_n$ is the estimated values of the sensing parameters in the nth experiment, calculated by estimation of signal parameters via rotational invariance techniques (ESPRIT) [3]. $d^\star$ is the accurate values of the sensing parameters. 1000 experiments are conducted in each round of simulation.

Table 4 Other system parameters

System Parameter	Value
Number of UEs in each sector	10
Base station downtilt	11°
Subcarrier spacing	60 kHz
RCS	0.01 m², 10 m²
Modulation scheme	QPSK
Sensing reference signal	DMRS
Number of experiments	1000

Figure 4 shows the relationship between the distance estimation accuracy and interference coordination factor in small-sized aerial object sensing (RCS: 0.01 m²). A smaller interference coordination factor leads to less system interference. In turn, this results in a much lower RMSE of distance estimation, and produces higher distance estimation accuracy in both high- and low-frequency systems. Furthermore, a smaller ISD leads to higher sensing accuracy when the same interference coordination factor is used in both high- and low-frequency systems. This is because although stronger interference occurs with a smaller ISD, the sensing receiving power is much higher, resulting in high SINR of the received sensing signal and therefore better collaborative sensing performance. In conclusion, densely deployed base stations are more advantageous for collaborative ISAC networks. Compared with low-frequency systems, the RMSE of distance converges to a higher value in the high-frequency system. This is because the collaborative ISAC network is noise-limited when the interference coordination factor is small enough.

In a high-frequency system, the sensing signal power loss and receiving noises are larger, leading to smaller SINR of the received sensing signal.

(a) Typical high-frequency system (b) Typical low-frequency system
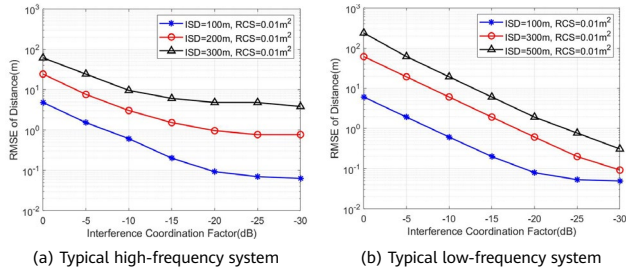
Figure 4 RMSE of distance over different interference coordination factors in small-sized aerial object sensing

Figure 5 shows the relationship between the distance estimation accuracy and interference coordination factor in medium-sized ground object sensing (RCS: 10 m^2) on high- and low-frequency networks. Similar to that in small-sized aerial object sensing, a smaller interference coordination factor leads to a significant decrease in the RMSE of distance estimation in both high- and low-frequency systems, and a smaller ISD leads to higher accuracy of distance estimation. Compared with Figure 4 and Figure 5, unlike small-sized aerial object sensing, the RMSE of distance estimation in medium-sized ground object sensing gradually converges as the interference coordination factor decreases. Furthermore, the sensing performance is significantly improved as the collaborative receiving power increases. In conclusion, the interference coordination factor needs to be selected based on a combination of factors, such as frequency and ISD, etc.

(a) Typical high-frequency system (b) Typical low-frequency system
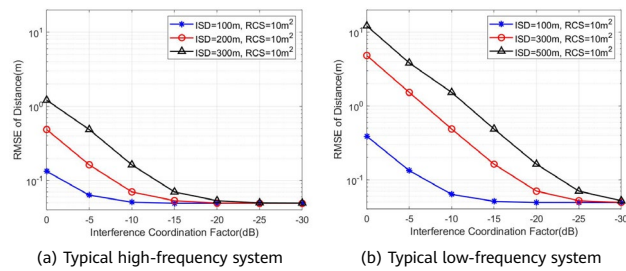
Figure 5 RMSE of distance over different interference coordination factors in medium-sized ground object sensing

Because long-distance base stations generate only weak interference, interference coordination is not mandatory for all interfering base stations. Figure 6 shows the relationship between the distance estimation accuracy and the number of interference management cells in small- and medium-sized object sensing on both high- and low-frequency networks. In a high-frequency system, when the number of interference management cells is sufficiently large, the distance estimation accuracy remains stable. This means that interference management for part of the

cells can achieve the expected sensing performance while also reducing the management overheads. Unlike high-frequency systems, the distance estimation accuracy in low-frequency systems is significantly improved when the number of interference management cells increases. Because the attenuation of the interference signal is weaker in a low-frequency system, significant improvements in the sensing performance can still be achieved even when performing the interference management on short-distance base stations.

(a) Typical high-frequency system (b) Typical low-frequency system

(c) Typical high-frequency system (d) Typical low-frequency system
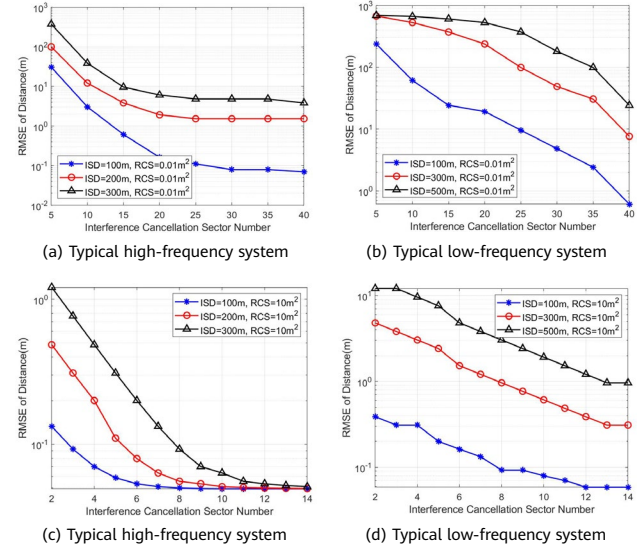
Figure 6 RMSE of distance versus the number of interference coordination sectors in different scenarios
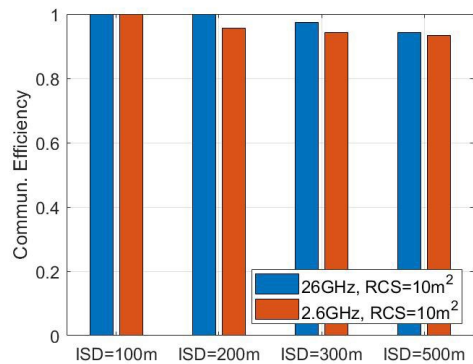
Figure 7 Communication efficiency at 1 m RMSE of distance

Figure 7 shows the communication efficiency loss caused by interference management when the RCS of the object is 10 m^2 and the RMSE of distance is 1 m on high- and low-frequency networks. Generally, communication efficiency decreases as the ISD increases in both high- and low-frequency systems. In particular, after applying interference management, the communication efficiency can still be higher than 97% in high-frequency systems. In low-frequency systems, when the ISD increases from 100 m to 500 m,

communication efficiency decreases from 100% to 93%. Generally, communication efficiency at a distance estimation accuracy of 1 m does not decrease significantly in either high-frequency or low-frequency collaborative ISAC networks.

5 Conclusion and Outlook

This paper analyzes and evaluates the intensity of inter-base station interference in collaborative ISAC networks. To ensure successful detection of sensing signals, interference generated by sensing receivers at the same site needs to be managed in long-distance small-sized aerial object sensing. Further increases in the number of base stations for interference management and the degree of interference management based on the high- and low-frequency interference characteristics can also improve the sensing performance. A system-level simulation verifies the sensing accuracy and communication efficiency of different interference management schemes.

Collaborative ISAC networks can solve more complex ISAC problems from a global perspective, facilitating the prototype research and the industry implementation of ISAC. In addition to interference management, coordinated cluster design, coordinated signal processing, synchronization error elimination, and coverage enhancement are also key topics that need to be researched in the field of collaborative ISAC. We believe that more research breakthroughs will be made in collaborative ISAC technology in the near future to transform the technology from a theoretical blueprint to real-world applications.

References

[1] L. Pucci, E. Paolini, and A. Giorgetti, "System-level analysis of joint sensing and communication based on 5G new radio," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2043–2055, Jul. 2022.

[2] L. Ma, C. Pan, Q. Wang, M. Lou, Y. Wang, and T. Jiang, "A downlink pilot based signal processing method for integrated sensing and communication towards 6G," in *Proc. IEEE VTC Spring*, 2022, pp. 1–5.

[3] X. Zhang, J. Jin, L. Ma, Y. Wang, M. Lou, T. Jiang, Q. Wang, G. Liu, and J. Wang, "Pilot-based delay and doppler estimation in 6G integrated communication and sensing networks," in *Proc. 14th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, 2022, pp. 478–482.

[4] L. Leyva, D. Castanheira, A. Silva, and A. Gameiro, "Two-stage estimation algorithm based on interleaved OFDM for a cooperative bistatic ISAC scenario," in *Proc. IEEE VTC spring*, 2022, pp. 1–6.

[5] L. Pucci, E. Matricardi, E. Paolini, W. Xu, and A. Giorgetti, "Performance analysis of a bistatic joint sensing and communication system," in *Proc. IEEE ICC Workshops*, 2022, pp. 73–78.

[6] A. Sakhnini, M. Guenach, A. Bourdoux, and S. Pollin, "A Cramér-Rao lower bound for analyzing the localization performance of a multistatic joint radar-communication system," in *Proc. IEEE JC&S*, 2021, pp. 1–5.

[7] R. Li, Z. Xiao, and Y. Zeng, "Beamforming towards seamless sensing coverage for cellular integrated sensing and communication," in *Proc. IEEE ICC Workshops*, 2022, pp. 492–497.

[8] L. Leyva, D. Castanheira, A. Silva, A. Gameiro, and L. Hanzo, "Cooperative multiterminal radar and communication: A new paradigm for 6G mobile networks," *IEEE Veh. Technol. Mag.*, vol. 16, no. 4, pp. 38–47, 2021.

[9] IMT-2020(5G), 5G-Adcanced, 2023.

[10] Study on channel model for frequencies from 0.5 to 100 GHz, document 3GPP TR 38.901, Mar. 2022.

Integrated Imaging and Communication: Advances, Challenges, and Prospects

Yixuan Huang [1], Jie Yang [2, 3], Shi Jin [1, 3]
[1] National Mobile Communications Research Laboratory, Southeast University
[2] Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education, Southeast University
[3] Frontiers Science Center for Mobile Information Communication and Security, Southeast University
{huangyx, yangjie, jinshi}@seu.edu.cn

Abstract

With the extensive development of integrated sensing and communication, high-precision sensing is anticipated for future wireless communication systems, where imaging is one of the important sensing forms. This study reviews the recent progress in integrated imaging and communication and outlines two technical roadmaps: beamforming-based range estimation and inverse scattering problem-solving. Specifically, two main approaches, Fourier transform and compressed sensing, for solving inverse scattering problems are discussed. Reconfigurable intelligent surface and artificial intelligence techniques are presented as two key enablers for imaging. The paper concludes by discussing the challenges and prospects of integrated imaging and communication.

Keywords

integrated imaging and communication, beamforming, inverse scattering, reconfigurable intelligent surface, artificial intelligence

1 Introduction

Integrated sensing and communication (ISAC) is a promising technique and application scenario in future wireless communication systems [1]. The sensing functions are realized using wireless communication devices, and the dual functions of communication and sensing are expected to be enhanced mutually. While previous research [2, 3] has focused on estimating simple parameters of the targets, such as locations and velocities, the study on more precise sensing abilities is growing. This has led to the development of ISAC systems that can capture finer environment features, such as radio maps and target images [4, 5]. Imaging functions are particularly valuable because they can provide rich target information and support applications such as wireless monitoring and augmented reality (AR). Therefore, this study focuses on the research and development of integrated imaging and communication systems.

Electromagnetic (EM) imaging, specifically synthetic aperture imaging, is a well-established field with numerous advantages, including strong penetration, good privacy, and independence of lighting conditions [6–9]. The primary goal of EM imaging is to obtain the geometric shape and scattering characteristics (such as the scattering coefficient and dielectric constant) of the targets within a specific region of interest (ROI). To achieve this, the ROI is discretized into numerous voxels, whose scattering coefficients or dielectric constants form the ROI image [10]. The scattering coefficient of a voxel is defined as the square root of the scattering energy when illuminated by a unit energy field, which is equivalent to the square root of the radar cross section [5, 11]. In integrated imaging and communication, as well as traditional synthetic aperture imaging fields, imaging means recovering the scattering coefficients of these voxels. However, an ROI can consist of tens of thousands of voxels, resulting in a significant amount of data that should be recovered, making the task of imaging more challenging than estimating locations or velocities.

The current study on integrated imaging and communication primarily involves two types of methods: (1) imaging based on beamforming and ranging and (2) imaging based on inverse scattering problem-solving. The former is a monostatic imaging technique that relies on the beamforming ability of large antenna arrays. By transmitting directional beams and receiving echoes, the range information about the voxels in the interested directions can be derived. The target shapes and scattering coefficients can be recovered by scanning the beams [12, 13]. In contrast, imaging based on inverse scattering problem-solving involves both monostatic and bistatic scenarios where the received echoes contain the scattered waves from the entire ROI [6]. Considering the large energy loss of multiple bounces, most studies on integrated imaging and communication utilize the Born approximation [14], assuming small scattering coefficients and ignoring multiple bounces among voxels. Consequently, the imaging problem is modeled as a linear inverse scattering problem, which can be solved by Fourier transform (FT) [6–8] and compressed sensing (CS) [9, 15], algorithms[16].

Imaging algorithms are developed with the advance of communication and signal processing techniques. The emergence of metamaterials, specifically the reconfigurable intelligent surface (RIS), has spurred numerous innovative designs of communication systems and imaging algorithms [17–20]. In addition, the powerful data fitting and image processing capabilities of artificial intelligence (AI) have been demonstrated [21]. By employing AI techniques in integrated imaging and communication systems, ultra-high-precision sensing performance is expected [22]. This study considers RIS and AI as two key enablers for imaging.

The research on integrated imaging and communication is still in its early stages. Most of the related literature focuses on realizing imaging in communication systems. Therefore, this paper mainly summarizes communication-assisted imaging techniques but does not elaborate on the benefits that imaging brings to communication. The rest of this paper is organized as follows. Section 2 provides an overview of the two major technical roadmaps of integrated imaging and communication. Section 3 highlights the two key enablers of integrated imaging and communication. Section 4 analyzes the current challenges and future directions. Section 5 finally concludes this paper.

2 Technical Roadmaps

This section provides an overview of the technical roadmaps of integrated imaging and communication, including beamforming-based range estimation and inverse scattering problem-solving. Both of these techniques require transmitting and receiving pilot signals in communication systems. In the case of beamforming-based range estimation, the echo signal is considered to contain only the scattered waves from a single voxel. In contrast, for inverse scattering problem-solving, the echo signal contains the scattered waves from the entire ROI.

2.1 Imaging Based on Beamforming and Ranging

One of the main tasks of imaging is to capture the target shapes. Traditional LiDAR accomplishes this task using highly directional laser beams to scan the ROI and estimate the range of targets [23]. With the increasing communication frequency bands and antenna sizes, high-precision beamforming and scanning can be achieved [13, 24]. Assume that the centers of both the transmitting and receiving beams point to the same direction (ϕ, θ), where ϕ and θ represent the azimuth and elevation angles, respectively. Considering the extremely narrow beams generated by a large antenna array, we assume that the transmitted pilot signal $p(k)$ illuminates only voxels in direction (ϕ, θ), and **the received echo only contains the scattered waves from the voxel in this direction**. Here, k represents the wave number. As shown in Figure 1a, the transmitted signal is scattered by the ROI and finally arrives at the receiver. According to the radar equation [11] and supposing that the transmitting and receiving antenna gains
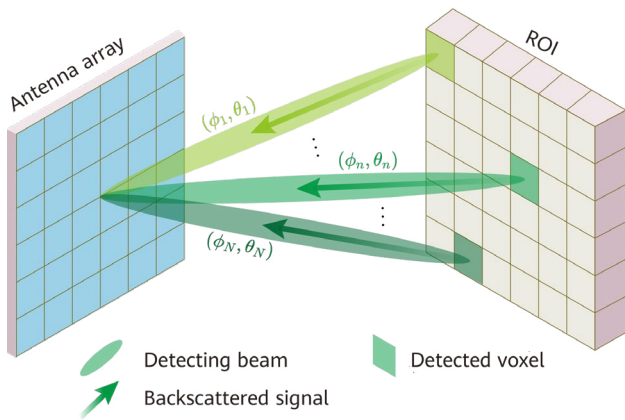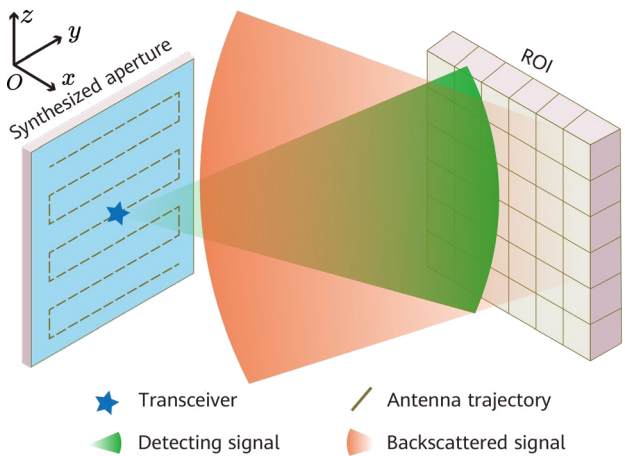
are 1, the received signal in the monostatic scenario can be expressed as

$$s(k) = \frac{\sigma(\phi, \theta)}{4\pi d(\phi, \theta)^2} e^{-j2kd(\phi,\theta)} p(k) + n(k), \quad (1)$$

where $d(\phi, \theta)$ and $\sigma(\phi, \theta)$ represent the distance and scattering coefficient of the voxel in direction (ϕ, θ), respectively; $n(k)$ is the additive noise at the receiver. For simplicity, only free-space path loss is considered in (1). Calculating the delay $\hat{\tau}$ between the transmitted and received signals, the voxel distance in direction (ϕ, θ) can be estimated by [13]

$$\hat{d}(\phi, \theta) = \hat{\tau}c/2, \quad (2)$$

where c denotes the speed of light. Then, the scattering coefficient of the voxel can be calculated by

$$\hat{\sigma}(\phi, \theta) = 4\pi \hat{d}(\phi, \theta)^2 \frac{|s(k)|}{|p(k)|}, \quad (3)$$

where $|\cdot|$ calculates the modulus. As shown in Figure 1a, **depth images** of the surrounding environments and voxel scattering coefficients can be obtained by scanning beams in all desired directions (ϕ, θ). The imaging method based on beamforming and ranging has been explored with personal mobile radars [12] and 5G BSs [13]. The imaging resolution is directly determined by the beam scanning resolution, whereas the resolution limit of range estimation is given as [13]

$$\delta_{\text{range}} = \frac{c}{2B}, \quad (4)$$

where B represents the signal bandwidth. Consequently, a higher bandwidth results in more precise ranging and imaging.

2.2 Imaging Based on Inverse Scattering Problem-Solving

The inverse scattering problem typically aims to recover the image information about an ROI from the received signals [22] and has been extensively studied in traditional synthetic aperture imaging fields [6, 7]. In the imaging scenario shown in Figure 1b[1], a single-antenna transceiver moves across the imaging plane and transmits and receives signals at each position, synthesizing a large imaging aperture. In communication systems, the transceiver scanning can be achieved using hand-held user equipment (UE) [25]. Assuming that $\sigma(x, y, z)$ represents the scattering

(a) Imaging based on beamforming and ranging

(b) Imaging based on inverse scattering problem-solving

Figure 1 Comparison between the two roadmaps

---

[1] Synthetic aperture imaging has been widely applied to various scenarios including monostatic and bistatic imaging with various aperture synthesis methods [8]. A simple scenario is used as an example herein.

coefficient function of the ROI, the synthesized aperture lies in the yOz plane, the aperture center is located at the origin, and the ROI is positioned right in front of the aperture. **The received signal at position $[0, y_\mathrm{a}, z_\mathrm{a}]^T$ is the superposition of all scattered waves from the voxels in the ROI**. In noise-free scenarios, the received signal can be expressed as [6, 7]

$$s^\circ(y_\mathrm{a}, z_\mathrm{a}, k) = \iiint \frac{\sigma(x,y,z)}{4\pi d(x,y,z)^2} p(k) e^{-j2kd(x,y,z)} dx dy dz, \quad (5)$$

where the distance between the transceiver at position $[0, y_\mathrm{a}, z_\mathrm{a}]^T$ and point $[x,y,z]^T$ is

$$d(x,y,z) = \sqrt{x^2 + (y-y_\mathrm{a})^2 + (z-z_\mathrm{a})^2}. \quad (6)$$

Based on the preceding model (5), FT and CS-based algorithms can be used for imaging [16].

1. *FT-based imaging:* By performing 2D FT on both sides of (5), we obtain [7]

$$S^\circ(k_y, k_z, k) = \iiint \frac{\sigma(x,y,z)}{4\pi} p(k) E(k_y, k_z, k) dx dy dz, \quad (7)$$

where, $S^\circ(k_y, k_z, k) = \mathrm{FT}_{2D}\{s^\circ(y,z,k)\}$,

$$E(k_y, k_z, k) = \frac{-4\pi^2}{k_x^2} e^{-jk_x x - jk_y y - jk_z z}, \quad (8)$$

where k_x, k_y, and k_z are the spatial Fourier domain variables corresponding to x, y, and z, respectively. Therefore, we have

$$S^\circ(k_y, k_z, k)\frac{k_x^2}{-\pi p(k)} = \Sigma(k_x, k_y, k_z) = \mathrm{FT}_{3D}\{\sigma(x,y,z)\}. \,(9)$$

Considering the influence of additive noise, the received signal can be represented as

$$s(y_\mathrm{a}, z_\mathrm{a}, k) = s^\circ(y_\mathrm{a}, z_\mathrm{a}, k) + n(y_\mathrm{a}, z_\mathrm{a}, k), \quad (10)$$

where $n(y_\mathrm{a}, z_\mathrm{a}, k)$ is additive white Gaussian noise (AWGN). Thus, the FT-based imaging method can be expressed as [6, 7]

$$\sigma(x,y,z) = \mathrm{FT}_{3D}^{-1}\left\{ \mathrm{FT}_{2D}\{s(y_\mathrm{a}, z_\mathrm{a}, k)\}\frac{k_x^2}{-\pi p(k)} \right\}. \quad (11)$$

In this approach, FT and inverse FT (IFT) operations can be realized by fast FT (FFT) and inverse FFT (IFFT), resulting in a significant reduction in computation time. Note that FT-based imaging requires the transceiver to perform uniform sampling on the imaging aperture, which is challenging when hand-held UEs are used as transceivers for imaging. Thus, imaging algorithms based on irregular sampling are desired [25–27]. The following subsection introduces CS-based imaging, which does not require uniform sampling on the aperture plane.

2. *CS-based imaging:* Suppose that the transceiver

transmits and receives signals at M positions, and the ROI is divided into N voxels at equal intervals. (10) can be expressed with discrete matrix forms, given as [9]

$$\mathbf{s} = \mathbf{A}\boldsymbol{\sigma} + \mathbf{n}, \quad (12)$$

where $\mathbf{s} \in \mathbb{C}^{M \times 1}$ is the received signals, $\mathbf{A} \in \mathbb{C}^{M \times N}$ is the sensing matrix, $\boldsymbol{\sigma} \in \mathbb{R}^{N \times 1}$ is the vector consists of voxel scattering coefficients, and $\mathbf{n} \in \mathbb{C}^{M \times 1}$ is the additive noise. The (m,n) element in $\mathbf{A}$ is given as

$$a_{m,n} = \frac{p(k)}{4\pi d_{m,n}^2} e^{-j2kd_{m,n}}, \quad (13)$$

where $d_{m,n}$ represents the distance between the m-th transceiver position and the n-th voxel. Since the targets typically occupy only a small part of the ROI, the vector $\boldsymbol{\sigma}$ is sparse in nature. Based on the theory of CS, fewer measurements than the Nyquist criterion can recover the sparse vector $\boldsymbol{\sigma}$ [15]. Specifically, the CS-based imaging problem can be presented as

$$\min_{\boldsymbol{\sigma}} \|\mathbf{s} - \mathbf{A}\boldsymbol{\sigma}\|_2 + \eta\|\boldsymbol{\sigma}\|_0, \quad (14)$$

or

$$\min_{\boldsymbol{\sigma}} \|\boldsymbol{\sigma}\|_0 \quad \text{subject to} \quad \|\mathbf{s} - \mathbf{A}\boldsymbol{\sigma}\|_2 \le \varepsilon, \quad (15)$$

where η and ε are hyperparameters involved in optimization. The above problems can be solved by various kinds of algorithms under the theory of CS, including the convex relaxation, greedy algorithm, and Bayesian estimation [28].

2.3 Discussions on Algorithm Performances

This subsection discusses the performances of the imaging methods mentioned above. In the case of imaging based on beamforming and ranging, the transmitted beamforming signals scan the ROI. The imaging performance is dependent on beamforming precision and delay estimation accuracy. Moreover, the computational complexity is related to the beam scanning times and complexity of delay estimation algorithms. Specifically, the beamforming precision depends on the number of antennas and codebook design. When antennas are insufficient, the beamforming precision is poor, and imaging is reduced to point target detection [13].

For imaging based on inverse scattering problem-solving, the transmitted signals act as broadcast signals for the ROI. The imaging performance and computational complexity primarily rely on the performance of the utilized algorithms.

Table 1 Comparison between FT- and CS-based imaging

Item	FT	CS
Required measurements	Many	Few
Required running memory	Low	High
Uniform sampling on the antenna aperture	Yes	No
Computational complexity	Low	High
Sidelobe reduction	No	Yes

The range resolution limit (along the x axis in Figure 1b) has been given in (4), whereas the cross-range resolution limit (along the y or z axis in Figure 1b) is given as

$$\delta_{\text{cross-range}} = \frac{\lambda_c}{4 \sin(\gamma/2)}, \quad (16)$$

where λ_c represents the wavelength of the center frequency, and γ is the angle subtended by the imaging aperture. FT and CS-based algorithms have unique advantages and disadvantages in imaging, and Table 1 compares their performances and requirements.

When the number of antennas is limited, beamforming-based range estimation is more suitable for less precise imaging in scenarios of large interested regions. In contrast, imaging based on inverse scattering problem-solving can potentially achieve high-precision imaging, owing to the implementation of the sparsity of the ROI with CS-based algorithms.

3 Key Enablers

3.1 RIS-assisted Imaging

RIS is a revolutionary technique with immense potential for future wireless communication systems [19, 20]. This subsection focuses on RIS-assisted imaging, where the RIS array is used as the imaging aperture. Since the RIS is equipped with a large antenna array, the requirement of multiple transmitting and receiving antennas can be released, allowing for imaging with a single-antenna transmitter and receiver. This helps lower system costs and energy consumption [19, 29]. Moreover, the RIS can intelligently tune the radio propagation environment to derive multiple measurements, eliminating the process of aperture synthesis required in traditional synthetic aperture imaging. This simplifies the process of measurement acquisition and enables real-time imaging owing to the fast reconfiguration ability of the RIS [30]. Additionally, since the RIS can enhance communication coverage [19], the imaging region can be extended and NLOS imaging can be

realized. Although a wide variety of designs and solutions can be developed in RIS-assisted integrated imaging and communication systems, the imaging roadmaps do not exceed the scope of Section 2. However, for common reflection-only RISs, their elements cannot receive or process signals [19], presenting additional challenges for RIS-assisted imaging.

- **RIS-assisted imaging based on beamforming and ranging.** The RIS consists of numerous sub-wavelength tunable elements, allowing their scattering phase shifts (and amplitudes) reconfigurable to intelligently control the EM propagation environments. Since the RIS possesses a large array, depth images can be obtained based on beamforming and ranging techniques with designed codebooks of RIS phase shifts [31], similar to the imaging techniques introduced in Section 2.1 with large antenna arrays.

- **RIS-assisted imaging based on inverse scattering problem-solving.** Because common reflection-only RISs cannot receive or process signals, RIS-assisted imaging is typically bistatic scenarios [32–35], where the signals are scattered by the RIS and received by a base station (BS), as depicted in Figure 2. According to channel models of RISs [19], the imaging problem can be modeled as the form of (12) by tuning RIS phase shifts, and CS-based algorithms are employed to realize imaging [32, 33]. Under the concept of inverse scattering problems, RIS phase shifts can be intelligently tuned to align multipath phases and enhance signal strength [34]. Moreover, the inverse scattering problem can be converted into a forward problem that is easier to solve [35].
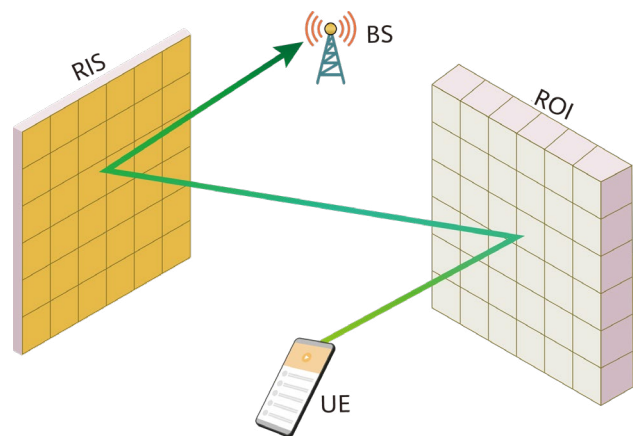
Figure 2 RIS-assisted integrated imaging and communication

3.2 AI-assisted Imaging

AI has been demonstrated to have powerful image

processing capabilities such as classification, generation, and super resolution [21]. Therefore, introducing AI techniques into integrated imaging and communication systems is expected to enhance imaging performances. According to the problem given in (12), the simplest idea is to train a neural network (NN) that directly maps the measurement $\mathbf{s}$ to the image $\boldsymbol{\sigma}$. However, this training method lacks the utilization of known physical knowledge, resulting in the need for a large amount of data, lengthy training time, and limited portability. Recently, model-driven NNs have made significant progress in addressing these challenges by leveraging known physical knowledge [22]. This subsection introduces the main methods of designing model-driven NNs for integrated imaging and communication.

- **NN-assisted image denoising.** This method uses traditional imaging algorithms to obtain initial images, which may contain many false detections and artifacts. Inputting initial images into an NN for denoising, imaging performances can be significantly improved [36].

- **Physics-assisted loss function design.** Considering the imaging problem given by (12) as an example, a physics-based loss function can be formulated as [37]

$$L = \beta \|\hat{\boldsymbol{\sigma}} - \boldsymbol{\sigma}\|_2^2 + \xi \|\mathbf{s} - \mathbf{A}\hat{\boldsymbol{\sigma}}\|_2^2, \tag{17}$$

where β and ξ are tunable parameters, and $\hat{\boldsymbol{\sigma}}$ represents the estimated result of $\boldsymbol{\sigma}$. This loss function effectively combines physical knowledge and data information during network training, leading to better training outcomes.

- **Unfolding of traditional iterative algorithms.** Most traditional CS-based algorithms are iterative in nature [28]. Each iteration of such an algorithm can be considered as one layer of an NN. Consequently, the iterative algorithm can be unfolded into a cascaded NN, whose parameters are learned from network training [38]. This approach integrates the physics-based model into the NN, significantly reducing the number of training parameters and achieving fast convergence.

4 Challenges and Future Outlook

The previous sections have reviewed the advances in integrated imaging and communication. However, this field is still developing, and numerous technical hurdles must be addressed. This section analyzes the research challenges and sheds light on potential directions for future research.

4.1 Challenges

- **Limited size of the imaging aperture.** Equation (16) suggests that a large aperture is necessary to enhance the cross-range resolution, which means that the aperture size limits the imaging resolution of communication systems. To overcome this limitation and improve imaging performances, widely distributed multi-user antennas have been used to synthesize a large aperture [10]. However, ultra-large apertures may lead to anisotropic scattering, where the scattering coefficients of the voxels vary at different observation angles [39], making the imaging problem more complex and difficult. Tackling the tradeoff between large aperture sizes and anisotropic scattering is a highly challenging task.

- **Limited bandwidth of communication systems.** According to (4), the range resolution is inversely proportional to the system bandwidth. However, the available bandwidth of communication systems is limited. The method of subband stitching proposed in [13] can obtain large effective bandwidths and partly overcome this limitation. However, the resulting complexities in system and algorithm design are relatively high. Furthermore, single-frequency 3D imaging is only feasible within a very short distance despite its proven feasibility [40]. Therefore, improving the range resolution is a crucial challenge.

- **Positioning error.** The current research on integrated imaging and communication assumes that the locations of transmitting and receiving antennas are already known [10, 32]. This assumption can be easily satisfied in monostatic imaging at the BS. However, in the imaging system shown in Figure 2, determining the accurate UE location is difficult. Any localization error leads to the inaccuracy of FT in (11) or modeling defects in the sensing matrix $\mathbf{A}$ in (12), which significantly degrades the imaging performances. Consequently, addressing the positioning errors for imaging is a key challenge.

- **System framework design and hardware implementation.** Implementing imaging functions requires a large number of pilot signals to capture the image information. Thus, it is crucial to allocate pilot and data symbols effectively within the limited time-frequency-space resources of the system. Additionally, CS-based imaging algorithms are characterized by high computational complexity, necessitating careful scheduling of computing resources.

In hardware implementation, system synchronization, phase, and quantization errors may occur [13], where phase errors substantially decline imaging performances. Thus, specifically designed algorithms are required to compensate for these hardware impairments.

4.2 Future Outlook

- **Application scenario design.** The current focus of integrated imaging and communication is mainly on imaging algorithm design, whereas the exploration of practical application scenarios is limited. Leveraging imaging results, various applications are anticipated, such as the 3D reconstruction of EM environments, the RCS estimation of far-field targets [41], the realization of AR, and the assistance of communication algorithm design. Some study employs imaging results for semantic segmentation and scene understanding [32, 42], which potentially contribute to security check and behavior detection fields [6]. Integrated imaging and communication is expected to empower a wide range of application scenarios.

- **Improvement and tradeoff of communication and imaging performances.** Existing research on integrated imaging and communication focuses on extracting image information from pilot-based channel estimation results. Thus, a large number of pilot signals are required owing to the high dimensions of the image to be recovered, probably resulting in reduced communication rates. The tradeoff between communication and imaging performances has been explored in [5]. Meanwhile, joint algorithm designs of imaging and communication are expected to enhance the overall estimation accuracy, considering that they both extract information from the received signals [10, 33]. Additionally, 3D EM environment reconstruction based on the imaging results can further aid in channel estimation and signal detection during communication, thereby degrading the bit error rate (BER).

- **Prototype setup and experimental validation.** As previously mentioned, the research on integrated imaging and communication is still in its early stages. Only a few studies have built prototype systems to verify the effectiveness of imaging algorithms [13, 25, 32]. Further experimental validations are required, and hardware impairments should be properly addressed.

5 Conclusion

In this paper, we have provided a comprehensive review of the advancements in integrated imaging and communication. We outlined two technical roadmaps: beamforming-based range estimation and inverse scattering problem-solving. Additionally, we showcased the use of advanced key enablers, such as RIS and AI techniques, to assist in imaging. Finally, we summarized the challenges and future directions. Integrated imaging and communication is an essential part of ISAC. It holds immense research potential, offers numerous opportunities for diverse applications, and is expected to play a pivotal role in future wireless communication systems.

References

[1] 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," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.

[2] J. Yang, S. Jin, C.-K. Wen *et al.*, "Model-based learning network for 3-D localization in mmWave communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5449–5466, Mar. 2021.

[3] C. D. Ozkaptan, E. Ekici, O. Altintas, and C.-H. Wang, "OFDM pilot-based radar for joint vehicular communication and radar systems," in Proc. *IEEE VNC*, Dec. 2018, pp. 1–8.

[4] J. Yang, C.-K. Wen, and S. Jin, "Hybrid active and passive sensing for SLAM in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2146–2163, Jul. 2022.

[5] N. Mehrotra and A. Sabharwal, "On the degrees of freedom region for simultaneous imaging & uplink communication," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1768–1779, Jun. 2022.

[6] D. M. Sheen, D. L. McMakin, and T. E. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," *IEEE Trans. Microwave Theory Tech.*, vol. 49, no. 9, pp. 1581–1592, Sep. 2001.

[7] X. Zhuge and A. G. Yarovoy, "Three-dimensional near-field MIMO array imaging using range migration techniques," *IEEE Trans. Image Process.*, vol. 21, no. 6, pp. 3026–3033, Jun. 2012.

[8] J. Gao et al., "An efficient algorithm for MIMO cylindrical millimeter-wave holographic 3-D imaging," *IEEE Trans. Microwave Theory Tech.*, vol. 66, no. 11, pp. 5065–5074, Nov. 2018.

[9] B. Zhang, W. Hong, and Y. Wu, "Sparse microwave imaging: Principles and applications," *Sci. China Inf. Sci.*, vol. 55, no. 8, pp. 1722–1754, Jul. 2012.

[10] X. Tong *et al.*, "Joint multi-user communication and sensing exploiting both signal and environment sparsity," *IEEE J. Sel. Top. Signal Process.*, vol. 15, no. 6, pp. 1409–1422, Nov. 2021.

[11] E. F. Knott, J. F. Schaeffer, and M. T. Tulley, *Radar cross section*, 2nd ed. Norwood, MA, USA: Artech House, 1993.

[12] F. Guidi, A. Guerra, and D. Dardari, "Personal mobile radars with millimeter-wave massive arrays for indoor mapping," *IEEE Trans. Mob. Comput.*, vol. 15, no. 6, pp. 1471–1484, Jun. 2016.

[13] J. Guan, A. Paidimarri, A. Valdes-Garcia, and B. Sadhu, "3-D imaging using millimeter-wave 5G signal reflections," *IEEE Trans. Microwave Theory Tech.*, vol. 69, no. 6, pp. 2936–2948, Jun. 2021.

[14] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[15] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[16] J. Fang *et al.*, "Fast compressed sensing SAR imaging based on approximated observation," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 1, pp. 352–363, Jan. 2014.

[17] M. F. Imani *et al.*, "Review of metasurface antennas for computational microwave imaging," *IEEE Trans. Antennas Propag.*, vol. 68, no. 3, pp. 1860–1875, Mar. 2020.

[18] L. Li, H. Ruan, C. Liu, Y. Li, Y. Shuang, A. Alù, C.-W. Qiu, and T. J. Cui, "Machine-learning reprogrammable metasurface imager," *Nat. Commun.*, vol. 10, no. 1, Mar. 2019, Art. no. 1082.

[19] W. Tang *et al.*, "Path loss modeling and measurements for reconfigurable intelligent surfaces in the millimeter-wave frequency band," *IEEE Trans. Commun.*, vol. 70, no. 9, pp. 6259–6276, Sep. 2022.

[20] Y. Huang, J. Yang, W. Tang, C.-K. Wen, S. Xia, and S. Jin, "Joint localization and environment sensing by harnessing NLOS components in RIS-aided mmWave communication systems," *IEEE Trans. Wireless Commun.*, early access, Apr. 2023.

[21] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.

[22] R. Guo, T. Huang, M. Li, H. Zhang, and Y. C. Eldar, "Physics-embedded machine learning for electromagnetic data imaging: Examining three types of data-driven imaging methods," *IEEE Signal Process Mag.*, vol. 40, no. 2, pp. 18–31, Mar. 2023.

[23] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-time loop closure in 2D LIDAR SLAM," in *Proc. IEEE ICRA*, May 2016, pp. 1271–1278.

[24] Z. Wang, X. Mu, and Y. Liu, "Terahertz near-field communications and sensing," [Online]. Available: https://arxiv.org/abs/2306.09723

[25] O. Li *et al.*, "Integrated sensing and communication in 6G: A prototype of high resolution THz sensing on portable device," in *Proc. EuCNC/6G Summit*, Jun. 2021, pp. 544–549.

[26] J. W. Smith and M. Torlak, "Efficient 3-D near-field MIMO-SAR imaging for irregular scanning geometries," *IEEE Access*, vol. 10, pp. 10 283–10 294, Jan. 2022.

[27] X. Li and Y. Chen, "Lightweight 2D imaging for integrated imaging and communication applications," *IEEE Signal Process Lett.*, vol. 28, pp. 528–532, Feb. 2021.

[28] Q. Zou and H. Yang, "A concise tutorial on approximate message passing," [Online]. Available: https://arxiv.org/abs/2201.07487

[29] J. Wang, W. Tang, J. C. Liang, L. Zhang, J. Y. Dai, X. Li, S. Jin, Q. Cheng, and T. J. Cui, "Reconfigurable intelligent surface: Power consumption modeling and practical measurement validation," [Online]. Available: https://arxiv.org/abs/2211.00323

[30] G. Castaldi, L. Zhang, M. Moccia, A. Y. Hathaway, W. X. Tang, T. J. Cui, and V. Galdi, "Joint multi-frequency beam shaping and steering via space–time-coding digital metasurfaces," *Adv. Funct. Mater.*, vol. 31, no. 6, Feb. 2021, Art. no. 2007620.

[31] A. Taha, H. Luo, and A. Alkhateeb, "Reconfigurable intelligent surface aided wireless sensing for scene depth estimation," [Online]. Available: https://arxiv.org/abs/2211.08210

[32] J. Hu, H. Zhang, K. Bian, Z. Han, H. V. Poor, and L. Song, "MetaSketch: Wireless semantic segmentation by reconfigurable intelligent surfaces," *IEEE Trans.*

Wireless Commun., vol. 21, no. 8, pp. 5916–5929, Aug. 2022.

[33] S. Zhu, Z. Yu, Q. Guo, J. Ding, Q. Cheng, and T. J. Cui, "RIS-assisted joint uplink communication and imaging: Phase optimization and Bayesian echo decoupling," [Online]. Available: https://arxiv.org/abs/2301.03817

[34] Y. He, D. Zhang, and Y. Chen, "High-resolution WiFi imaging with reconfigurable intelligent surfaces," *IEEE Internet Things J.*, vol. 10, no. 2, pp. 1775–1786, Jan. 2023.

[35] Y. Jiang, F. Gao, Y. Liu, S. Jin, and T. Cui, "Near field computational imaging with RIS generated virtual masks," [Online]. Available: https://arxiv.org/abs/2304.11510

[36] Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1849–1860, Sep. 2018.

[37] Y. Jin, Q. Shen, X. Wu, J. Chen, and Y. Huang, "A physics-driven deep-learning network for solving nonlinear inverse problems," *Petrophysics*, vol. 61, no. 1, pp. 86–98, Feb. 2020.

[38] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Model-driven deep learning for MIMO detection," *IEEE Trans. Signal Process.*, vol. 68, pp. 1702–1715, Feb. 2020.

[39] M. Çetin *et al.*, "Sparsity-driven synthetic aperture radar imaging: Reconstruction, autofocusing, moving targets, and compressed sensing," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 27–40, Jul. 2014.

[40] T. Fromenteze, M. Boyarsky, J. Gollub, T. Sleasman, M. Imani, and D. R. Smith, "Single-frequency near-field MIMO imaging," in *Proc. EUCAP*, Mar. 2017, pp. 1415–1418.

[41] A. Broquetas, J. Palau, L. Jofre, and A. Cardama, "Spherical wave near-field imaging and radar cross-section measurement," *IEEE Trans. Antennas Propag.*, vol. 46, no. 5, pp. 730–735, May 1998.

[42] X. Tong, Z. Zhang, and Z. Yang, "Multi-view sensing for wireless communications: Architectures, designs, and opportunities," *IEEE Commun. Mag.*, vol. 61, no. 5, pp. 40–46, May 2023.

Information Theory Limits for Integrated Sensing and Communication over Gaussian Channels

Fan Liu [1], Yifeng Xiong [2], Weijie Yuan [1], Yuanhao Cui [1], Xiao Han [3]

[1] Southern University of Science and Technology

[2] Beijing University of Posts and Telecommunications

[3] Wireless Technology Lab, Huawei Technologies Co., Ltd.

1 Introduction

The next-generation wireless communication system is making a wealth of promising technologies a reality, ranging from smart cities to smart factories, Internet of Vehicles (IoV), drone-based low-altitude economy, etc. [1]. These technologies have a common trait that demands high-quality communication and robust, high-precision sensing. IoV, for example, requires a communication latency of less than 10 ms and a positioning precision down to the centimeter level to implement vehicle-vehicle and vehicle-road collaboration and high-precision sensing of the traffic environment. However, existing 4G/5G networks cannot cope with such service requirements effectively. For example, the LTE network can achieve a communication rate of 100 Mbps and an end-to-end (E2E) latency of hundreds of milliseconds. The 5G radio access technology introduced in 3GPP Release 16 can achieve an indoor positioning precision of less than 3 m and an outdoor positioning precision of less than 10 m. Even so, meeting the communication and sensing requirements described above is difficult. Deploying ubiquitous wireless sensing functionality in future 5G-Advanced (5G-A) and 6G wireless networks is considered an effective approach to these challenges. Under this background, integrated sensing and communication (ISAC) is envisioned to be one of the game-changing technologies for 5G-A and 6G networks [2].

From the technical trend perspective, the existing radar and communication systems are becoming strikingly similar in their hardware architecture, channel characteristics, and signal processing method. Therefore, a combination or even integration of them is deemed inevitable. From the application requirement perspective, many emerging 5G-A and 6G applications, for example, IoT applications like smart factories and smart homes, and smart transportation applications like IoV and autonomous driving, require joint sensing and communication design. Being able to efficiently utilize radio resources such as spectrum and power in emerging application scenarios, ISAC is well-recognized as a low-cost solution that can greatly improve resource usage. In light of this, in the *Report on Global Engineering Frontier 2020*, the Chinese Academy of Engineering ranked ISAC as the top 2 technology priority in the information and electronics engineering field. In 2021, the Ministry of Science and Technology of China listed the 6G communication-sensing-computing converged network architecture as one of the key technologies in the guideline for the multimodal network and communication project of the "14th Five-Year" National Key R&D Plan. In the industry, Huawei proposed its three visions for 5G-A at Global MBB Forum 2020. One of them is the implementation of ISAC, which requires cellular networks to provide both communication and sensing capabilities. Besides, communication and sensing are expected to be collaborative with each other for mutual benefit. According to China Mobile, service requirements in the 6G era are characterized by deep sensing and understanding of the physical world. Meanwhile, international standardization organizations such as ITU, 3GPP, and IEEE have also conducted extensive ISAC research. ITU and 3GPP treat ISAC as a work item and a study item, respectively, and undertake research on scenario requirements, use cases, signal designs, and transmission protocols. IEEE has started the research and drafting of the IEEE 802.11bf WLAN sensing standard, aiming to standardize the Wi-Fi sensing technology. On top of that, Next G Alliance in the US, Hexa-X in the EU, and IMT-2030 (6G) Promotion Group in China have developed a myriad of standards proposals and industry white papers concerning ISAC. In June 2023, ITU defined the overall vision of 6G in IMT Vision – Framework and overall objectives of the future development of IMT for 2030 and beyond, marking the official opening of the 6G era [3]. This report formally establishes ISAC as one of the six typical application scenarios of 6G. ISAC is capable of wide-area multi-dimensional sensing, which helps to obtain spatial information about unknown objects, connected devices, and ambient environments. It applies to innovative applications such as automation, safe driving, and digital twins and can be combined with AI to further enhance sensing of the physical environment. Hence, ISAC is anticipated to be a key enabling technology for new use cases.

Generally, ISAC systems use two types of performance evaluation metrics [4]. In an ISAC system, communication and sensing subsystems share radio resources and hardware platforms. This requires the ISAC system to fulfill communication information transmission and target information sensing at the same time under constraints such as time, frequency, and power. However, it is impossible to achieve the optimal communication performance and sensing performance at the same time because of mutual restriction between performance indicators such as the sensing precision, detection probability, communication rate, and bit error rate (BER). For this reason, a performance tradeoff is inevitable. That's also why it is unrealistic to analyze the performance bound of the ISAC system based on classical theories. A basic

theoretical problem that needs to be resolved urgently is how to depict the performance tradeoff and coupling between communication and sensing. To reveal the performance tradeoff mechanism, the core idea is to study the performance limits of the communication and sensing subsystems and then define their performance metrics and achievable regions. This is an open issue in the ISAC field, and there is no definite general conclusion in academia [5, 6].

This paper introduces the latest research progress of the author team in ISAC performance limits. In particular, this paper focuses on the Gaussian channel, the most widely applicable channel model, and considers point-to-point (P2P) communication in monostatic or bistatic sensing. For the first time, this paper defines a Miller-Chang type, namely, Cramér-Rao bound (CRB) for stochastic ISAC signals and provides the lower error bound for unbiased estimation of target parameters. Further, it analyzes the achievable region (that is, CRB-rate region) of the CRB and communication rate and especially two important corner points — sensing-optimal corner point P_{SC} and communication-optimal corner point P_{CS} — on the Pareto boundary of the region. Specifically, this paper provides the signal structure, high-SNR communication capacity, and achieving strategy for optimal sensing at P_{SC}, and the upper and lower bounds of the sensing CRB at P_{CS}. Our research shows a two-fold tradeoff, namely, subspace tradeoff and deterministic-random tradeoff, between the communication and sensing performance in ISAC systems. The former mainly depends on the allocation strategy for radio resources between the communication and sensing subsystems, and the latter depends on the modulation and coding scheme for ISAC signals. Particularly, the deterministic-random tradeoff points out that the randomness of ISAC signals is conducive to communication, while the determinism is conducive to sensing. This empirical criterion is widely seen in ISAC signal design strategies but has not yet been proven by mathematical analysis. The conclusion of this paper provides rigorous theoretical support for this criterion and is expected to play a guiding role in the design and implementation of ISAC systems.

2 ISAC System Model

This paper bases on a typical monostatic or bistatic P2P downlink ISAC system as shown in Figure 1. The system consists of an ISAC Tx, a sensing Rx, a communication Rx, and one or more sensing targets. The ISAC Tx transmits
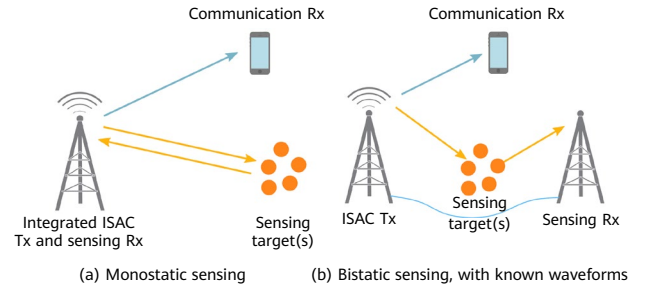
Figure 1 Typical monostatic or bistatic P2P downlink ISAC system

ISAC signals to simultaneously perform target sensing and downlink communication. Over a vector Gaussian channel, the ISAC signals received at the sensing Rx and communication Rx can be respectively modeled as:

$$\begin{aligned}
\mathbf{Y}_s &= \mathbf{H}_s\left(\boldsymbol{\eta}\right)\mathbf{X} + \mathbf{Z}_s \\
\mathbf{Y}_c &= \mathbf{H}_c\mathbf{X} + \mathbf{Z}_c
\end{aligned} \tag{1}$$

where, $\mathbf{X} \in \mathbb{C}^{M \times T}$ is the ISAC signal matrix transmitted from the ISAC Tx, with M being the number of antennas at the ISAC Tx and T being the number of discrete samples; $\mathbf{H}_s \in \mathbb{C}^{N_s \times M}$ and $\mathbf{H}_c \in \mathbb{C}^{N_c \times M}$ are ISAC channel matrices, with N_s and N_c being the numbers of antennas at the sensing Rx and communication Rx, respectively; $\mathbf{Z}_s \in \mathbb{C}^{N_s \times T}$ and $\mathbf{Z}_c \in \mathbb{C}^{N_c \times T}$ are zero mean white Gaussian noise matrices with variances σ_s^2 and σ_c^2, respectively. In particular, $\boldsymbol{\eta} \in \mathbb{R}^K$ represents the target parameters of interest, e.g., angle, range, and Doppler velocity. Without loss of generality, the sensing channel matrix $\mathbf{H}_s : \mathbb{R}^K \to \mathbb{C}^{N_s \times M}$ is assumed to be a deterministic function of $\boldsymbol{\eta}$. A typical example of the sensing channel model is the multi-target angle estimation model of the collocated multiple-input multiple-output (MIMO) radar. Suppose L targets are involved in monostatic MIMO sensing, the angle of each target relative to the antenna array of the ISAC base station is θ_l $(l = 0, 1, 2, \ldots, L)$, and the complex amplitude generated by the product of the path loss between each target and the ISAC base station and the radar cross section (RCS) of the target is denoted as β_l $(l = 0, 1, 2, \ldots, L)$. The sensing channel matrix $\mathbf{H}_s$ can be modeled as:

$$\mathbf{H}_s = \sum_{l=1}^{L} \beta_l \mathbf{b}\left(\theta_l\right)\mathbf{a}^T\left(\theta_l\right) \tag{2}$$

where, $\mathbf{a}\left(\theta\right)$ and $\mathbf{b}\left(\theta\right)$ represent array vectors of the Tx and Rx antennas, respectively. In this case, the sensing parameters are defined as $\boldsymbol{\eta} = \left[\theta_1, \theta_2, \ldots, \theta_L, \beta_1, \beta_2, \ldots, \beta_L\right]^T$.

In a typical sensing system, the sensing Rx is usually collocated with the ISAC Tx (monostatic sensing) or is connected to the ISAC Tx via an optical fiber (bistatic

sensing), as shown in Figure 1. Therefore, the ISAC signal $\mathbf{X}$ is known to both the ISAC Tx and sensing Rx as a reference sensing signal. On the other hand, as $\mathbf{X}$ contains useful information intended for the communication Rx, it is a random signal unknown to the communication Rx. Therefore, we model $\mathbf{X}$ as a random matrix following a distribution $p_{\mathbf{X}}(\boldsymbol{X})$, whose realization is known to the ISAC Tx and sensing Rx, but not to the communication Rx. We also assume $\mathbb{E}\{\mathbf{X}\} = \mathbf{0}$, and denote the sample and statistical covariance matrices of $\mathbf{X}$ as $\mathbf{R}_X = T^{-1}\mathbf{X}\mathbf{X}^H$ and $\tilde{\boldsymbol{R}}_X = \mathbb{E}\{\mathbf{R}_X\}$, respectively.

Without loss of generality, we define ISAC tasks in the ISAC system as:

- **Sensing task**: Estimate $\boldsymbol{\eta}$ from the observation $\mathbf{Y}_s$ at the sensing Rx, with the knowledge of the probing signal $\mathbf{X}$.

- **Communication task**: Recover the useful information contained in $\mathbf{X}$ from the received signal $\mathbf{Y}_c$ at the communication Rx, with the knowledge (or statistical knowledge) of the channel $\mathbf{H}_c$.

Besides, we assume $\boldsymbol{\eta} \sim p_{\boldsymbol{\eta}}(\boldsymbol{\eta})$, which varies every T samples in an i.i.d. manner. Accordingly, $\mathbf{H}_c$ is also assumed to vary in an i.i.d. manner every T samples.

3 Communication and Sensing Performance Metrics

3.1 Communication Performance Metrics

For P2P channels, the communication performance can be measured by the achievable rate. Since $\mathbf{H}_c$ is assumed to be a block-fading channel, the achievable rate is equal to the ergodic rate, which is defined as:

$$I_c = \max_{p_{\mathbf{X}}(\boldsymbol{X})} T^{-1} I(\mathbf{Y}_c; \mathbf{X} | \mathbf{H}_c), \ \text{s.t.} \ p_{\mathbf{X}}(\boldsymbol{X}) \in \mathcal{F} \quad (3)$$

where, $I(\mathbf{Y}_c; \mathbf{X} | \mathbf{H}_c)$ stands for the mutual information between $\mathbf{X}$ and $\mathbf{Y}_c$ conditioned on $\mathbf{H}_c$, and $\mathcal{F}$ represents the feasible set of the distribution $p_{\mathbf{X}}(\boldsymbol{X})$ under conditions such as power and sensing performance constraints.

3.2 Sensing Performance Metrics

The estimation performance of target parameters is usually measured by mean squared error (MSE) in sensing tasks. In model (1), however, the MSE is jointly determined by the functional relationship between $\mathbf{H}_s$ and $\boldsymbol{\eta}$, probability distribution of $\boldsymbol{\eta}$, and estimator in use, and therefore cannot be represented by closed-form expressions in most cases. To this end, we consider using Bayesian CRB (BCRB) as the sensing performance metric. It is the MSE lower bound obtained by performing weakly unbiased estimation on the random parameter $\boldsymbol{\eta}$ and may be asymptotically achieved by the maximum a posteriori (MAP) estimator at a high SNR. In other words, the BCRB is equivalent to the MSE obtained by the MAP estimator at a high SNR.

Given a random implementation of ISAC signals, the MSE of any weakly unbiased estimator has the following BCRB:

$$\mathrm{MSE}_{\boldsymbol{\eta}|\mathbf{X}} \geq \mathrm{Tr}\left(\mathbf{J}_{\boldsymbol{\eta}|\mathbf{X}}^{-1}\right) \quad (4)$$

where, $\mathbf{J}_{\boldsymbol{\eta}|\mathbf{X}}$ is a Bayesian Fisher information matrix (BFIM) about $\boldsymbol{\eta}$ under a random implementation $\boldsymbol{X}$, and may be represented by the following affine transformation:

$$\begin{aligned}\mathbf{J}_{\boldsymbol{\eta}|\mathbf{X}} &= \boldsymbol{\Phi}(\mathbf{R}_X) \\ &\triangleq \frac{T}{\sigma_s^2}\left(\sum_{i=1}^{r_1} \mathbf{F}_i \mathbf{R}_X \mathbf{F}_i^H + \sum_{j=1}^{r_2} \mathbf{G}_j \mathbf{R}_X^T \mathbf{G}_j^H\right) + \mathbf{J}_P\end{aligned} \quad (5)$$

where, $\mathbf{R}_X$ is a sample covariance matrix, r_1, r_2, $\mathbf{F}_i$ and $\mathbf{G}_j$ are partitioned from the Jacobian matrix $\tilde{\mathbf{F}} = \frac{\partial \operatorname{vec}(\mathbf{H}_s)}{\partial \boldsymbol{\eta}}$, and $\mathbf{J}_P$ is the prior Fisher information contributed by the parameter prior distribution $p_{\boldsymbol{\eta}}(\boldsymbol{\eta})$, which may be expressed as:

$$\mathbf{J}_P = \mathbb{E}_{\boldsymbol{\eta}}\left\{\frac{\partial \ln p_{\boldsymbol{\eta}}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \frac{\partial \ln p_{\boldsymbol{\eta}}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}^T}\right\} \quad (6)$$

Due to limited space, this paper does not provide a detailed analysis on BFIM's structure. Interested readers can refer to [7].

It is noted that the lower bound in Equation (4) can only describe the sensing performance at a random implementation of the given ISAC signal. Therefore, it may vary with the random change of the ISAC signal. To measure the overall estimation performance of the ISAC system, we consider calculating the expectation of $\mathbf{X}$ on both the left and right sides of Equation (4), as shown below:

$$\begin{aligned}\mathrm{MSE}_{\boldsymbol{\eta}} &\triangleq \mathbb{E}_{\mathbf{X}}\left\{\mathrm{MSE}_{\boldsymbol{\eta}|\mathbf{X}}\right\} \\ &\geq \mathbb{E}_{\mathbf{X}}\left\{\mathrm{Tr}\left(\mathbf{J}_{\boldsymbol{\eta}|\mathbf{X}}^{-1}\right)\right\} = \mathbb{E}_{\mathbf{X}}\left\{\mathrm{Tr}\left(\left[\boldsymbol{\Phi}(\mathbf{R}_X)\right]^{-1}\right)\right\}\end{aligned} \quad (7)$$

The preceding equation is referred to as Miller-Chang type CRB (MCB) [8] for estimating the parameter $\boldsymbol{\eta}$. It treats $\mathbf{X}$ as a nuisance parameter and calculates the

expectation of the CRB over $\mathbf{X}$. We point out again that the ISAC signal must be a **random but known** signal for the sensing Rx because the sensing Rx collaborates with the ISAC Tx, and the ISAC signal needs to carry useful information. Therefore, the MCB is especially suitable for the ISAC system. In other words, because the ISAC system performs sensing by using random signals, the sensing performance of the ISAC system definitely depends on the probability distribution of the signals. Unless otherwise specified, Equation (7) is referred to as the CRB for short.

4 Information Theory Limits of the ISAC System

The performance analysis of the ISAC system is quite different from that of the discrete communication or sensing system. First, the ISAC system uses random signals that carry useful communication information for sensing. In contrast, the sensing system uses traditional radar signals, where are deterministic or pseudo-random signals that do not carry information. Therefore, sensing performance metrics, for example, MCB described in this paper, need to be defined for the ISAC system. Second, in the ISAC system, the communication and sensing subsystems share radio resources but use different performance metrics. As a result, a performance tradeoff is inevitable between the communication and sensing subsystems; that is, the two subsystems cannot achieve optimal performance simultaneously. As shown in Figure 2, this tradeoff and the corresponding ISAC performance bound can be measured by the Pareto frontier of the achievable rate (3) and the sensing CRB (7). A CRB-rate pair is considered achievable when and only when being located in the Pareto boundary's internal area shown in the figure. Otherwise, the corresponding communication and/or sensing performance cannot be achieved in the ISAC system.
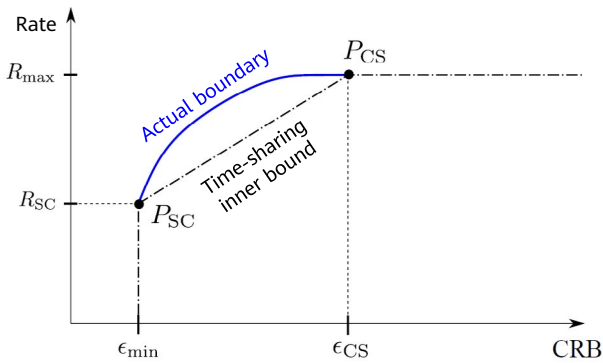
Figure 2 CRB-rate region

Specifically, the measurement of the Pareto boundary can be modeled as the following functional optimization problem:

$$\max_{p_{\mathbf{X}}(\boldsymbol{X})} \quad I\left(\mathbf{Y}_c; \mathbf{X} \,|\, \mathbf{H}_c\right)$$
$$\text{s.t.} \quad \mathbb{E}_{\mathbf{X}}\left\{\text{Tr}\left(\left[\boldsymbol{\Phi}\left(\mathbf{R}_X\right)\right]^{-1}\right)\right\} \leq \epsilon, \quad (8)$$
$$\mathbb{E}_{\mathbf{X}}\left\{\text{Tr}\left(\mathbf{R}_X\right)\right\} \leq P_T.$$

Under a given sensing CRB constraint, the probability distribution $p_{\mathbf{X}}(\boldsymbol{X})$ of ISAC signals can be optimized to maximize the mutual information for communication. Although the probability distribution function $p_{\mathbf{X}}(\boldsymbol{X})$ is a convex optimization problem, its solution is quite complex. The reason is that the functional optimization problem is actually an **infinite-dimensional** optimization problem. It usually needs to go through quantization before being solved by using numerical methods such as the Blahut-Arimoto algorithm, and moreover, obtaining a closed-form expression is difficult. To reveal the essential insights of the performance tradeoff, this paper considers the two corner points on the Pareto boundary of the CRB-rate region, that is, the minimum CRB P_{CS} when the communication rate is the maximum and the maximum communication rate P_{SC} when the CRB is minimum. The time-sharing inner bound of the CRB-rate region can be obtained through time-sharing between the two corner points. In what follows, we present a comparative analysis on the performance at the two corner points.

4.1 Communication and Sensing Performance at P_{SC}

We first look at the communication and sensing performance at P_{CS}. It is noted that the communication performance and achieving strategy of this point are well known. Under the constraint of the average power $\mathbb{E}_{\mathbf{X}}\left\{\text{Tr}\left(\mathbf{R}_X\right)\right\} \leq P_T$, the capacity of P2P Gaussian channels is achieved when the channel input distribution is equal to Gaussian distribution. In other words, at P_{CS}, each column of $\mathbf{X}$ should be independently and identically distributed in the circular symmetric complex Gaussian distribution $\mathcal{CN}\left(\mathbf{0}, \tilde{\boldsymbol{R}}_{\text{CS}}^\star\right)$. The statistical covariance matrix $\tilde{\boldsymbol{R}}_{\text{CS}}^\star$ is given by the optimal solution of the following rate maximization problem:

$$R_{\max} = \max_{\boldsymbol{R} \succcurlyeq 0, \boldsymbol{R} = \boldsymbol{R}^H} \mathbb{E}_{\mathbf{H}_c}\left\{\log\left|\mathbf{I} + \sigma_c^{-2}\mathbf{H}_c\boldsymbol{R}\mathbf{H}_c^H\right|\right\}$$
$$\text{s.t.} \quad \text{Tr}\left(\boldsymbol{R}\right) \leq P_T \quad (9)$$
$$= \mathbb{E}_{\mathbf{H}_c}\left\{\log\left|\mathbf{I} + \sigma_c^{-2}\mathbf{H}_c\tilde{\boldsymbol{R}}_{\text{CS}}^\star\mathbf{H}_c^H\right|\right\}$$

According to the preceding solution, we can easily get the following eigenvalue decomposition structure:

$$\tilde{R}_{\mathrm{CS}}^{\star} = \mathbf{U}_c \mathbf{\Lambda}_c^{\star} \mathbf{U}_c^H \tag{10}$$

where, $\mathbf{U}_c$ denotes the right singular matrix of the communication channel $\mathbf{H}_c$, and $\mathbf{\Lambda}_c^{\star}$ denotes the optimal eigenvalue matrix, in which the eigenvalue can be solved by the well-known Water-Filling power allocation algorithm. On this basis, it is easy to know that the ISAC signal at P_{CS} should have the following structure:

$$\mathbf{X}_{\mathrm{CS}}^{\star} = \left(\tilde{R}_{\mathrm{CS}}^{\star} \right)^{1/2} \mathbf{D} = \mathbf{U}_c (\mathbf{\Lambda}_c^{\star})^{1/2} \mathbf{D} \tag{11}$$

where, $\mathbf{D}$ is a random matrix, each column of which is independently and identically distributed in the standard complex Gaussian distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I})$.

The natural question is, "What is the sensing performance at P_{CS}?" Because the achieving strategy for P_{CS} is Gaussian distribution, this question is equivalent to "What is the CRB when a Gaussian signal is used for sensing?" It can be seen from Equation (7) that the CRB depends on the sample covariance matrix $\mathbf{R}_X = T^{-1} \mathbf{X} \mathbf{X}^H$ other than the statistical covariance matrix $\tilde{R}_X = \mathbb{E}\{\mathbf{R}_X\}$. Particularly, at P_{CS}, each column of $\mathbf{X}$ is independently and identically distributed in $\mathcal{CN}(\mathbf{0}, \tilde{R}_{\mathrm{CS}}^{\star})$. Therefore, $\mathbf{R}_X$ obeys the complex Wishart distribution. The CRB at P_{CS} is denoted as ϵ_{CS}. Note that $\mathrm{Tr}\left([\mathbf{\Phi}(\mathbf{R}_X)]^{-1}\right)$ is a convex function of $\mathbf{R}_X$, and by the Jensen's inequality, the CRB has the following lower bound:

$$\epsilon_{\mathrm{CS}} \triangleq \mathbb{E}_{\mathbf{X}}\left\{ \mathrm{Tr}\left([\mathbf{\Phi}(\mathbf{R}_X)]^{-1}\right) \right\} \geq \mathrm{Tr}\left\{ (\mathbf{\Phi}[\mathbb{E}(\mathbf{R}_X)])^{-1} \right\}$$
$$= \mathrm{Tr}\left\{ \left[\mathbf{\Phi}\left(\tilde{R}_{\mathrm{CS}}^{\star}\right)\right]^{-1} \right\} \tag{12}$$

It is noted that the preceding equation holds for arbitrarily distributed $\mathbf{R}_X$. We can see that, for $\mathbf{R}_X$ that obeys the Wishart distribution, there is a deterioration in the CRB performance. This is because the Jensen lower bound is obtained only when $\mathbf{R}_X = \mathbb{E}(\mathbf{R}_X)$, and the Wishart matrix meets this condition only when $T \to \infty$.

On the other hand, it is proved in [7] that the sensing CRB at P_{CS} also has the following upper bound:

$$\epsilon_{\mathrm{CS}} \leq \frac{T}{T - \min\{K, M_{\mathrm{CS}}\}} \mathrm{Tr}\left\{ \left[\mathbf{\Phi}\left(\tilde{R}_{\mathrm{CS}}^{\star}\right)\right]^{-1} \right\} \tag{13}$$

The upper bound given by Equation (13) indicates that the sensing performance loss at P_{CS} is determined by the number of sensing parameters K and the rank $M_{\mathrm{CS}} = \mathrm{rank}\left(\tilde{R}_{\mathrm{CS}}^{\star}\right)$ of the statistical covariance matrix $\tilde{R}_{\mathrm{CS}}^{\star}$. In particular, we have $\mathrm{rank}\left(\tilde{R}_{\mathrm{CS}}^{\star}\right) = \mathrm{rank}(\mathbf{H}_c)$ when the communication SNR is sufficiently high. Once again, we can see that the ISAC

system does not incur sensing performance loss when and only when $T \to \infty$.

4.2 Communication and Sensing Performance at P_{SC}

At this point, the achieving strategy for P_{CS} is known, but the communication and sensing performance and achieving strategy for P_{SC} are still unknown. To answer this question, we need to consider the Jensen's inequality again. It is noted that for any complex semi-positive definite matrix $\mathbf{R}_X$ whose average power is less than P_T, we have:

$$\mathbb{E}_{\mathbf{X}}\left\{ \mathrm{Tr}\left([\mathbf{\Phi}(\mathbf{R}_X)]^{-1}\right) \right\} \geq \mathrm{Tr}\left\{ (\mathbf{\Phi}[\mathbb{E}(\mathbf{R}_X)])^{-1} \right\}$$
$$\geq \mathrm{Tr}\left\{ \left[\mathbf{\Phi}\left(\tilde{R}_{\mathrm{SC}}^{\star}\right)\right]^{-1} \right\} \triangleq \epsilon_{\min} \tag{14}$$

where, $\tilde{R}_{\mathrm{SC}}^{\star}$ is the optimal solution to the following deterministic optimization problem:

$$\tilde{R}_{\mathrm{SC}}^{\star} = \underset{\mathbf{R} \succeq \mathbf{0}, \ \mathbf{R} = \mathbf{R}^H}{\arg\min} \ \mathrm{Tr}\left\{ [\mathbf{\Phi}(\mathbf{R})]^{-1} \right\} \ \mathrm{s.\,t.} \ \mathrm{Tr}(\mathbf{R}) \leq P_T \tag{15}$$

Note that Equation (15) is a convex optimization problem. Specifically, the problem is a semi-definite programming (SDP) problem, which, however, is not strictly convex, and its optimal solution is not unique. It is known that the optimal solutions for an SDP problem are all in a subspace formed by a maximum-rank solution. To this end, the optimal solution to the problem can be parameterized as follows:

$$\tilde{R}_{\mathrm{SC}}^{\star} = \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^H \tag{16}$$

where, $\mathbf{U}_s$ is an eigenvector matrix of a maximum-rank solution, and $\mathbf{\Lambda}_s$ is a semi-definite complex symmetric matrix.

In most cases, it can be proved that the preceding problem has a unique solution. If the uniqueness holds, the equal sign in Equation (14) holds if and only if:

$$\mathbf{R}_X = \mathbb{E}(\mathbf{R}_X) = \tilde{R}_{\mathrm{SC}}^{\star} \tag{17}$$

According to the preceding equation, to obtain CRB's global minimum value $\epsilon_{\min}$, the sample covariance matrix $\mathbf{R}_X$ must be a **deterministic matrix** [9, 10]. This result is non-trivial because in most cases, $\mathbf{R}_X = T^{-1} \mathbf{X} \mathbf{X}^H$ should be a random matrix due to the reason that $\mathbf{X}$ is a random matrix. The sample covariance matrix $\mathbf{R}_X$ is deterministic, but this does not mean that the ISAC signal is also deterministic. Conversely, $\mathbf{X}$ may still be a random signal carrying information and may be expressed as:

$$\mathbf{X}_{\mathrm{SC}}^{\star} = \sqrt{T}\left(\tilde{R}_{\mathrm{SC}}^{\star}\right)^{1/2} \mathbf{Q} = \sqrt{T} \mathbf{U}_s \mathbf{\Lambda}_s^{1/2} \mathbf{Q} \tag{18}$$

where, $\mathbf{Q} \in \mathbb{C}^{M_{\mathrm{SC}} \times T}$ is a random semi-unitary matrix satisfying $\mathbf{Q}\mathbf{Q}^H = \mathbf{I}$ and $M_{\mathrm{SC}} = \mathrm{rank}\left(\tilde{\boldsymbol{R}}_{\mathrm{SC}}^{\star}\right)$. Note that due to $\left(\tilde{\boldsymbol{R}}_{\mathrm{SC}}^{\star}\right)^{1/2}$ being a deterministic matrix, the communication's degrees of freedom (DoF) at P_{SC} is contributed only by the probability distribution of the $\mathbf{Q}$ matrix.

Based on the preceding conclusion, the problem of solving the achievable communication rate at P_{SC} may be summarized as searching for a probability distribution $p_{\mathbf{Q}}\left(\boldsymbol{Q}\right)$ on a set formed by all $M_{\mathrm{SC}} \times T$-dimensional semi-unitary matrices, so as to maximize the mutual information $I\left(\mathbf{Q}; \mathbf{Y}_c \,|\, \mathbf{H}_c\right)$. In the differential geometry theory, this set is also called $M_{\mathrm{SC}} \times T$-dimensional Stiefel manifold, which is generally denoted as $\mathcal{S}\left(T, M_{\mathrm{SC}}\right)$. More precisely, this problem is equivalent to the sphere packing problem on the Stiefel manifold. According to [7], it can be proved that at P_{SC}, the asymptotic achievable rate at a high SNR is:

$$R_{\mathrm{SC}} = \left(1 - \frac{M_{\mathrm{SC}}}{2T}\right) \mathbb{E}_{\mathbf{H}_c} \left\{\log\left|\sigma_c^{-2}\mathbf{H}_c\tilde{\boldsymbol{R}}_{\mathrm{SC}}^{\star}\mathbf{H}_c^H\right| + c_0\right\} + \mathcal{O}\left(\sigma_c^2\right) \quad (19)$$

where,

$$c_0 = \frac{L}{T}\left[\left(T - \frac{L}{2}\right)\log\frac{T}{e} - \log\Gamma\left(T\right) + \log 2\sqrt{\pi}\right] \quad (20)$$

and $c_0 \to 0$, $T \to \infty$.

It can be immediately observed from Equation (19) that there is no loss of communication freedom when $T \to \infty$. This is because even for a Gaussian matrix, its rows are asymptotically orthogonal as T increases; that is, the Gaussian matrix is asymptotically equivalent to a semi-unitary matrix.

5 Discussion

In the preceding theoretical analysis, both the CRB in Equation (7) and the communication rate in Equation (19) are progressively achievable at a high SNR. From the preceding theoretical results, we can conclude two important tradeoff mechanisms, deterministic-random tradeoff and subspace tradeoff, in the high-SNR scenario. In the following, we'll briefly introduce the two types of tradeoff mechanisms.

5.1 Deterministic-Random Tradeoff

Intuitively, to improve the communication rate, Tx signals are expected to be as random as possible to carry more

information. If sensing performance needs to be considered, then the Tx signals need to be deterministic to ensure stable sensing performance. On the one hand, by comparing Equation (9) with Equation (19), we can see that to achieve the optimal sensing performance, the use of deterministic signals leads to a loss of the DoF, which is $M_{\mathrm{SC}}/2T$. On the other hand, by comparing Equation (13) with Equation (14), we can see that to achieve optimal communication performance, the use of Gaussian signals also causes a certain loss of the sensing DoF, which is $\min\{K, M_{\mathrm{CS}}\}$ at maximum. It is noteworthy that when the ratio of the sensing duration T to M_{SC} is sufficiently large, the deterministic-random tradeoff becomes no longer significant. This is because even if $\mathbf{X}$ is the communication-optimal Gaussian signal, its rows are asymptotically orthogonal when $T/M_{\mathrm{SC}} \to \infty$ so that the signal is asymptotically equivalent to a semi-unitary signal. In addition, if the number of sensing parameters K is far less than the sensing duration T, the loss of the sensing DoF caused by the randomness of Gaussian signals can almost be negligible. In other words, if the load of sensing tasks is relatively light, the communication-optimal Gaussian distribution signals can be directly used for sensing. On the other hand, if T/M_{SC} or $T/\min\{K, M_{\mathrm{CS}}\}$ remains constant, the deterministic-random tradeoff will always exist even if $T \to \infty$.

To further illustrate the significance of the deterministic-random tradeoff in an ISAC system, we can consider such a simple case: $M = 1$ and $T = 1$. In this case, the Tx signal is a scalar, which may be denoted as x. To achieve the optimal sensing performance, $R_X = |x|^2$ needs to be a definite value. Hence, the Tx signal needs to have a constant amplitude. To achieve the maximum communication rate, the empirical distribution of the Tx signal needs to approach the Gaussian distribution. Figure 3 shows the signal constellation diagrams in the two cases. As we can see, there is a communication DoF loss due to constant amplitude for the sensing-optimal Tx signal, thereby weakening its capability of carrying information. At the same time, the constant-modulus feature
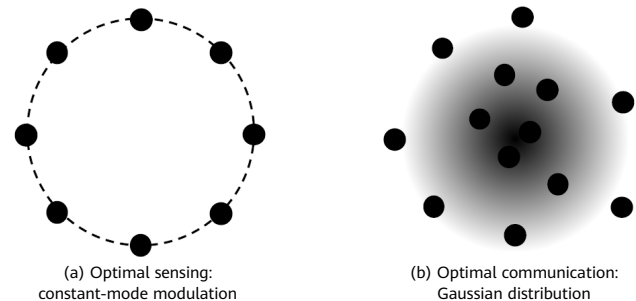

(a) Optimal sensing: constant-mode modulation

(b) Optimal communication: Gaussian distribution

Figure 3 Tx signal constellation diagrams for optimal sensing and communication performance when $M = 1$ and $T = 1$

ensures a constant sensing performance for x in each implementation. If x obeys the Gaussian distribution, its sensing performance fluctuates between implementations, and on average, the performance is not as good as that of a constant-modulus signal.

5.2 Subspace Tradeoff

By comparing Equation (11) with Equation (18), we can see that the column space of the sensing-optimal Tx signal is aligned with $\mathbf{U}_s$, and the column space of the communication-optimal Tx signal is aligned with $\mathbf{U}_c$. Therefore, as described in [7], column spaces of $\mathbf{U}_s$ and $\mathbf{U}_c$ are respectively referred to as the sensing subspace and communication subspace. The lessons learned are that when the design objective of an ISAC system is not to optimize the communication or sensing performance but to achieve some balance between them, we can try to allocate signal power properly between the communication and sensing subspaces. It is further pointed out in [7] that the power allocation may be implemented through statistical covariance shaping, that is, by solving the following multi-objective optimization problem:

$$
\begin{aligned}
&\max_{\tilde{\boldsymbol{R}}_X} \ (1-\alpha)\mathrm{Tr}\left\{\left[\boldsymbol{\Phi}\left(\tilde{\boldsymbol{R}}_X\right)\right]^{-1}\right\} - \alpha \log\left|\boldsymbol{I} + \sigma_c^{-2}\mathbf{H}_c\tilde{\boldsymbol{R}}_X\mathbf{H}_c^H\right| \\
&\text{s.t.} \ \ \mathrm{Tr}\left\{\tilde{\boldsymbol{R}}_X\right\} = P_T, \ \tilde{\boldsymbol{R}}_X \succcurlyeq \boldsymbol{0} \ \tilde{\boldsymbol{R}}_X = \tilde{\boldsymbol{R}}_X^H
\end{aligned}
\tag{21}
$$

where, $\alpha \in [0,1]$ is a parameter for adjusting the tradeoff between the communication and sensing performance.

The significance of the subspace tradeoff is closely related to the coupling strength between communication and sensing subspaces. Intuitively, if the communication user is the sensing target, the column space of $\mathbf{U}_c$ is the same as that of $\mathbf{U}_s$. In this case, there is no subspace tradeoff, and the performance tradeoff is completely determined by the deterministic-random tradeoff. In an extreme case, if the communication and sensing subspaces are orthogonal to each other, Gaussian signals can be transmitted in the communication subspace, and semi-unitary signals can be transmitted in the sensing subspace. This shows that there is no deterministic-random tradeoff in such a situation, and optimal adjustment of the performance tradeoff can be achieved by completely relying on the statistical covariance shaping of Equation (21).

The preceding analysis shows that the communication and sensing performance tradeoff in the ISAC system is essentially a two-fold tradeoff. Specifically, when the ISAC system operational mode transits from the communication-optimal corner point P_{CS} to the sensing-optimal corner point P_{SC}, the ISAC signal transits from completely random (Gaussian distribution) to relatively deterministic (uniform distribution on the Stiefel manifold). On the other hand, the power of the ISAC signal flows from the communication subspace to the sensing subspace, and the coupling strength between the subspaces determines the utilization efficiency of power resources. The higher the coupling strength, the more power can be reused between communication and sensing.

5.3 Guiding Principles for ISAC System Design

Some guiding principles for ISAC system design can be drawn from the theoretical analysis in this paper, as described below:

- **Identification criteria for use of unified waveforms:** The ISAC system generally uses two types of signal design: orthogonal resource allocation signal and unified signal. The former means that communication and sensing signals occupy different orthogonal resource blocks, such as frequency division, time division, space division, and code division. The latter means that all radio resources are reused between communication and sensing, that is, the so-called unified waveform design. In the former scheme, communication and sensing signals do not interfere with each other, and implementation is relatively simple, but resource utilization efficiency is relatively low. In the latter scheme, the design is relatively complex, but resource utilization efficiency is relatively high. The analysis in this paper points out that the resource utilization efficiency in the ISAC system is determined by the coupling strength between the communication and sensing channels. If the coupling strength is relatively high, unified signals should be used to improve efficiency. If the coupling strength is relatively low, orthogonal resource allocation solutions, such as frequency division, time division, space division, and code division, should be used to reduce the design complexity.

- **Identification criteria for use of communication-centric waveforms:** This paper points out that randomness of ISAC signals is conducive to communication, while determinism is conducive to sensing. There have always been no definite guiding principles for the use of waveforms in different application scenarios. The analysis on P_{CS} in this paper shows that when the ratio T/K of the number of signal samples to the number of sensing parameters is relatively large, the communication-optimal Gaussian codebook (which is usually obtained through constellation shaping on high-order quadrature amplitude modulation) may be directly used for sensing. In this case, the implementation cost is relatively low, and the sensing performance loss is relatively small. Conversely, if T/K is relatively small, the ISAC signal needs to be designed to maintain proper randomness for acceptable sensing performance.

- **Design for sensing-centric waveforms:** This paper provides the achievable rate for high-SNR communication at P_{SC}. To achieve this rate, a semi-unitary matrix needs to be uniformly selected as a codebook on the Stiefel manifold induced by the sensing-optimal sample covariance matrix. This scheme can facilitate the sensing-centric ISAC waveform design for improved communication performance. For example, in references [11, 12], the attempt to use metrics such as the MIMO radar antenna and signal frequency to modulate information carried in signals is essentially equivalent to using a permutation matrix to carry communication symbol. Since the permutation matrix is a unitary matrix, its signal sample covariance remains unchanged, and therefore the radar performance remains stable. However, this solution is unable to achieve the optimal performance at P_{SC} because the set of all permutation matrices is a discrete subset of the Stiefel manifold, and the communication performance of the codebook constructed on the permutation matrix is definitely weaker than that of the uniform codebook on the Stiefel manifold.

6 Conclusion

ISAC is a key enabling technology for the next-generation wireless network, and the performance tradeoff between communication and sensing is one of the most important basic theoretical problems arising therefrom. Based on information and estimation theories, this paper deep dives into the performance limits and achievable bounds of communication and sensing and introduces the latest research progress of the author team in the ISAC field and the insights into the tradeoff between communication and sensing performance. In particular, this paper discusses the upper and lower bounds of the sensing CRB at the communication-optimal corner point P_{CS} and provides the signal structure, high-SNR communication capacity, and achievability for optimal sensing at the sensing-optimal corner point P_{SC}. The research in this paper shows that a two-fold-tradeoff is present between the communication and sensing performance in ISAC systems, namely, deterministic-random tradeoff and subspace tradeoff. The disclosure of this tradeoff mechanism is expected to play a guiding role in the ISAC system design.

References

[1] W. Tong and P. Zhu, *6G: The Next Horizon: From Connected People and Things to Connected Intelligence*. Cambridge University Press, 2021.

[2] 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," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, June 2022.

[3] ITU-R WP5D, "IMT framework for 2030 and beyond," ITU, Tech. Rep., 2023.

[4] F. Liu, L. Zheng, Y. Cui, C. Masouros, A. P. Petropulu, H. Griffiths, and Y. Eldar, "Seventy years of radar and communications: The road from separation to integration," *IEEE Signal Process. Mag.*, vol. 40, no. 5, pp. 106-121, July 2023.

[5] Y. F. Xiong, F. Liu, W. J. Yuan, Y. H. Cui, K. Yang, G. Wu, Q. X. Zhang, Z. Q. Wei, Z. Y. Feng, P. Zhang, "Tongxin ganzhi yitihua de xinxi lilun jixian" [Information-theoretic limits of integrated sensing and communications], SCIENTIA SINICA Informationis, 2023. [Online]. Available: https://doi.org/10.1360/SSI-2023-0056

[6] M. Ahmadipour, M. Kobayashi, M. Wigger, and G. Caire, "An information-theoretic approach to joint sensing and communication," *IEEE Trans. Inf. Theory*, pp. 1–1, *early access*, 2022.

[7] Y. Xiong, F. Liu, Y. Cui, W. Yuan, T. X. Han, and G. Caire, "On the fundamental tradeoff of integrated sensing and communications under Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 69, no. 9, pp. 5723-5751, Sept. 2023.

[8] R. Miller and C. Chang, "A modified Cramér-Rao bound and its applications (corresp.)," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 398–400, May 1978.

[9] Y. Xiong, F. Liu, Y. Cui, W. Yuan, and T. X. Han, "Flowing the information from Shannon to Fisher: Towards the fundamental tradeoff in ISAC," in *Proc. IEEE GLOBECOM*, 2022.

[10] F. Liu, Y. Xiong, K. Wan, T. X. Han, and G. Caire, "Deterministic-random tradeoff of integrated sensing and communications: A rate-distortion perspective," in *Proc. IEEE ISIT*, 2023.

[11] E. BouDaher, A. Hassanien, E. Aboutanios, and M. G. Amin, "Towards a dual-function MIMO radar-communication system," in *Proc. 2016 IEEE Radar Conf. (RadarConf)*, Philadelphia, PA, USA, May 2016, pp. 1–6.

[12] T. Huang, N. Shlezinger, X. Xu, Y. Liu, and Y. C. Eldar, "MAJoRCom: A dual-function radar communication system using index modulation," *IEEE Trans. Signal Process.*, vol. 68, pp. 3423–3438, 2020.

Time-Frequency-Space Signal Design of Terahertz Integrated Sensing and Communication

Yongzhi Wu, Chong Han
Shanghai Jiao Tong University

Abstract

The Terahertz (0.1–10 THz) band holds enormous potential for supporting unprecedented data rates and millimeter-level accurate sensing thanks to its ultra-broad bandwidth. Terahertz integrated sensing and communication (ISAC) is viewed as a game-changing technology to realize connected intelligence in the sixth generation (6G) and beyond systems.

In this article, waveform design and receiver algorithms to unleash the full potential of THz ISAC are investigated from the perspectives of the time-frequency (TF) domain, delay-Doppler (DD) domain, and time-frequency-space (TFS) domain.

1 Introduction

With the revolutionary enhancement of wireless data rates and the increasing demand for highly accurate sensing, a promising blueprint is expected to be realized: everything will be sensed, connected, and intelligent [1]. The Terahertz (THz) band and integrated sensing and communication (ISAC) are two descriptive technologies for achieving such a vision in sixth generation (6G) and beyond wireless systems, by exploring new spectrum and new sensing functions, respectively. Two forces drive together to motivate the study of integration of these two technologies [2]. It can be foreseen that wireless links with Terabit-per-second (Tbps) will come true in the future intelligent information society. Moreover, the ultra-broad bandwidth in the THz band also provides huge potential to realize ultra-accurate sensing, e.g., millimeter-level and millidegree-level accuracy.

Thanks to the abundant continuous bandwidth in the THz band [3], THz integrated sensing and communication (THz ISAC) can enable simultaneously transmitting billions of data streams and ultra-accurate sensing by sharing the spectrum, hardware, and signal processing modules. In addition to enhancing the performance and functionality, THz ISAC is able to effectively reduce the hardware cost, and improve the spectral and energy efficiency [4]. In light of these, THz ISAC is envisioned to provide high quality of experience (QoE) for various future services and applications demanding ultra-high data rates and ultra-high resolution sensing.

Waveform design plays a key role in ISAC systems, which mainly focuses on designing a commonly used waveform to achieve the integration gain [4]. Straightforwardly, sensing and communication can be directly scheduled on the non-overlapped wireless resources, such as time-division and frequency-division ISAC. However, the fully unified waveform design can achieve higher spectral and energy efficiency. In this article, we focus on the dual-functional waveform design of THz ISAC, which is also the fundamental of physical-layer transmission. Specifically, we discuss the waveform design for THz ISAC from the perspectives of the time-frequency (TF) domain, delay-Doppler (DD) domain, and time-frequency-space (TFS) domain. Meanwhile, generic receive sensing algorithms for single and multiple targets are studied.

2 Waveform Design

Sensing and communication can fully share the hardware and signal processing modules when jointly designing the ISAC transmit waveform, which thereby reduces power consumption and signal processing complexity. Existing communication-centric ISAC waveforms can be classified into two classes, i.e., working in the TF domain or the DD domain. In addition, since THz ISAC systems are usually enabled with large-scale antenna arrays, the compatibility between ultra-massive multiple-input multiple-output (UM-MIMO) and transmit waveform needs to be considered, i.e., TFS domain design.

2.1 Time-Frequency Domain Processing

As used in 4G and 5G standards, orthogonal frequency division multiplexing (OFDM) is a promising candidate for ISAC, which has attracted tremendous attention [5–9]. A baseband OFDM transmit signal is given by

$$s(t) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} X[m,n] \text{rect}\left(\frac{t-nT_o}{T_o}\right) e^{j2\pi m \Delta f(t-T_{cp}-nT_o)}, \quad (1)$$

where t denotes the time instant, M and N stand for the number of subcarriers and symbols, respectively. $X[m,n](m = 0, 1, \cdots, M-1; n = 0, 1, \cdots, N-1)$ describes the information data symbols and the $\text{rect}(t)$ denotes the rectangular pulse. Δf refers to the subcarrier spacing. The symbol duration T equals $\frac{1}{\Delta f}$ and the total symbol duration is expressed as $T_o = T + T_{cp}$ with the cyclic prefix (CP) duration of T_{cp}. Authors in [6] propose an OFDM ISAC framework with the two-dimensional (2D) discrete Fourier transform (DFT) method, which can be implemented with computationally-efficient fast Fourier transform (FFT) algorithms. Hereby, a multiple signal classification (MUSIC)-based method is also developed to realize OFDM target sensing, which achieves higher resolution while increasing computational complexity and requiring a high signal-to-noise ratio (SNR). However, OFDM waveforms inherently suffer from a high peak-to-average power ratio (PAPR) and might induce undesirable clipping distortions in THz power amplifiers (PAs) and degrade the energy efficiency [10].

While maintaining the advantages of frequency-domain processing, some single-carrier waveforms, such as DFT-

s-OFDM (aka SC-FDMA) and its variants, are regarded as more potential candidate waveforms for THz ISAC, thanks to their lower PAPR compared with multi-carrier waveforms. Since DFT-s-OFDM introduces a DFT spreading operation on information symbols before OFDM modulation, it presents Gaussian randomness in the frequency domain [11]. To follow the classical OFDM sensing algorithms, we need to modify the element-wise division in [6] into the conjugated Hadamard product or employ the constant-enveloped pilot signal [12].

2.2 Delay-Doppler Domain Processing

Being proportional to the carrier frequency, the Doppler spread effect becomes even stricter in the THz band, especially in high-mobility scenarios. Nevertheless, the Doppler shifts are difficult to handle in the TF domain and may degrade the link performance for high-mobility communications. Recently, an orthogonal time-frequency-space (OTFS) modulation is developed to deal with Doppler effects and conveniently accommodate the channel dynamics in the DD domain [13], since it transforms a time-variant channel into a 2D quasi-time-invariant channel in the DD domain. A baseband OTFS transmit signal is expressed as

$$s(t) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} X^{\mathrm{TF}}[m,n]\mathrm{rect}\left(\frac{t-nT}{T}\right)e^{j2\pi m\Delta f(t-nT)}, \quad (2)$$

where $X^{\mathrm{TF}}[m,n]$ refers to the TF domain signal generated by applying the inverse symplectic finite Fourier transform (ISFFT) [14]:

$$X^{\mathrm{TF}}[m,n] = \frac{1}{\sqrt{MN}}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1}X[l,k]e^{j2\pi\left(\frac{nk}{N}-\frac{ml}{M}\right)}. \quad (3)$$

While OFDM inserts a CP for each symbol, OTFS uses a CP for each frame. The effectiveness of OTFS for ISAC is validated in [15], which requires constructing the DD domain channel matrix in the sensing estimator that is not highly efficient for THz ISAC.

Behaving like OFDM, the time-domain transmit samples

of OTFS are equivalent to the output of inverse DFT (IDFT) of the information symbols in the DD domain. Therefore, the PAPR of OTFS is still not satisfactory for THz PAs. We proposed a DFT-spread OTFS (DFT-s-OTFS) waveform in [16] to reduce the PAPR of OTFS, by developing a DFT precoding operation on the information symbols along the Doppler axis. Meanwhile, based on an exact DD domain input-output (I/O) relation, a two-phase estimation method with DD domain processing is developed and can be implemented by the 2D circular shift and the FFT algorithm.

2.3 Time-Frequency-Space Domain Processing

While the DD domain processing provides stronger robustness against Doppler effects, it increases the processing complexity at the receiver side. Moreover, to compensate for such severe path loss in the THz band, UM-MIMO antenna arrays with beamforming are used to generate directional beams for THz ISAC systems. The high complexity of data detection for MIMO-OTFS constitutes a serious problem. Despite the PAPR issue, OFDM is still a potential waveform in the THz band, since it has good compatibility with UM-MIMO and enables flexible TF-domain resource allocation among multiple users [2]. MIMO-OFDM systems can perform subcarrier-wise processing on the TFS domain signal, which decouples the dimensions of the frequency domain and spatial domain. With the usage of hybrid beamforming and OFDM modulation, the time-domain transmit vector is given by [17]:

$$\mathbf{s}_n(t) = \sum_{m=0}^{M-1}\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}[m]\mathbf{x}_{\mathrm{S}}[m,n]e^{j2\pi m\Delta ft}, \quad (4)$$

where $\mathbf{x}_{\mathrm{S}}[m,n] \in \mathbb{C}^{N_s\times 1}$ denotes the data symbols generated from N_s data streams at the mth subcarrier and nth symbol, $\mathbf{F}_{\mathrm{BB}}[m] \in \mathbb{C}^{N_{\mathrm{RF}}\times N_s}$ and $\mathbf{F}_{\mathrm{RF}} \in \mathbb{C}^{N_t\times N_{\mathrm{RF}}}$ refer to the digital precoder and the analog precoder, respectively. N_{RF} stands for the number of transmit radio frequency (RF) chains and N_t describes the number of transmit antennas.

Table 1 Signal models in matrix forms of the TF domain, DD domain, and TFS domain

Domain	CPs	Time-Domain Transmit Signal	I/O Relation	Approximated I/O Relation
TF	N CPs	$\mathbf{S} = \mathbf{F}_M^H \mathbf{X}^{\mathrm{TF}}$	$\mathbf{Y}^{\mathrm{TF}} = \alpha \mathbf{F}_M \mathrm{vec}^{-1}(\boldsymbol{\Delta}_1^{(\nu)} \mathrm{vec}(\mathbf{F}_M^H \mathbf{b}_\tau \mathbf{X}^{\mathrm{TF}}))$	$\mathbf{Y}^{\mathrm{TF}} \approx \alpha \mathbf{X}^{\mathrm{TF}} \odot \boldsymbol{\Psi}_{\tau,\nu}$
DD	1 CP	$\mathbf{S} = \mathbf{X}^{\mathrm{DD}} \mathbf{F}_N^H$	$\mathbf{Y}^{\mathrm{DD}} = \alpha \mathrm{vec}^{-1}(\boldsymbol{\Delta}_2^{(\nu)} \boldsymbol{\Pi}_{MN}^{l_\tau} \mathrm{vec}(\mathbf{F}_M^H \mathbf{B}_\tau \mathbf{F}_M \mathbf{X}^{\mathrm{DD}} \mathbf{F}_N^H)) \mathbf{F}_N$	$\mathbf{Y}^{\mathrm{DD}} \approx \alpha \boldsymbol{\Pi}_M^{l_\tau} \mathbf{X}^{\mathrm{DD}} \boldsymbol{\Pi}_N^{k_\nu}$
TFS	N CPs	$\mathbf{x}^{\mathrm{TF}}[m,n] = \mathbf{F}_{\mathrm{RF}} \mathbf{F}_{\mathrm{BB}}[m] \mathbf{x}_{\mathrm{S}}[m,n],$ $\mathbf{S}_u = \mathbf{F}_M^H \mathbf{X}_u^{\mathrm{TF}}, X_{u,m,n}^{\mathrm{TF}} = x_u^{\mathrm{TF}}[m,n]$	$\hat{\mathbf{x}}[m,n] = \mathbf{W}_{\mathrm{RF}}^H \mathbf{a}_r(\theta,\phi) \mathbf{a}_t^T(\theta,\phi) \mathbf{x}^{\mathrm{TF}}[m,n],$ $\mathbf{Y}_u^{\mathrm{TF}} = \alpha \hat{\mathbf{X}}_u \odot \boldsymbol{\Psi}_{\tau,\nu}, \hat{X}_{u,m,n} = \hat{x}_u[m,n]$	

Notations of Table 1: $\mathbf{F}_M \in \mathbb{C}^{M \times M}$ denotes the DFT matrix. $\mathbf{X}^{\mathrm{TF}}$ and $\mathbf{X}^{\mathrm{DD}}$ refer to the TF- and DD-domain transmit signals in matrix forms, respectively. $\mathbf{Y}^{\mathrm{TF}}$ and $\mathbf{Y}^{\mathrm{DD}}$ are respectively the TF- and DD-domain received signals. $\boldsymbol{\Delta}_1^{(\nu)} = \mathrm{diag}(\mathrm{vec}(\mathbf{V}))$ with $\mathbf{V}_{m,n} = e^{j2\pi\nu(nT_o + T_{\mathrm{cp}} + \frac{m}{M}T)}$ and $\boldsymbol{\Delta}_2^{(\nu)} = \mathrm{diag}\{\delta^0, \delta^1, \cdots, \delta^{MN-1}\}$ with $\delta = e^{j2\pi\nu\frac{T}{M}}$. $\boldsymbol{\Pi}$ stands for the permutation matrix (forward cyclic shift). $\boldsymbol{\Psi}_{\tau,\nu} = \boldsymbol{\Psi}_\tau \boldsymbol{\Psi}_\nu^T$ with $\boldsymbol{\Psi}_\tau = [e^{-j2\pi0\Delta f\tau}, e^{-j2\pi1\Delta f\tau}, \cdots, e^{-j2\pi(M-1)\Delta f\tau}]^T$ and $\boldsymbol{\Psi}_\nu = [e^{j2\pi0T_o\nu}, e^{j2\pi1T_o\nu}, \cdots, e^{j2\pi(N-1)T_o\nu}]^T$. $\mathbf{b}_\tau = \mathrm{diag}(\boldsymbol{\Psi}_\tau)$ and $\mathbf{B}_\tau = \mathrm{diag}\{b^0, b^1, \cdots, b^{M-1}\}$ with $b = e^{j2\pi\left(\frac{l_\tau}{M} - \frac{\tau}{T}\right)}$, $l_\tau = \lceil \frac{M\tau}{T} \rceil$, and $k_\nu = \lceil \frac{N\nu}{\Delta f} \rceil$. $\mathbf{W}_{\mathrm{RF}}$ describes the analog combiner and $\mathbf{a}(\theta,\phi)$ represents the array steering vector. $m = 0, 1, \cdots, M-1$, $n = 0, 1, \cdots, N-1$, and $u = 1, 2, \cdots, N_{\mathrm{RF}}$.

As shown in Table 1, we can summarize the signal models of the TF domain, DD domain, and TFS domain. In the TF domain, when there is no inter-carrier interference, the I/O relation can be approximated as a Hadamard product, i.e., a single-tap model. Therefore, with one-tap frequency-domain equalization, data detection can be performed with high computational efficiency. In the DD domain, the I/O relation can be approximated as a 2D circular shift, which relates to the delay and Doppler taps. With the derived relations, we can estimate the sensing parameters, given the knowledge of transmit signals and received signals in different domains.

3 Receiver Design

In this section, we discuss the sensing algorithm design in the TF domain, DD domain, and TFS domain. We first focus on the single-target estimation and then consider the multi-target estimation.

3.1 Single-Target Estimation

Without loss of generality, the vectorized I/O relation with additive white Gaussian noise (AWGN) in a certain domain (TF, DD or TFS) can be expressed as:

$$\mathbf{y} = \alpha \mathbf{H}(\tau, \nu)\mathbf{x} + \mathbf{w} \tag{5}$$

where α refers to the complex channel coefficient, $\mathbf{H}(\tau, \nu)$ denotes the channel matrix related to the delay τ and Doppler ν of the sensing channel, and $\mathbf{w}$ stands for the noise vector. The sensing algorithm aims to estimate the delay and Doppler parameters, given the knowledge of $\mathbf{x}$ and $\mathbf{Y}$. The maximum likelihood (ML) estimator is given by:

$$(\hat{\tau}, \hat{\nu}) = \arg \min_{(\alpha, \tau, \nu)} \|\mathbf{y} - \alpha \mathbf{H}(\tau, \nu)\mathbf{x}\|^2. \tag{6}$$

By deriving the partial derivative with respect to the channel coefficient α and obtaining an estimate of α, this problem boils down to the maximization problem:

$$(\hat{\tau}, \hat{\nu}) = \arg \max_{(\tau, \nu)} \frac{|(\mathbf{H}(\tau, \nu)\mathbf{x})^H \mathbf{y}|^2}{\mathbf{x}^H \mathbf{H}(\tau, \nu)^H \mathbf{H}(\tau, \nu)\mathbf{x}}. \tag{7}$$

With different I/O relations in different domains, this ML estimator turns into different algorithms.

To reduce the search complexity, we can design a two-stage method for all cases of TF, DD, and TFS domains, i.e., an on-grid estimation in the first stage and an off-grid search in the second stage. The on-grid estimation is performed on the discretized grid $\boldsymbol{\Gamma}_1 = \{(\frac{m}{M\Delta f}, \frac{n}{NT'}), m = 0, \cdots, M-1; n = -\frac{N}{2}, \cdots, \frac{N}{2} - 1\}$ (T' equals T_o for the TF domain and T for the DD domain). The region $\boldsymbol{\Gamma}_2$ of the off-grid search is limited to the neighborhood of the on-grid estimation result. In this case, the fractional part of the delay and Doppler parameters can be estimated by a 2D GSS method.

Table 2 Sensing algorithms of the TF domain, DD domain, and TFS domain processing

Domain	Stage		Estimator	Complexity
TF	On-grid	2D DFT:	$(\hat{m}, \hat{n}) = \arg \max_{(m,n)\in\mathbf{\Gamma}_1} \left\| g_{m,[n]_N} \right\|^2,$ $\mathbf{g} = \mathbf{F}_M^H \left(\mathrm{conj}(\mathbf{X}^{\mathrm{TF}}) \odot \mathbf{Y}^{\mathrm{TF}} \right) \mathbf{F}_N$	$\mathcal{O}(MN \log(MN))$
	Off-grid	2D GSS:	$(\hat{\tau}, \hat{\nu}) = \arg \max_{(\tau,\nu)\in\mathbf{\Gamma}_2} \left\| \mathrm{vec}(\mathbf{F}_M \mathrm{vec}^{-1}(\mathbf{\Delta}_1^{(\nu)} \mathrm{vec}(\mathbf{F}_M^H \mathbf{b}_\tau \mathbf{X}^{\mathrm{TF}})))^H \mathbf{y}^{\mathrm{TF}} \right\|^2$	$\mathcal{O}(N_{\mathrm{iter}} MN \log(MN))$
DD	On-grid	2D circular shift:	$(\hat{l}, \hat{k}) = \arg \max_{(l,k)\in\mathbf{\Gamma}_1} \left\| \mathrm{vec}(\mathbf{\Pi}_M^l \mathbf{X} \mathbf{\Pi}_N^k)^H \mathbf{y}^{\mathrm{DD}} \right\|^2$	$\mathcal{O}((MN)^2)$
	Off-grid	2D GSS:	$(\hat{\tau}, \hat{\nu}) = \arg \max_{(\tau,\nu)\in\mathbf{\Gamma}_2} \left\| \mathrm{vec}(\mathrm{vec}^{-1}(\mathbf{\Delta}_2^{(\nu)} \mathbf{\Pi}_{MN}^{l_\tau} \mathrm{vec}(\mathbf{F}_M^H \mathbf{B}_\tau \mathbf{F}_M \mathbf{X}^{\mathrm{DD}} \mathbf{F}_N^H)) \mathbf{F}_N)^H \mathbf{y}^{\mathrm{DD}} \right\|^2$	$\mathcal{O}(N_{\mathrm{iter}} MN \log(MN))$
TFS	On-grid	Sum-DFT:	$(\hat{m}, \hat{n}) = \arg \max_{(m,n)\in\mathbf{\Gamma}_1} \left\| g_{m,[n]_N} \right\|^2,$ $\mathbf{g} = \sum_{u=1}^{N_{\mathrm{RF}}} \mathbf{F}_M^H \left(\mathrm{conj}(\hat{\mathbf{X}}_u) \odot \mathbf{Y}_u^{\mathrm{TF}} \right) \mathbf{F}_N$	$\mathcal{O}(N_{\mathrm{RF}} MN \log(MN))$
	Off-grid	2D GSS:	$(\hat{\tau}, \hat{\nu}) = \arg \max_{(\tau,\nu)\in\mathbf{\Gamma}_2} \left\| \sum_{u=1}^{N_{\mathrm{RF}}} \mathrm{Tr} \left(\left(\mathbf{\Psi}_{\tau,\nu} \odot \hat{\mathbf{X}}_u \right)^H \mathbf{Y}_u^{\mathrm{TF}} \right) \right\|^2$	$\mathcal{O}(N_{\mathrm{iter}} N_{\mathrm{RF}} MN)$

Notations: N_{iter} denotes the iterations of golden section search (GSS).

As summarized in Table 2, the sensing estimators in different domains can be further transformed into different algorithms by utilizing corresponding I/O relations and approximated I/O relations. In the TF domain, we can use the 2D DFT algorithm to perform the on-grid estimation. In the DD domain, we can conduct the 2D circular shift for on-grid search. In the TFS domain, we propose a sum-DFT (S-DFT) method to estimate the sensing parameters. Meanwhile, the computational complexity of the sensing algorithms is also given in Table 2.

3.2 Multi-Target Estimation

With the sensing algorithms for single-target estimation, we can obtain the solutions for multi-target estimation by directly removing the interference signal among different targets. For the set of 3 unknown vectors with $3P$ parameters, $(\boldsymbol{\alpha}, \boldsymbol{\tau}, \boldsymbol{\nu})$, the maximum likelihood estimator of these parameters is given by:

$$(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\tau}}, \hat{\boldsymbol{\nu}}) = \arg \min_{(\boldsymbol{\alpha}, \boldsymbol{\tau}, \boldsymbol{\nu})} \left\| \mathbf{y} - \sum_{i=1}^{P} \alpha_i \mathbf{H}_i(\tau_i, \nu_i) \mathbf{x} \right\|^2. \quad (8)$$

We perform the interference cancellation mechanism [16], i.e., when estimating the parameters of ith target, eliminating the interference signal from the $(i-1)$ targets that have been estimated. In this case, the estimator for the ith target becomes

$$(\hat{\tau}_i, \hat{\nu}_i) = \arg \max_{(\tau_i, \nu_i)} \left| (\mathbf{H}_i \mathbf{x})^H \left(\mathbf{y} - \sum_{j=1}^{i-1} \hat{\alpha}_j \mathbf{H}_j(\hat{\tau}_j, \hat{\nu}_j) \mathbf{x} \right) \right|^2, \quad (9)$$

and the estimated channel coefficient α_i is given by:

$$\hat{\alpha}_i = \frac{(\mathbf{H}_i(\hat{\tau}_i, \hat{\nu}_i) \mathbf{x})^H \left(\mathbf{y} - \sum_{j=1}^{i-1} \hat{\alpha}_j \mathbf{H}_j(\hat{\tau}_j, \hat{\nu}_j) \mathbf{x} \right)}{\mathbf{x}^H \mathbf{H}_i^H(\hat{\tau}_i, \hat{\nu}_i) \mathbf{H}_i(\hat{\tau}_i, \hat{\nu}_i) \mathbf{x}}. \quad (10)$$

4 Simulation Results

In this section, we provide some simulation results to quantitatively illustrate the performance of the aforementioned waveforms with signal processing in the TF and DD domain. In Figure 1, the transceiver block diagrams of these waveforms are described, including OFDM, DFT-s-OFDM, OTFS, and DFT-s-OTFS. To list the key simulation parameters, the carrier frequency equals 0.3 THz. The subcarrier spacing is set as 1.92 MHz. The 4-QAM modulation scheme is employed. The PAPRs of these waveforms are evaluated by conducting 4-times oversampling on the discrete-time baseband signal. In Figure 2, we demonstrate that the PAPR of DFT-s-OTFS approaches that of DFT-s-OFDM, while their PAPRs are approximately 3 dB lower than those of OTFS and OFDM.

Next, we set a sensing target with a distance of 10 m and a relative velocity of 20 km/h. As shown in Figure 3, we learn that all of these waveforms can achieve millimeter-level
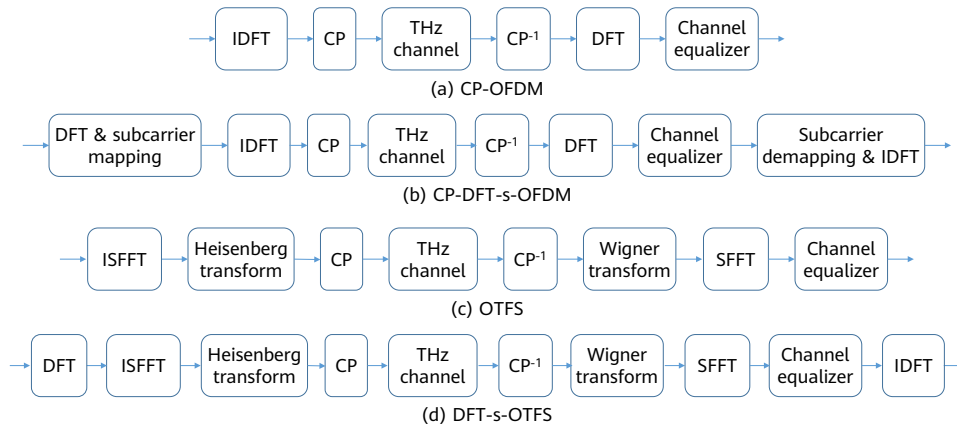
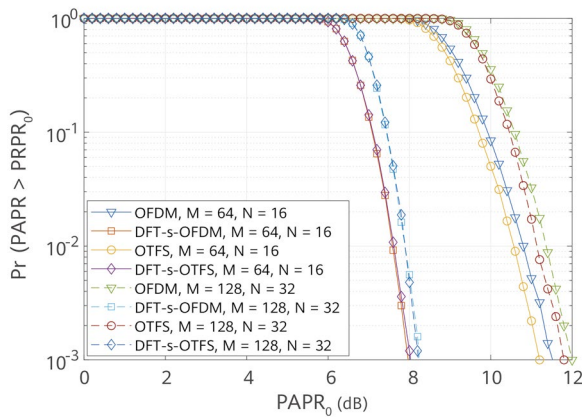**Figure 1** Waveform transceiver block diagrams

Figure 2 PAPR comparison for the transmit signals of OFDM, DFT-s-OFDM, OTFS and DFT-s-OTFS with respect to different numbers of subcarriers M and symbols N.
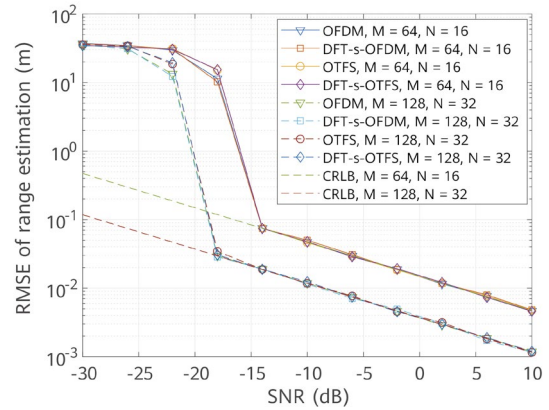
Figure 3 The root mean square errors (RMSEs) of target range estimation are compared among various waveforms with respect to different numbers of subcarriers M and symbols N. The RMSEs can achieve the Cramér-Rao lower bound (CRLB) when the SNR is above −15 dB.

range estimation accuracy. Moreover, the range estimation error can be reduced by increasing the subcarrier number from 64 to 128 and the symbol number from 16 to 32.

5 Conclusion

THz ISAC is one of the key promising enablers to meet demands for next-generation wireless systems. In this paper, waveform design and sensing algorithms are investigated for THz ISAC from the perspectives of TF, DD, and TFS domains. In a certain domain (TF, DD, or TFS), the sensing algorithm can be implemented efficiently with the derived I/O relation. Simulation results indicate that THz ISAC can achieve millimeter-level sensing accuracy with both TF and DD domain processing.

In summary, in low-mobility scenarios, OFDM is still an excellent candidate for THz ISAC, and achieves good sensing and communication performance, since it enables good compatibility with UM-MIMO, flexible multi-user scheduling, and resource allocation. In high-mobility scenarios, DFT-s-OTFS may provide stronger robustness against severe Doppler effects, while it comes at the price of increased detection complexity. For energy-constrained links, DFT-s-OFDM and DFT-s-OTFS are shown to realize high energy efficiency due to their low PAPR characteristics.

References

[1] W. Tong and P. Zhu, "6G: The next horizon: From connected people and things to connected intelligence," Cambridge University Press, 2021.

[2] C. Han, Y. Wu, Z. Chen, Y. Chen, and G. Wang, "THz ISAC: A physical-layer perspective of terahertz integrated sensing and communication," IEEE Communications Magazine, to appear, 2023.

[3] Z. Chen, C. Han, Y. Wu, L. Li, C. Huang, Z. Zhang, G. Wang, and W. Tong, "Terahertz wireless communications for 2030 and beyond: A cutting-edge frontier," IEEE Communications Magazine, vol. 59, no. 11, pp. 66–72, 2021.

[4] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," IEEE Journal on Selected Areas in Communications, vol. 40, no. 6, pp. 1728–1767, 2022.

[5] C. R. Berger, B. Demissie, J. Heckenbach, P. Willett, and S. Zhou, "Signal processing for passive radar using OFDM waveforms," IEEE Journal of Selected Topics in Signal Processing, vol. 4, no. 1, pp. 226–238, 2010.

[6] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for fusion of wireless communications and radar sensing," Proceedings of the IEEE, vol. 99, no. 7, pp. 1236–1259, 2011.

[7] J. Johnston, L. Venturino, E. Grossi, M. Lops, and X. Wang, "MIMO OFDM dual-function radar-communication under error rate and beampattern constraints," IEEE Journal on Selected Areas in Communications, vol. 40, no. 6, pp. 1951–1964, 2022.

[8] J. A. Zhang, X. Huang, Y. J. Guo, J. Yuan, and R. W. Heath, "Multibeam for joint communication and radar sensing using steerable analog antenna arrays," IEEE Transactions on Vehicular Technology, vol. 68, no. 1, pp. 671–685, 2019.

[9] K. Wu, J. A. Zhang, X. Huang, and Y. J. Guo, "Integrating low-complexity and flexible sensing into communication systems," IEEE Journal on Selected Areas in Communications, vol. 40, no. 6, pp. 1873–1889, 2022.

[10] T. Mao, J. Chen, Q. Wang, C. Han, Z. Wang, and G. K. Karagiannidis, "Waveform design for joint sensing and communications in millimeter-wave and low terahertz bands," IEEE Transactions on Communications, vol. 70, no. 10, pp. 7023–7039, 2022.

[11] K. Wu, J. A. Zhang, X. Huang, and Y. J. Guo, "Integrating low-complexity and flexible sensing into communication systems," IEEE Journal on Selected Areas in Communications, vol. 40, no. 6, pp. 1873–1889, 2022.

[12] Y. Wu, F. Lemic, C. Han, and Z. Chen, "Sensing integrated DFT-spread OFDM waveform and deep learning-powered receiver design for terahertz integrated sensing and communication systems," IEEE Transactions on Communications, vol. 71, no. 1, pp. 595–610, 2023.

[13] Z. Wei, W. Yuan, S. Li, J. Yuan, G. Bharatula, R. Hadani, and L. Hanzo, "Orthogonal time-frequency space modulation: A promising next-generation waveform," IEEE Wireless Communications, vol. 28, no. 4, pp. 136–144, 2021.

[14] P. Raviteja *et al.,* "Interference cancellation and iterative detection for orthogonal time frequency space modulation," IEEE Transactions on Wireless Communications, vol. 17, no. 10, pp. 6501–6515, 2018.

[15] L. Gaudio *et al.,* "On the effectiveness of OTFS for joint radar parameter estimation and communication," IEEE Transactions on Wireless Communications, vol. 19, no. 9, pp. 5951–5965, 2020.

[16] Y. Wu, C. Han, and Z. Chen, "DFT-spread orthogonal time frequency space system with superimposed pilots for terahertz integrated sensing and communication," IEEE Transactions on Wireless Communications, 2023.

[17] Y. Wu and C. Han, "Time-frequency-space signal design with dynamic subarray for terahertz integrated sensing and communication," in Proc. of IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), to appear, 2023.
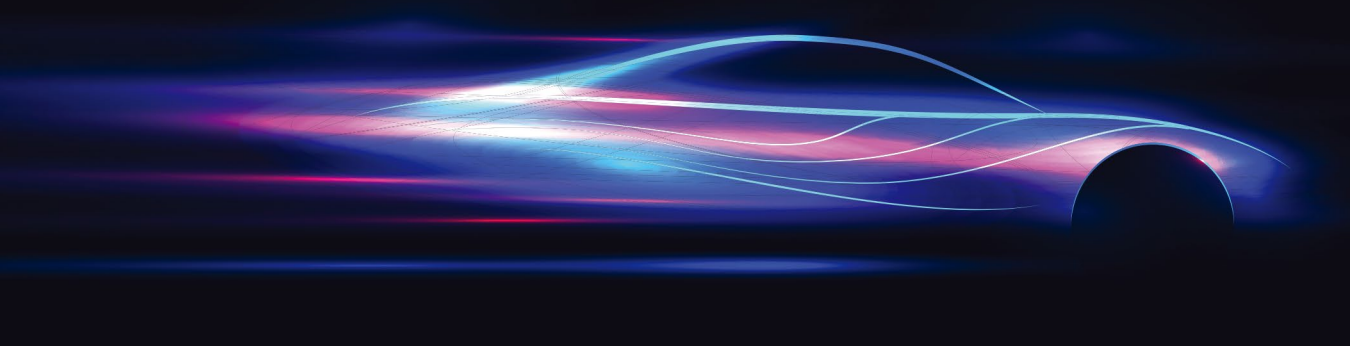
Multi-Base Station Cooperative Sensing with AI-aided Tracking

Elia Favarelli [1], Elisabetta Matricardi [1], Lorenzo Pucci [1], Enrico Paolini [1], Wen Xu [2], Andrea Giorgetti [1]

[1] Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi" (DEI), University of Bologna, and the Wireless Communications Laboratory (WiLab), CNIT, Italy

[2] Advanced Wireless Technology Laboratory, Munich Research Center, Huawei Technologies Duesseldorf GmbH, Munich, Germany

Abstract

In this work, we investigate the performance of a sensing and communication network consisting of multiple base stations (BSs) that cooperate through a fusion center (FC) to exchange information about the sensed environment while concurrently establishing communication links with a set of user equipment (UE) devices. Each BS within the network operates as a monostatic sensing system, enabling comprehensive scanning of the monitored area and generating range-angle maps that provide information regarding the position of a group of heterogeneous objects. The acquired maps are subsequently fused in the FC. Then, a convolutional neural network (CNN) is employed to infer the category of the targets, e.g., pedestrians or vehicles, and such information is exploited by an adaptive clustering algorithm to group the detections originating from the same target more effectively. Finally, two multi-target tracking algorithms, the probability hypothesis density (PHD) filter and multi-Bernoulli mixture (MBM) filter, are applied to estimate the state of the targets. Numerical results demonstrated that our framework could provide remarkable sensing performance, achieving an optimal sub-pattern assignment (OSPA) less than 60 cm, while keeping communication services to UEs with a reduction of the communication capacity in the order of 10% to 20%. The impact of the number of BSs engaged in sensing is also examined, and we show that in the specific case study, 3 BSs ensure a localization error below 1 m.

Keywords

integrated sensing and communication, tracking, orthogonal frequency division multiplexing, millimeter-wave, artificial intelligence, convolutional neural network

1 Introduction

The forthcoming generation of mobile radio networks is poised to offer a range of emerging functionalities, including innovative services. Notably, the ability to perform effective sensing using radio frequency (RF) signals has become feasible due to the evolution toward larger antenna arrays, namely massive multiple-input multiple-output (mMIMO), and higher frequency bands [1, 2]. The joint sensing and communication (JSC) approach leverages existing communication infrastructure to provide sensing capabilities, offering advantages such as reduced costs and improved spectral and energy efficiency when compared to dedicated spectrum- and transceiver-dependent systems like radar [3]. This convergence of sensing and communication systems envisioned for future networks will enable ubiquitous sensing services that rely on capturing reflections from non-collaborative objects, thus playing a critical role, e.g., in intelligent vehicular networks [4]. Furthermore, the growing interest in sensing stems from its potential to support various applications, such as traffic monitoring, autonomous driving, safety in industrial environments, and environmental mapping [5, 6].

The advent of mMIMO technology in millimeter-wave (mmWave) bands facilitates the detection, tracking, and precise localization of pedestrians, vehicles, drones, and other moving objects in real-time scenarios [7]. This enables the acquisition of range profiles of targets, a kind of target fingerprint or signature [9], as scatterers in complex objects may be resolved into different range cells. At the same time, the enormous advancement of artificial intelligence (AI), and particularly image identification, has generated a vast and solid portfolio of solutions that could also be exploited in the field of integrated sensing and communication (ISAC) [8–10].

This work aims to investigate the possibility of using multisensor fusion techniques combined with multi-target tracking algorithms, to exploit range-angle maps obtained through a set of cooperating BSs with monostatic sensing capability and orthogonal frequency division multiplexing (OFDM) signals. The main contributions can be summarized as follows:

- We propose a soft map fusion strategy based on range-angle maps obtained at each BS.
- We present an AI-based approach to infer the target category that is then exploited by an adaptive clustering methodology capable of managing point-like and extended targets.

- The adaptive clustering is then combined with tracking algorithms to perform target state estimation and prediction. Two different tracking algorithms, the probability hypothesis density (PHD) and multi-Bernoulli mixture (MBM) filter, are compared.
- We propose the optimal sub-pattern assignment (OSPA) metric and aggregate downlink capacity to evaluate the sensing and communication capabilities.
- Finally, we investigate the impact of the number of cooperative BSs performing sensing on the localization and communication performance.

In this work, capital and lowercase boldface letters represent matrices and vectors, respectively; $\mathbf{D}_{q,t}$ stands for a matrix dependent on indexes q and t, while $\mathbf{v}_{t,p}$ represents the pth column selected by the matrix $\mathbf{V}_t$. $\mathbf{I}_n$ is the $n \times n$ identity matrix; $\|\cdot\|_p$ stands for the p-norm; $|\cdot|$ represents the cardinality of a set; $\delta(\cdot)$ is the Dirac delta function; $\lfloor \cdot \rceil$ represents the round operator; $(\cdot)^c$ stands for conjugate; $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma})$ denotes a zero-mean circularly symmetric complex Gaussian random vector with covariance $\boldsymbol{\Sigma}$; and $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the real-valued Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$.

The rest of the paper is organized as follows. Section 2 presents the JSC model. Section 3 describes the data fusion strategy, target identification methodology, clustering scheme, and tracking algorithms. System performance is evaluated in Section 4, and conclusions are drawn in Section 5.

2 System Model

This work considers a JSC network, and a scenario, like the one portrayed in Figure 1. In particular, the considered system consists of several monostatic JSC BSs transmitting OFDM signals at mmWave using mMIMO technology. Each of these BSs is connected to the FC via backhaul; the FC allows them to cooperate in performing the detection and tracking of targets in the surveillance area. As shown later, the sensing task is accomplished through range-Doppler maps that each BS can generate by scanning the environment using a dedicated sensing beam. Moreover, to ensure communication functionality, each BS scans the environment for sensing and communicates with UEs in its respective cell using the same time-frequency resources via multiple beams. To keep interference among the sensing beams of different BSs at a negligible level, we consider the proper use of frequency division (FD) or time division (TD) through coordination.
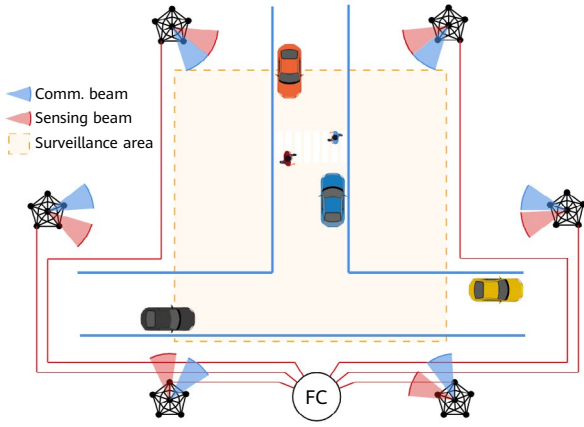
Figure 1 An urban scenario with 6 monostatic JSC base stations (BSs) aiming at monitoring pedestrians (point-like targets) and vehicles (extended targets) in a surveillance area. BSs communicate with their user equipment (UE) while simultaneously sensing the surrounding environment via dedicated sensing beams. The fusion center (FC) collects measurements from the BSs via the backhaul network, fuses them to create likelihood maps, and performs detection, target identification, and multiple target tracking.

Each monostatic BS is equipped with two separate uniform linear arrays (ULAs), one for transmission and one for reception, with N_{T} and N_{R} antennas respectively, and both with a half-wavelength separation between the elements. In particular, the transmitted waveform is used for communication and sensing, while the sensing receiver (Rx) only collects backscattered signals. More specifically, considering the downlink communication toward UEs, each BS transmits frames consisting of M OFDM symbols and K subcarriers, and the same signals are simultaneously used to sense the environment. A multi-beam radiation pattern is used to split the power between sensing and communication by exploiting spatial diversity, as explained later. In particular, each BS uses a communication beam for the UE while steering a sensing beam scanning the environment within the angular interval $[-\Theta_0, \Theta_0]$ with steps $\Delta\Theta$. In each sensing direction, a subset $M_{\mathrm{s}} < M$ of OFDM symbols is collected by the Rx.

The OFDM time-frequency grid containing the transmitted (complex) symbol for each sensing direction can be represented by a matrix $\mathbf{X}_{\mathrm{s}} \in \mathbb{C}^{K \times M_{\mathrm{s}}}$ with elements $x_k^{(m)}$, where k is the subcarrier index and m is the OFDM symbol (or time) index.

Starting from this grid, a precoding operation is performed on its elements with the beamformer $\mathbf{w}_{\mathrm{T}} \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$ to map each complex symbol to each antenna and obtain the vector of the transmitted symbols $\tilde{\mathbf{x}}_k^{(m)} = \mathbf{w}_{\mathrm{T}} x_k^{(m)}$. As previously mentioned, a multi-beam radiation pattern is considered at the Tx to split the total available power between the communication and sensing directions. Hence, the beamforming vector $\mathbf{w}_{\mathrm{T}}$ is defined as follows

$$\mathbf{w}_{\mathrm{T}} = \frac{\sqrt{P_{\mathrm{T}} G_{\mathrm{T}}^{\mathrm{a}}}}{N_{\mathrm{T}}} \left(\sqrt{\rho_{\mathrm{p}}} \mathbf{a}_{\mathrm{T}}^c(\theta_{\mathrm{T,s}}) + \sqrt{1 - \rho_{\mathrm{p}}} \mathbf{a}_{\mathrm{T}}^c(\theta_{\mathrm{T,c}}) \right) \quad (1)$$

where $\rho_{\mathrm{p}} \in [0, 1]$ is the fraction of power reserved for the sensing beam, P_{T} is the transmit power, $G_{\mathrm{T}}^{\mathrm{a}}$ is the transmit array gain along the beam steering direction, and $\mathbf{a}_{\mathrm{T}}(\theta_{\mathrm{T,c}}) \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$ and $\mathbf{a}_{\mathrm{T}}(\theta_{\mathrm{T,s}}) \in \mathbb{C}^{N_{\mathrm{T}} \times 1}$ are the steering vectors associated with the communication and sensing directions, respectively, being $\theta_{\mathrm{T,c}}$ and $\theta_{\mathrm{T,s}}$ the respective directions of departure (DoDs).

Starting from the vector of the transmitted symbols $\tilde{\mathbf{x}}_k^{(m)}$, the vector $\tilde{\mathbf{y}}_k^{(m)} \in \mathbb{C}^{N_{\mathrm{R}} \times 1}$ of symbols received at each antenna, after OFDM demodulation, is given by

$$\tilde{\mathbf{y}}_k^{(m)} = \mathbf{H}_k^{(m)} \tilde{\mathbf{x}}_k^{(m)} + \tilde{\mathbf{n}}_k \quad (2)$$

where $\mathbf{H}_k^{(m)} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}}}$ is the channel matrix for the mth symbol and the kth subcarrier, which will be defined later, and $\tilde{\mathbf{n}}_k \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathrm{N}}^2 \mathbf{I}_{N_{\mathrm{R}}})$ is the noise vector.[1]

Spatial combining is then performed in the considered sensing direction, $\theta_{\mathrm{R,s}} = \theta_{\mathrm{T,s}}$, by using the receiving beamforming vector $\mathbf{w}_{\mathrm{R}} = \mathbf{a}_{\mathrm{R}}^c(\theta_{\mathrm{R,s}})$. This yields the grid of the received symbols $\mathbf{Y}_{\mathrm{s}} \in \mathbb{C}^{K \times M_{\mathrm{s}}}$, whose (k, m) elements are defined as $y_k^{(m)} = \mathbf{w}_{\mathrm{R}}^T \tilde{\mathbf{y}}_k^{(m)}$. The received symbols grid collected in each sensing direction is then used to generate range-angle maps, as explained in Section 2.2.

2.1 Target Models

This work considers both point-like targets, such as pedestrians, and extended targets, such as vehicles. Specifically, vehicles are represented by a model comprising 12 reflection points. These include 4 points to capture planar reflections originating from the front, back, and sides of the vehicle (characterized by a narrow visibility function and a substantial radar cross section, aka RCS), 4 points to account for the wheelhouses, and 4 points to simulate the corners [11–13].

Now, considering L as the total number of reflections from both extended and point-like targets, the channel matrix already introduced in Equation (2) is given by

$$\mathbf{H}_k^{(m)} = \sum_{l=1}^{L} \beta_l e^{j2\pi m T_{\mathrm{s}} f_{\mathrm{D},l}} e^{-j2\pi k \Delta f \tau_l} \mathbf{a}_{\mathrm{R}}(\theta_l) \mathbf{a}_{\mathrm{T}}^T(\theta_l) \quad (3)$$

where $\Delta f = 1/T$ is the subcarrier spacing, $T_{\mathrm{s}} = T + T_{\mathrm{cp}}$ is the total OFDM symbol duration including the cyclic prefix time T_{cp}. Additionally, $f_{\mathrm{D},l}$ refers to the Doppler

---

[1] Both inter-carrier interference (ICI) and inter-symbol interference (ISI) are considered negligible.

shift, τ_l represents the round-trip delay, θ_l denotes the direction of arrival (DoA), and $\mathbf{a}_{\mathrm{R}}(\theta_l)$ represents the array response vector at the Rx for the lth backscattered signal. The complex term $\beta_l = |\beta_l|\,e^{j\phi_l}$ includes phase shift and attenuation along the lth propagation path. The signal-to-noise ratio (SNR) at each receiving antenna related to the lth reflection point (hence the sensing SNR) becomes

$$
\begin{aligned}
\mathrm{SNR}_l^{(\mathrm{s})} &= \rho_{\mathrm{P}} \cdot \gamma_l \cdot \frac{P_{\mathrm{T}} G_{\mathrm{T}}^{\mathrm{a}} G_{\mathrm{R}}}{\sigma_{\mathrm{N}}^2} |\beta_l|^2 \\
&= \rho_{\mathrm{P}} \cdot \gamma_l \cdot \frac{P_{\mathrm{T}} G_{\mathrm{T}}^{\mathrm{a}} G_{\mathrm{R}}}{N_0 K \Delta f} \frac{c^2 \sigma_{\mathrm{rcs},l}}{(4\pi)^3 f_c^2 d_l^4}
\end{aligned}
\tag{4}
$$

where G_{R} represents the gain of a single antenna element at the Rx, $\gamma_l = |\mathrm{AF}(\theta_{\mathrm{T},\mathrm{s}} - \theta_l)|^2 \in [0,1]$ denotes the normalized array gain at the Tx, which considers the imperfect alignment between the sensing direction and the target DoA, N_0 is the one-sided noise power spectral density (PSD) at the Rx, d_l represents the distance between the lth reflection point and the BS, $\sigma_{\mathrm{rcs},l}$ corresponds to RCSs, f_{c} is the carrier frequency, and c is the speed of light.

The $\sigma_{\mathrm{rcs},l}$ RCSs of scatterers for both pedestrians and vehicles are random and modeled according to a Swerling I type distribution whose mean value, $\bar{\sigma}_{\mathrm{rcs}}$, can be found in Table 1 [14]. It is important to note that the number of backscattered signals L depends on the relative angular position with respect to the BS and varies over time according to a visibility function [11], as objects are moving.

2.2 Measurement Model

As mentioned, each BS detects objects in the environment by scanning using a multi-beam pattern defined in Equation (1). Specifically, the communication beam is directed toward a UE, while the sensing direction changes over time, sequentially pointing toward various directions following a predefined angular increment. In each direction, a set of M_{s} OFDM symbols is collected to form the grid of received symbols $\mathbf{Y}_{\mathrm{s}}$, which is then used to obtain a range-angle map. The period required to complete a full scan, denoted as T_{scan}, depends on the chosen number of sensing directions and on the symbol duration T_{s}. Once all the symbols are acquired and assembled into the matrix $\mathbf{Y}_{\mathrm{s}}$, the first step involves an element-wise division between $\mathbf{Y}_{\mathrm{s}}$ and $\mathbf{X}_{\mathrm{s}}$, an operation often indicated as reciprocal filtering [15, 16]. This division aims to eliminate the influence of the transmitted symbols and generate a new matrix denoted as $\mathbf{G}_{\mathrm{s}}$. Subsequently, a double-periodogram is performed on the rows and columns of $\mathbf{G}_{\mathrm{s}}$ to obtain a range-Doppler map [15]. From this map, a range-angle map $\mathbf{D}_{q,t}$ is derived

at the qth BS and tth scan by selecting the column of the periodogram with the maximum value and uniquely associating it with the corresponding scan direction.[2]

Table 1 Average RCS for different reflection points

Reflection	$\bar{\sigma}_{\mathrm{rcs}}\,(\mathrm{m}^2)$
Pedestrians	1
Surfaces	20
Wheelhouses	0
Corners	5

3 Data Fusion, Target Classification, and Target-oriented Processing

According to the block diagram depicted in Figure 2, each BS exchanges the range-angle map (denoted as $\mathbf{D}_{q,t}$) with the FC. The FC employs a linear uniform grid, with resolution Δ_{x} and Δ_{y} (with N_{x} and N_{y} points) as a baseline. The received maps are rotated and translated according to the specific BS position and ULA orientation, and resampled at the baseline grid to ensure consistent map fusion. Subsequently, the resampled range-angle maps, represented as $\overline{\mathbf{D}}_{q,t}$, are combined via element-wise summation to yield the soft map $\mathbf{L}_t = \sum_{q=1}^{N_{\mathrm{s}}} \overline{\mathbf{D}}_{q,t}$, where N_{s} is the number of BSs performing sensing.[3]

3.1 Target Identification

Each target exhibits a different reflection pattern related to its geometrical shape and RCS, namely its reflection fingerprint [9]. To this end, a convolutional neural network (CNN) is adopted to infer the target category (pedestrian or vehicle) directly from the resampled and fused soft maps $\mathbf{L}_t$ which contain such information.

Following Figure 2, a first step named image cropping is required to isolate each target from the others. A square window with side W_{size} pixels is selected to frame each target. Such windows are centered in the predicted target position at time t, inferred by the tracking algorithms exploiting information extracted during the previous time

---

[2] The estimation of target parameters turns out to be a frequency estimation problem. Since the periodogram represents (asymptotically) the log-likelihood, the column with the maximum value is selected.

[3] Since $\mathbf{L}_t$ are obtained via periodogram estimation, they can be interpreted as target log-likelihood maps, hence their summation results from noise independence among BSs.
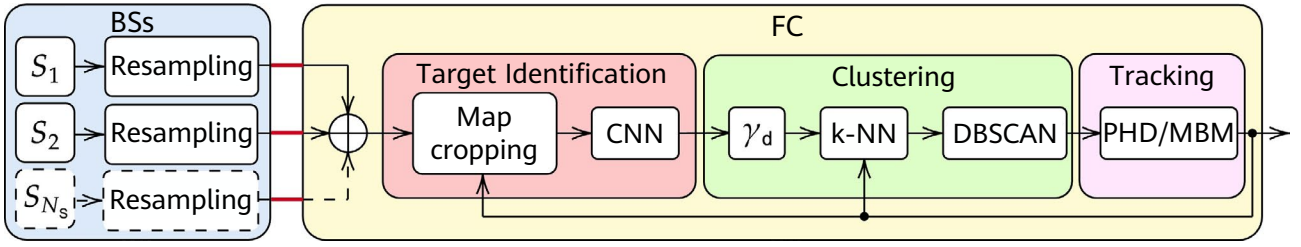
Figure 2 Block diagram of the sensing processing chain exploiting BS cooperation, target classification, and target-specific tracking. The BSs scan the environment generating range-angle maps and resample them according to a predefined grid. Resampled range-angle maps are then shared with the FC and fused in a single map. Target identification is performed at the FC through map cropping and classification (red block). Then clustering is performed to merge detections generated by the same target (green block). Finally, tracking algorithms perform target state estimation (pink block).

step $t-1$. To generate the training set for the CNN, we consider a scenario where actual target positions and categories are known. To increase the classifier performance and robustness in the presence of imperfect target state predictions, which result in a misalignment between targets and relative frames, during training, the real target position is perturbed, adding Gaussian noise (which acts as a random displacement) with standard deviation σ_{w} in both x and y directions. This solution leads to more accurate target classification, reducing the generalization error. At the end of the training phase, the CNN can infer the target category in real time and in a different scenario.

3.2 Adaptive Clustering

A three-step clustering procedure is employed to extract detections from the soft maps, enabling effective handling of extended and point-like objects (refer to the green block in Figure 2 for a visual representation of the clustering procedure). The main steps of the proposed strategy can be summarized as follows:

1. An excision filter is implemented with threshold γ_{d}, to remove points with low values from the $\mathbf{L}_t$ maps which are likely produced by noise.

2. A k-nearest neighbors (k-NN) algorithm with $k=1$ and adaptive gate ξ_{k} (to ignore residual points distant from each target) is applied to cluster data that likely belong to a previously detected target [17]. It is important to highlight that the parameter ξ_{k} can be adapted and varied depending on the target category. Section 4 compares the solution with fixed values of ξ_{k} and the adaptive solution.

3. The remaining points (i.e., map points larger than γ_{d} and outside the gate ξ_{k}) are clustered through the density-based spatial clustering of applications with noise (DBSCAN) algorithm, with a maximum distance between points belonging to the same cluster ξ_{d}, and a minimum number of points to form a cluster N_{d} [18].

Finally, each cluster centroid is stored in the matrix $\mathbf{Z}_t$, representing target detections extracted from the soft maps.

3.3 Tracking Algorithms

For all the tracking algorithms, we adopt the following state vector to represent the state of each target:

$$\mathbf{s}_{t,n} = \left(s_{t,n,\mathrm{x}}, s_{t,n,\mathrm{y}}, s_{t,n,v_{\mathrm{x}}}, s_{t,n,v_{\mathrm{y}}} \right)^{\mathrm{T}} \quad (5)$$

where t and n are the time and target indexes. The first two elements of the vector correspond to the target position coordinates, while the last two represent the target velocity components. To update the target position coordinates, we use the information extracted from the map, while the target velocity components are inferred by considering both the previous target position at time $t-1$ and the current target position.

The PHD filter is a widely adopted algorithm in literature [19, 20]. One possible implementation suggests approximating the target intensity function as a Gaussian mixture (GM) with a predefined number of components, which takes the following form:

$$D_{t-1|t-1}(\mathbf{x}) = \sum_{h=1}^{\mathcal{H}_{t-1|t-1}} w_{t-1|t-1}^{(h)} \mathcal{N}_{\mathbf{x}}(\boldsymbol{\mu}_{t-1|t-1}^{(h)}, \mathbf{P}_{t-1|t-1}^{(h)}) \quad (6)$$

where $\mathbf{x}$ is a generic random finite set (RFS), $\mathcal{H}_{t-1|t-1}$ represents the number of Gaussian components in the intensity function, $w_{t-1|t-1}^{(h)}$ is the hth component weight, and $\boldsymbol{\mu}_{t-1|t-1}^{(h)}$ and $\mathbf{P}_{t-1|t-1}^{(h)}$ represent the mean and covariance of the considered component. The intensity function can be interpreted as an atypical probability density function (p.d.f.) whose integral returns the estimated number of targets in the scenario.

The prediction step infers the intensity function in the consecutive time step, i.e., $D_{t|t-1}(\mathbf{x})$, through a linear Kalman predictor [21]. During prediction, the probability of survival P_{s} is considered constant, so are the transition

matrix $\mathbf{F}$ and the process noise covariance matrix $\mathbf{Q}$ — the last one represents the motion uncertainty. A set of $\mathcal{B}$ birth components is added to the predicted intensity function $D_{t|t-1}(\mathbf{x})$ to represent the possibility of new targets spawning in the surveillance area. The total number of components after prediction is then $\mathcal{H}_{t|t-1} = \mathcal{H}_{t-1|t-1} + \mathcal{B}$.

In the update step, the predicted components are updated through the Kalman update equations, as in [21], with the measurements $\mathbf{Z}_t$ extracted from the maps $\mathbf{L}_t$. During this step, the detection probability P_{d} is considered constant, and the covariance matrix $\mathbf{R}_t$ for each measurement is estimated from the selected map detection points, as will be highlighted in Equation (15). The overall amount of components in the posterior can be written as $\mathcal{H}_{t|t} = \mathcal{H}_{t|t-1}(M_t + 1)$, where M_t denotes the number of measurements at time instant t.

To estimate the number of targets from the PHD posterior, it is enough to sum the weight of the components and round it to the closest integer

$$\hat{N}_{\mathrm{obj}} = \left\lceil \sum_{h=1}^{\mathcal{H}_{t|t}} w_{t|t}^{(h)} \right\rfloor \tag{7}$$

while for the nth target state estimation, we extract the mean value of the nth most likely component

$$\hat{\mathbf{s}}_{t,n} = \arg\max_{w_{t|t}^{(h)}} \boldsymbol{\mu}_{t|t}^{(h)} \tag{8}$$

The MBM filter is an alternative to the PHD filter for multi-target tracking problems that exploit the association probability between measurements and targets [22, 23]. The MBM filter is used to approximate the target multi-object p.d.f.

$$\mathrm{MBM}_{t-1|t-1}(\mathbf{x}) = \sum_{g=1}^{\mathcal{G}_{t-1|t-1}} w_{t-1|t-1}^{(g)} \mathrm{MB}_{t-1|t-1}^{(g)}(\mathbf{x}) \tag{9}$$

where $\mathcal{G}_{t-1|t-1}$ represents the number of multi-Bernoulli (MB) components or global hypothesis in the MBM distribution, and $w_{t-1|t-1}^{(g)}$ stands for the gth MBM component weight. The MB distribution in Equation (9) can be written as

$$\mathrm{MB}_{t-1|t-1}^{(g)}(\mathbf{x}) = \sum_{\uplus \mathbf{x}_l = \mathbf{x}} \prod_{l=1}^{\mathcal{L}_{t-1|t-1}^{(g)}} \mathrm{B}_{t-1|t-1}^{(g,l)}(\mathbf{x}_l) \tag{10}$$

where $\mathcal{L}_{t-1|t-1}^{(g)}$ represents the number of Bernoulli components or local hypothesis in the MB distribution, and the summation is performed for all the possible unions

of mutually disjoint RFS that generate $\mathbf{x}$, which means to evaluate all the possible data associations between measurements and targets [23]. The single Bernoulli component in Equation (10) can be written as

$$\mathrm{B}_{t-1|t-1}^{(g,l)}(\mathbf{x}_l) = r_{t-1|t-1}^{(g,l)} \mathcal{N}_{\mathbf{x}_l}(\boldsymbol{\mu}_{t-1|t-1}^{(g,l)}, \mathbf{P}_{t-1|t-1}^{(g,l)}) \tag{11}$$

where $r_{t-1|t-1}^{(g,l)}$ represents the existence probability of the lth local hypothesis in the gth global hypothesis, and $\boldsymbol{\mu}_{t-1|t-1}^{(h)}$ and $\mathbf{P}_{t-1|t-1}^{(h)}$ represent the mean and covariance of the considered component, respectively.

During prediction, linear Kalman prediction is performed again to infer the parameters in the consecutive time step. To account for new spawning objects, a set of $\mathcal{B}$ Bernoulli components is added to each global hypothesis. For both algorithms, to exploit the prior information about the environment, the components are generated following the scenario layout, i.e., the number of hypotheses, their mean value, covariance, and weight are based on the lanes and crosswalk positions in the environment. The overall number of components after the prediction step can be evaluated as $(\mathcal{L}_{t-1|t-1} + \mathcal{B})\mathcal{G}_{t-1|t-1}$.

In the update phase, a linear Kalman update is performed to derive the updated parameters, considering the most likely association between measurements and targets [22]. Estimations $\hat{\mathbf{s}}_{t,n}$ are then extracted from the posterior distribution, considering the mean value $\boldsymbol{\mu}_{t|t}^{(i,j)}$ of the MB components with existence probability $r_{t|t}^{(i,j)} \geq \gamma_{\mathrm{e}}$ from the MBM component with the highest probability $w_{t|t}^{(i)}$.

3.4 Motion and Measurement Model

To model clutter measurements representing false alarm detection extracted by the clustering procedure, a Poisson point process (PPP) is considered, whose intensity is defined as λ_{c}.

Target death is modeled through a constant probability of survival P_{s}. During prediction, if a component is associated with a missed detection, its weight is multiplied by a factor proportional to P_{s}, which means that consecutive missed detections lead to unlikely target state components.

A linear prediction model is selected to track the behavior of both extended and point-like targets. This is justified by the low value of T_{scan} compared to the target velocity, which allows to approximate target motions as piecewise linear among consecutive acquisitions. The corresponding transition matrix and process noise covariance matrix are

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & T_{\text{scan}} & 0 \\ 0 & 1 & 0 & T_{\text{scan}} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (12)$$

$$\mathbf{Q} = \alpha_{\text{q}} \, T_{\text{scan}} \cdot \mathbf{I}_4 \qquad (13)$$

where α_{q} is a parameter that represents the prediction uncertainty about the target motion.

In this work, only position information about the targets is estimated through measurements, while consecutive position measurements are used to infer the velocity.[4] With these assumptions, the following measurement matrix is considered:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \qquad (14)$$

Because of the high-resolution maps, multiple detections (closely spaced map pixels) from each target are generated, leading to a non-diagonal measurement covariance matrix. Thus such a matrix needs to be estimated. Let us define the set of map points $\mathbf{L}_t^{(\mathbf{z}_{t,m})}$ extracted after clustering, specified by $\mathbf{z}_{t,m}$, from the measurement matrix (map) $\mathbf{L}_t$. Indicating with $\mathbf{V}_t^{(\mathbf{z}_{t,m})}$, the $2 \times N_{\mathbf{z}_{t,m}}$ matrix containing the pixel coordinates relative to $\mathbf{L}_t^{(\mathbf{z}_{t,m})}$, the sample covariance measurement matrix can be calculated as

$$\mathbf{R}_t = \frac{1}{N_{\mathbf{z}_{t,m}} - 1} \sum_{p=1}^{N_{\mathbf{z}_{t,m}}} (\mathbf{v}_{t,p}^{(\mathbf{z}_{t,m})} - \mathbf{z}_{t,m})(\mathbf{v}_{t,p}^{(\mathbf{z}_{t,m})} - \mathbf{z}_{t,m})^{\mathrm{T}} \qquad (15)$$

where $N_{\mathbf{z}_{t,m}}$ represents the number of map points associated to the mth measurement $\mathbf{z}_{t,m}$, and $\mathbf{v}_{t,p}^{(\mathbf{z}_{t,m})}$ stands for the pth map point coordinates in the matrix $\mathbf{V}_t^{(\mathbf{z}_{t,m})}$.

3.5 Post-Processing

A set of post-processing procedures are implemented to manage the complexity of the algorithms and ensure good estimation accuracy. In the PHD filter, pruning, capping, and merging are implemented sequentially to reduce the number of components in the posterior intensity function. Pruning removes all the components in the posterior whose weights $w_{t|t}^{(h)}$ are under a predefined threshold γ_{P} [20]. Then, capping is realized on the remaining components selecting the γ_{q} components with the greatest $w_{t|t}^{(h)}$, by fixing the maximum number of components in the posterior to γ_{q} [22].

Finally, on the remaining components in the set ζ_{m}, merging of those whose average distance, defined in the following equation, is lower than a predefined threshold γ_{s} is performed:

$$d(\boldsymbol{\mu}_{t|t}^{(i)}, \boldsymbol{\mu}_{t|t}^{(j)}) = \|\boldsymbol{\mu}_{t|t}^{(i)} - \boldsymbol{\mu}_{t|t}^{(j)}\|_2 \qquad (16)$$

where weights, mean, and covariance are updated as follows:

$$w_{t|t}^{(k)} = \sum_{i \in \zeta_{\text{m}}} w_{t|t}^{(i)}$$

$$\boldsymbol{\mu}_{t|t}^{(k)} = \sum_{i \in \zeta_{\text{m}}} w_{t|t}^{(i)} \boldsymbol{\mu}_{t|t}^{(i)}$$

$$\mathbf{P}_{t|t}^{(k)} = \sum_{i \in \zeta_{\text{m}}} w_{t|t}^{(i)} \mathbf{P}_{t|t}^{(i)} + (\boldsymbol{\mu}_{t|t}^{(i)} - \boldsymbol{\mu}_{t|t}^{(k)})(\boldsymbol{\mu}_{t|t}^{(i)} - \boldsymbol{\mu}_{t|t}^{(k)})^{\mathrm{T}}$$

where k represents the new index assigned to the derived component, and i represents the index of the merged component.

In the MBM filter, during the update phase, a gate for eligible data association allows pruning all the weak association hypotheses with $w_{t|t}^{(k)} < \xi_{\text{a}}$. Both the MBM and the MB components are pruned with the threshold γ_{g} and γ_{l}, respectively. Then, the residual MBM components are capped with a threshold γ_{c}. To increase the estimation accuracy, in the most likely MBM components, the MB components closer than γ_{m} are merged as previously described.

4 Numerical Results

4.1 Performance Metrics

From the communication perspective, the aggregate network capacity, intended as the sum rate of each BS in the downlink, is considered to assess the communication performance. Considering N_{s} BSs dedicated for both sensing and communications among the N_{tot} available BSs (so $N_{\text{tot}} - N_{\text{s}}$ are for communications only) and a fraction of power dedicated for sensing ρ_{P} (1), the overall aggregate network capacity can be written as

$$C(\rho_{\text{P}}) = (N_{\text{s}} \, \Delta f \, K \, \log_2(1 + (1 - \rho_{\text{P}}) \, \text{SNR}^{(\text{c})})$$
$$+ (N_{\text{tot}} - N_{\text{s}}) \, \Delta f \, K \, \log_2(1 + \text{SNR}^{(\text{c})}))/N_{\text{tot}} \qquad (17)$$

where $\text{SNR}^{(\text{c})}$ is the communication SNR experienced by the UEs.[5]

To evaluate the network localization capability, the OSPA is selected as a single-value metric [24, 25]:

---

[4] Although possible, we consider the BSs do not estimate target Doppler.

[5] To keep the presentation of numerical results simple, we consider all the UEs experience the same SNR.

$$\mathrm{OSPA} =$$

$$\sqrt[\mathrm{p}]{\frac{1}{N_\mathrm{c}}\left(\sum_{(i,j)\in\boldsymbol{\zeta}_\mathrm{g}^*} d(\underline{\mathbf{s}}_{t,i}, \hat{\underline{\mathbf{s}}}_{t,j})^\mathrm{p} + \frac{\xi_\mathrm{g}^\mathrm{P}}{2}(|\underline{\mathbf{S}}_t| + |\hat{\underline{\mathbf{S}}}_t| - 2|\boldsymbol{\zeta}_\mathrm{g}^*|)\right)}$$

where $\hat{\underline{\mathbf{S}}}_t$ contains the position coordinates for all the estimated targets in the scenario representing the first two rows of $\hat{\mathbf{S}}_t$ inferred by the algorithms, and N_c is the number of elements in the OSPA metric given by $N_\mathrm{c} = |\underline{\mathbf{S}}_t| + |\hat{\underline{\mathbf{S}}}_t| - |\boldsymbol{\zeta}_\mathrm{g}^*|$, where $|\cdot|$ represents the RFS cardinality. The parameter P is the OSPA order, while ξ_g represents the OSPA gate. Estimations beyond the gate threshold ξ_g are categorized as false alarms, while actual target positions not linked to estimations within the gate are classified as missed detections. The set $\boldsymbol{\zeta}_\mathrm{g}^*$ represents the best assignment between the estimated set of objects $\hat{\underline{\mathbf{S}}}_t$ and the ground truth one $\underline{\mathbf{S}}_t$. The best assignment is selected as the one that minimizes the OSPA error. Note that the first term in the OSPA metric, i.e., $d(\underline{\mathbf{s}}_{t,i}, \hat{\underline{\mathbf{s}}}_{t,j})^\mathrm{p}$, can be interpreted as the distance between the estimated target positions and the real ones (and if $p = 2$, this term corresponds to the square error of position estimation). The rest of the OSPA metric can be rewritten as

$$\frac{\xi_\mathrm{g}^\mathrm{P}}{2}(|\underline{\mathbf{S}}_t| - |\boldsymbol{\zeta}_\mathrm{g}^*|) + \frac{\xi_\mathrm{g}^\mathrm{P}}{2}(|\hat{\underline{\mathbf{S}}}_t| - |\boldsymbol{\zeta}_\mathrm{g}^*|) \tag{18}$$

where the first term is proportional to the missed detections, while the second term is related to the false alarms.

Finally, to evaluate the target classification performance, we define the classification accuracy as

$$\mathrm{Accuracy} = \frac{T_\mathrm{P} + T_\mathrm{N}}{T_\mathrm{P} + T_\mathrm{N} + F_\mathrm{P} + F_\mathrm{N}} \tag{19}$$

where T_P, T_N, F_P, and F_N stand for true positive, true negative, false positive, and false negative classifications, respectively.

4.2 Parameter Setup

We considered a scenario with $N_\mathrm{tot} = 6$ BSs, and a set of 4 extended and 4 point-like targets. Pedestrians and vehicles move as a mix of constant turns, linear accelerated/decelerated, static, and uniform linear motions. The area monitored is of size $x \in [-20, 20]$ m and $y \in [-20, 20]$ m. The BSs are positioned on a circumference of radius 50 m centered in the surveillance area, with the axes normal to the ULAs pointing toward the center, a scanning aperture of 120° (i.e., $\Theta_0 = 60°$), and scan step $\Delta\Theta = 2.4°$.

The transmission parameters are: quadrature phase shift keying (QPSK) modulation, $f_\mathrm{c} = 28\,\mathrm{GHz}$, $\Delta f = 120\,\mathrm{kHz}$, $K = 3168$ (i.e., about 400 MHz bandwidth), $M = 1120$, and $M_\mathrm{s} = 112$. The effective isotropic radiated power (EIRP) is set to 30 dBm, and the noise PSD is $N_0 = 4 \cdot 10^{-20}$ W/Hz. All the 6 BSs are equipped with $N_\mathrm{T} = N_\mathrm{R} = 50$ antennas. At each BS, the scan duration is $T_\mathrm{scan} = 50$ ms, and the overall scenario is monitored for 10s, resulting in $N_\mathrm{m} = 200$ measurements (maps collected). The grid resolution for map fusion is set to $\Delta_\mathrm{x} = 0.1$ m and $\Delta_\mathrm{y} = 0.1$ m. The fraction of power dedicated to sensing ρ_P is set to 0.3; the same for all BSs.

For the target identification task, the window size is set to $W_\mathrm{size} = 6$ m, and the position perturbation standard deviation is $\sigma_\mathrm{w} = 0.5$ m. The CNN is composed of a 2D convolutional layer with 20 random mask filters of dimension 5×5 and the rectified linear unit (ReLU) activation function, a consecutive 2D max pooling layer which performs a down-sampling of a factor 2, and a fully connected layer with the softmax activation function whose output dimension is 2, to map the extracted features in the target classes.

In the clustering algorithm, the detection threshold is set to $\gamma_\mathrm{d} = 2 \cdot 10^{-7}$. The k-NN gate ξ_k is tested for values between 4 and 6. For DBSCAN, the cutoff distance is $\xi_\mathrm{d} = 3$, and the minimum number of points to form clusters is set to $N_\mathrm{d} = 50$.

In both tracking algorithms, the clutter intensity is $\lambda_\mathrm{c} = 0.1$, the prediction uncertainty is $\alpha_\mathrm{q} = 5$, and the initial component covariance is set to $\mathbf{P} = 0.5 \cdot \mathbf{I}_4$. The probabilities of detection and survival are $P_\mathrm{d} = 0.99$ and $P_\mathrm{s} = 0.9$, respectively. The merging threshold is set to $\gamma_\mathrm{m} = 5$. In the PHD filter, the pruning threshold of the components is $\gamma_\mathrm{p} = 100 \cdot 10^{-6}$, while the maximum number of components is fixed to $\gamma_\mathrm{q} = 10$. In the MBM filter, the pruning threshold on the probability of existence is $\gamma_1 = 100 \cdot 10^{-6}$ while the pruning threshold on the MBM components is set to $\gamma_\mathrm{g} = 10^{-15}$. The maximum number of MBM components is $\gamma_\mathrm{c} = 10$. The gate for the admissible associations is set to $\xi_\mathrm{a} = 14$. The existence threshold is $\gamma_\mathrm{e} = 0.99$. For both the algorithms, the birth components for new appearing objects are initialized with covariance $\mathbf{P}^{(b)} = 0.1 \cdot \mathbf{I}_4$, with position $\boldsymbol{\mu}^{(b)}$ reflecting the possible target spawn position. A recovery component is initialized centered in the scenario with covariance $\mathbf{P}^{(b)} = 5 \cdot \mathbf{I}_4$.

4.3 Target Classification Performance

In Figure 3, the classification accuracy for $\rho_P = 0.3$, and varying the number of BSs selected for sensing N_s, is reported in green for both PHD and MBM, on the top and bottom plots, respectively. It can be noticed that both the algorithms ensure a target classification accuracy greater than 0.85 when the number of BSs is $N_s \geq 3$. Classification performance is highly influenced by the number of BSs adopted for sensing; reducing the number of BSs reduces the number of detected reflection points of extended targets (because of the reduction of spatial diversity), resulting in more similar target fingerprints between pedestrians and vehicles in the fused maps. It is also interesting to notice that with $N_s = 6$ BSs, the target classification accuracy is greater than 98%, representing a remarkable result.

4.4 Sensing and Communication Performance

In Figure 3, the OSPA metric for $\rho_P = 0.3$, and varying the number of sensors N_s, is illustrated for PHD and MBM, on the top and bottom, respectively. Blue dotted curves represent the algorithm performance with $\xi_k = 4$ for both pedestrians and vehicles; red dashed curves refer to the performance with $\xi_k = 6$ again for both pedestrians and vehicles. Solid yellow curves represent the performance of the AI-based solution, whose gates are adapted for

pedestrians ($\xi_k = 4$) and vehicles ($\xi_k = 6$) based on the target identification.

As can be noticed, the adaptive gate achieves a lower localization error for both algorithms. For the PHD filter, the solution with adaptive gating presents an error lower than 1 m when the number of sensors is $N_s \geq 3$. Similarly, the MBM filter exhibits an OSPA lower than 0.7 m considering $N_s \geq 3$. The performance degradation experienced when $N_s < 3$ is due to target misclassification. In this case, the adaptive solution is affected by the mismatch between the real target classes and the estimated ones, resulting in an incorrect assignment of the gating parameter ξ_k.

To emphasize the benefit produced by the adoption of adaptive gating (see Figure 4), the number of BSs devoted to sensing is fixed. At the same time, the OSPA metric is reported over the first 100 acquisitions. Blue areas represent the OSPA produced by a fixed gate $\xi_k = 4$ for both pedestrians and vehicles, red areas refer to the solution with $\xi_k = 6$, and yellow areas represent the adaptive solution. It is important to highlight the increase in the localization performance thanks to adaptive gating for both algorithms, which results in reduced OSPA peaks.

From a communication perspective, the BS aggregate capacity is evaluated with Equation (17), considering $\rho_P = 0.3$. The worst case for communication is when all the BSs are performing joint communication and sensing, i.e., $N_s = 6$. In this situation, the downlink capacity is $C = 0.9$ Gbit/s. On the contrary, without performing sensing
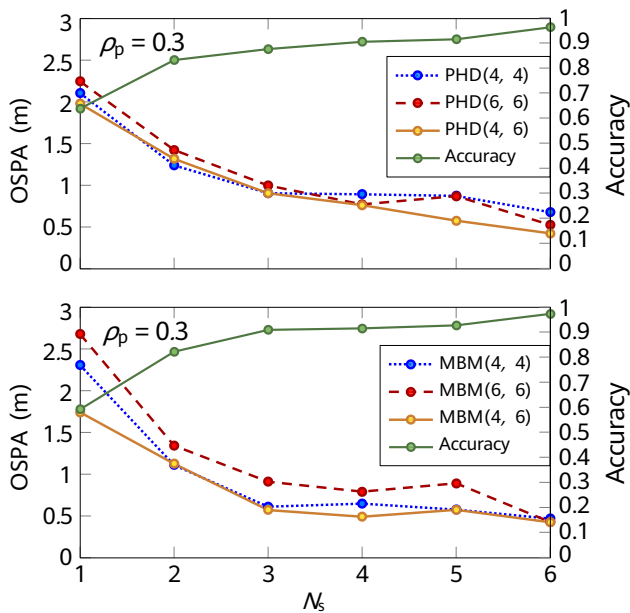

Figure 3 Localization performance and classification accuracy varying the number of BSs devoted for sensing N_s for the PHD (top) and MBM (bottom).
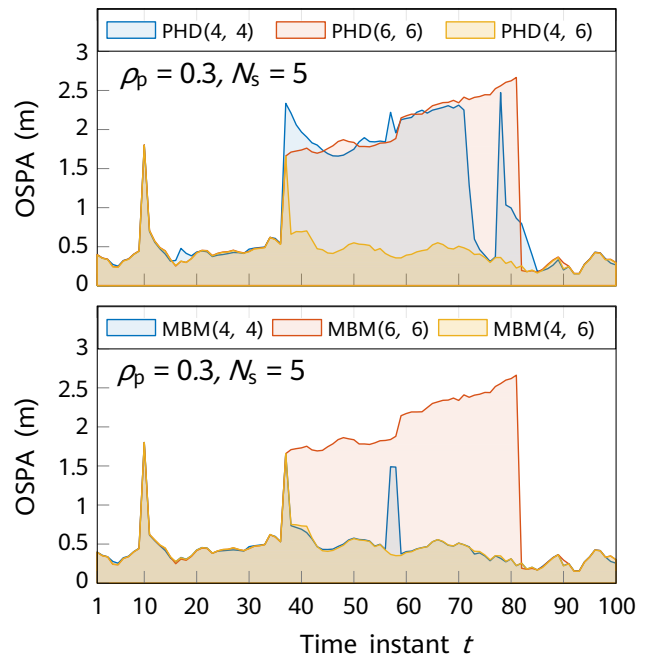

Figure 4 Localization performance over time for the PHD (top) and MBM (bottom).

$(N_\text{s} = 0)$, the downlink capacity C can reach 1.1 Gbit/s. As a compromise, using 3 BSs for JSC and 3 for communications only, the downlink capacity can be maintained greater than 1 Gbit/s.

5 Conclusion

In this work, we presented a framework to perform JSC with OFDM waveforms exploiting cooperation and data fusion among BSs to improve localization performance. Furthermore, leveraging different target reflection fingerprints in the soft maps, we developed a CNN classifier to identify the object type and adapt the multi-target tracking to the specific object type.

A three-step clustering strategy based on adaptive gating is proposed to manage point-like and extended targets and exploit target identification. Then, two multi-target tracking algorithms are used, the PHD and MBM filters, to track all the targets in the surveillance area.

The overall system is tested in a vehicular scenario with two types of targets, pedestrians and vehicles. To explore the communication/sensing trade-off, we investigated the sensing performance varying the number of cooperating BSs, considering that a fraction of transmit power is devoted to the sensing beams.

The system performance has been evaluated through the OSPA metric, target classification accuracy, and communication performance via aggregate downlink capacity. Numerical results show that adaptive gating aided by target identification performs better than the simpler target-agnostic solution when the target classification accuracy is greater than 90%. For example, by choosing $N_\text{s} = 3$ BSs, a classification accuracy around 0.9 is reached, with an OSPA error lower than 1 m for the PHD filter and around 0.7 m for the MBM filter, while also ensuring a downlink capacity greater than 1 Gbit/s. With $N_\text{s} = 6$ sensing BSs, a target classification accuracy larger than 98% is reached, with a localization error lower than 0.7 m for both tracking algorithms, and with 10% penalty on downlink capacity, i.e., from 1 Gbit/s to 0.9 Gbit/s.

References

[1]   R. Thomä, T. Dallmann, S. Jovanoska, P. Knott, and A. Schmeink, "Joint communication and radar sensing: An overview," in *Europ. Conf. on Ant. and Prop. (EuCAP)*, Dusseldorf, Germany, Mar. 2021, pp. 1–5.

[2]   S. Schieler, C. Schneider, C. Andrich, M. Döbereiner, J. Luo, A. Schwind, R. S. Thomä, and G. Del Galdo, "OFDM waveform for distributed radar sensing in automotive scenarios," *Int. J. of Microw. and Wireless Tech.*, vol. 12, no. 8, p. 716–722, 2020.

[3]   J. Zhang, X. Wang, P. Liu, J. Huang, and Z. Zheng, "Joint resource allocation and user association for multi-cell integrated sensing and communication systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2023, 07 2023.

[4]   Z. Wang, K. Han, J. Jiang, F. Liu, and W. Yuan, "Multi-vehicle tracking and ID association based on integrated sensing and communication signaling," *IEEE Wireless Commun. Letters*, vol. 11, no. 9, pp. 1960–1964, 2022.

[5]   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 J. of Sel. Topics in Signal Process.*, vol. 15, no. 6, pp. 1295–1315, 2021.

[6]   Y. Cui, F. Liu, X. Jing, and J. Mu, "Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges," *IEEE Netw.*, vol. 35, pp. 158–167, 2021.

[7]   R. Liu, M. Jian, D. Chen, X. Lin, Y. Cheng, W. Cheng, and S. Chen, "Integrated sensing and communication based outdoor multi-target detection, tracking and localization in practical 5G networks," *arXiv preprint arXiv:2305.13924*, 2023.

[8]   R. Chauhan, K. K. Ghanshala, and R. Joshi, "Convolutional neural network (CNN) for image detection and recognition," in *2018 1th*

Int. Conf. on Secure Cyber Comput. and Commun. (ICSCCC), 2018, pp. 278–282.

[9] O. Kechagias-Stamatis and N. Aouf, "Automatic target recognition on synthetic aperture radar imagery: A survey," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 36, no. 3, pp. 56–81, 2021.

[10] Y. Tian, "Artificial intelligence image recognition method based on convolutional neural network algorithm," *IEEE Access*, vol. 8, pp. 125 731–125 744, 2020.

[11] M. Bühren and B. Yang, "Simulation of automotive radar target lists using a novel approach of object representation," in *IEEE Intell. Veh. Symp.*, 2006, pp. 314–319.

[12] E. Favarelli, E. Matricardi, L. Pucci, E. Paolini, W. Xu, and A. Giorgetti, "Sensor fusion and extended multi-target tracking in joint sensing and communication networks," in *IEEE Int. Conf. on Commun. (ICC)*, accepted, Rome, Italy, May 2023.

[13] ——, "Map fusion and heterogeneous objects tracking in joint sensing and communication networks," in *20th European Radar Conference (EuRAD),* Berlin, Germany, Sep. 2023.

[14] M. I. Skolnik, *Radar handbook*. McGraw-Hill Education, 2008.

[15] L. Pucci, E. Paolini, and A. Giorgetti, "System-level analysis of joint sensing and communication based on 5G new radio," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, July 2022, pp. 2043–2055.

[16] J. T. Rodriguez, F. Colone, and P. Lombardo, "Supervised reciprocal filter for OFDM radar signal processing," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1–22, 2023.

[17] J. Watt, R. Borhani, and A. K. Katsaggelos, *Machine Learning Refined*. Cambridge University Press, 2016.

[18] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Int. Conf. on Know. Disc. in Data Mining*, Portland, Oregon, 1996, pp. 226–231.

[19] R. Mahler, "PHD filters of higher order in target number," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 4, pp. 1523–1543, 2007.

[20] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, 2006.

[21] Q. Li, R. Li, K. Ji, and W. Dai, "Kalman filter and its application," in *8th International Conference on Intell. Networks and Intell. Systems (ICINIS)*, 2015, pp. 74–77.

[22] A. F. García-Fernández, Y. Xia, K. Granström, L. Svensson, and J. L. Williams, "Gaussian implementation of the multi-Bernoulli mixture filter," in *22th Int. Conf. on Inf. Fusion (FUSION)*, 2019, pp. 1–8.

[23] A. F. García-Fernández, J. L. Williams, K. Granström, and L. Svensson, "Poisson multi-Bernoulli mixture filter: Direct derivation and implementation," *IEEE Trans. on Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1883–1901, 2018.

[24] A. S. Rahmathullah, A. F. García-Fernández, and L. Svensson, "Generalized optimal sub-pattern assignment metric," in *20th Int. Conf. on Inf. Fusion (FUSION)*, 2017, pp. 1–8.

[25] M. Beard, B. T. Vo, and B.-N. Vo, "OSPA(2): Using the OSPA metric to evaluate multi-target tracking performance," in *Int. Conf. on Control, Autom. and Inf. Sci. (ICCAIS)*, 2017, pp. 86–91.

On the Performance of SWIPT Integrated Sensing and Communication Systems

Meng Liu, Minglei Yang, Nan Liu

National Key Laboratory of Radar Signal Processing, Xidian University

liumeng2021@163.com, mlyang@xidian.edu.cn, liunaneoe@163.com

Abstract

This paper investigates the performance of simultaneous wireless information and power transfer (SWIPT) non-orthogonal downlink transmission integrated sensing and communication (ISAC) system. Particularly, residual hardware impairments, channel estimation errors, and imperfect successive interference cancelation are taken into account. The expressions of exact and asymptotic outage probabilities, probability of detection (PoD), and energy efficiency (EE) are derived for communication and sensing (C&S) performance evaluation. The simulation results verify the accuracy of our theoretical analysis and reveal that introducing SWIPT into ISAC systems can improve both C&S performance in terms of fairness, PoD as well as EE.

Keywords

integrated sensing and communication, non-orthogonal downlink transmission, simultaneous wireless information and power transfer

1 Introduction

With numerous devices access to the communication networks, the traditional communication frequency band has been unable to meet the basic needs of communication users, which leads to the proximity of communication and radar frequency bands [1]. Furthermore, the accelerated development of intelligent scenarios, e.g., smart city, intelligent transportation, and so on, has also led to the injection of novel capabilities into communication systems, i.e., sensing capabilities [2]. To address this issue, integrated sensing and communication (ISAC) is considered as an effective and feasible solution due to the realization of communication and sensing (C&S) simultaneously as well as economization of hardware cost [3]. Specifically, an ISAC base station (BS) transmits a unified signal, such that the purpose of users communication and sensing the non-cooperative targets can be achieved simultaneously [2–5].

Previous works [3–5] have laid solid foundation for the investigation of ISAC. Nevertheless, we notice that these works were based on traditional orthogonal transmission, where only one user can be served in the same time/frequency/code domain. By contrast, non-orthogonal transmission (NO-T) is regarded as a promising technology in the way of resource reuse at the transmitters and successive interference cancellation (SIC) at the receivers, which can serve multiple users while guaranteeing the users' quality of service [6]. Against this background, the outage probability (OP) of the communication users and ergodic sensing rate of the echoes were derived for evaluating the C&S performance of the non-orthogonal downlink transmission (NO-DLT) and non-orthogonal uplink transmission (NO-ULT) ISAC systems in [7] and [8]. These mentioned works revealed that NO-T provides higher performance gains compared with traditional orthogonal transmission in terms of C&S performance.

Apart from the pursuit of high spectral efficiency, the improvement of energy efficiency (EE) is also worth attention in NO-T ISAC networks. As a remedy, simultaneous wireless information and power transfer (SWIPT), harvesting energy from radio frequency (RF) signals, can further enhance the resource utilization and prolong the lifetime of ISAC networks [9]. Generally, power splitting (PS) and time switching (TS) are applied at receivers for the purpose of energy harvesting (EH). The OPs of users and throughput of the SWIPT NO-DLT system were analyzed in [10], where the TS scheme was exploited at the relay for EH. The power allocation scheme was formulated in [11] for maximizing

EE of the millimeter-wave SWIPT NO-DLT system, where PS was performed at the communication receivers. The power allocation schemes were designed in [12] for minimizing the Cramér-Rao bound of the SWIPT ISAC system, where PS was performed at the ISAC BS. Indeed, both TS and PS schemes take full advantage of the time slot or power of the RF signals, which can improve the EE of the NO-T ISAC systems. Furthermore, the introduction of SWIPT into the ISAC system can further save energy costs, meet the needs of green communication, smart energy and other communication scenarios, and improve EE.

Inspired by these works, the C&S performance of the SWIPT NO-DLT ISAC system is investigated in this paper, where the near user is regarded as a full-duplex EH relay that provides assistance to the far user and BS in terms of C&S. To be more practical, both residual hardware impairments (RHIs) and channel estimation errors (CEEs), as well as imperfect SIC (ipSIC) are taken into account in the system considered. Finally, the OP of the communication users, the probability of detection (PoD) and probability of false alarm (PoFA) of the BS, and the EE of the system are derived for characterizing the performance of the system considered.

2 System Model

As shown in Figure 1, we consider a SWIPT NO-DLT ISAC system, which consists one ISAC BS S, one near DL communication user D_n and one far DL communication user D_f. Specifically, S transmits signals to simultaneously detect target T and communicate with D_f and D_n. Furthermore, it is assumed that SWIPT can be performed at D_n, which indicates that the received signal at D_n is divided into two parts, one for information decoding (ID) and the other for EH. As a further step, D_n can also act as a full-
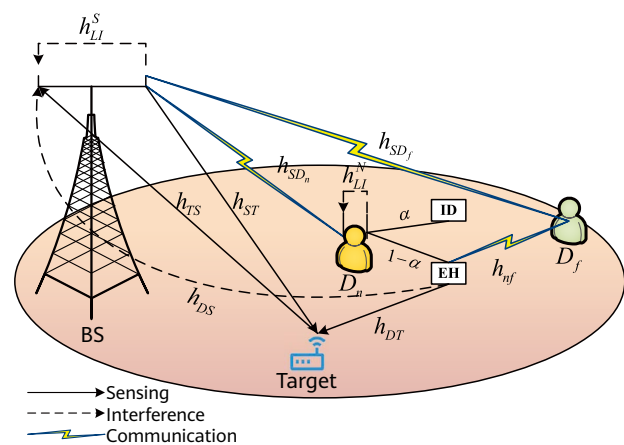

Figure 1 Illustration of the SWIPT NO-DLT ISAC system

duplex decoding-and-forwarding (DF) relay to assist communication for D_f and sensing for S by using the harvested energy. For the sake of convenience, we adopt h_{SD_n}, h_{SD_f}, and h_{nf} to denote the communication channels of the $S \to D_n$, $S \to D_f$, and $D_n \to D_f$, respectively. h_{ST}, h_{TS}, and h_{DT} denote the sensing channels of the $S \to T$, $T \to S$, and $D_n \to T$, respectively. Furthermore, h_{DS} is the interference channel of the $D_n \to S$, $w_{LI}^S \sim \mathcal{CN}\left(0, \varpi_{LI}^S\right)$ and $w_{LI}^N \sim \mathcal{CN}\left(0, \varpi_{LI}^N\right)$ represent the loop self-interference (LSI) channel caused by simultaneously transmitting and receiving signals at the BS and D_n, respectively [13]. For the considered system's operation, we make the following assumptions: (1) The signals transmitted from S and D_n can achieve both C&S functions [3]; (2) Both S and D_n operate with the full-duplex mode, while D_f operates with the half-duplex mode; (3) The unidirectional channels in the ISAC systems are considered to follow Nakagami-m fading [14].

2.1 Communication Model

For the direct links, the BS transmits the superimposed signals $y_{BS} = \sqrt{a_n P_S} x_n + \sqrt{a_f P_S} x_f$ to D_n and D_f, where P_S denotes the transmit power of the BS, a_n and a_f are the power allocation coefficients of D_n and D_f with $a_n < a_f$ and $a_n + a_f = 1$. Furthermore, x_n and x_f respectively represent the desired signals of D_n and D_f with $\mathbb{E}\left[|x_n|^2\right] = \mathbb{E}\left[|x_f|^2\right] = 1$. Therefore, the harvested energy at D_n can be expressed as

$$P_{EH} = (1 - \alpha) P_S |h_{SD_n}|^2, \tag{1}$$

where $\alpha \in [0, 1]$ denotes the PS ratio. Then, the received signal at D_n and D_f can be respectively expressed as

$$y_{SD_n} = \sqrt{\alpha} h_{SD_n} \left(y_{BS} + \eta_{SD_n}\right) + w_{LI}^N \sqrt{P_N} x_{LI}^N + n_{SD_n}, \tag{2}$$

$$y_{SD_f} = h_{SD_f} \left(y_{BS} + \eta_{SD_f}\right) + n_{SD_f}, \tag{3}$$

where $\eta_{SD_n} \sim \mathcal{CN}\left(0, \kappa_{SD_n}^2 P_S\right)$ and $\eta_{SD_f} \sim \mathcal{CN}\left(0, \kappa_{SD_f}^2 P_S\right)$ are the RHIs caused by the links of $S \to D_n$ and $S \to D_f$, while κ_{SD_n} and κ_{SD_f} denote the levels of RHIs at D_n and D_f, respectively. Furthermore, $P_N = \xi P_{EH}$ is the transmit power of D_n and $\xi \in [0, 1]$ denotes the energy conversion efficiency, x_{LI}^N represents the LSI signal at D_n with $\mathbb{E}\left[|x_{LI}^N|^2\right] = 1$, $n_{SD_n} \sim \mathcal{CN}(0, N_0)$ and $n_{SD_f} \sim \mathcal{CN}(0, N_0)$ are the additive white Gaussian noise (AWGN) at D_n and D_f, and N_0 is the noise power.

According to the NO-DLT protocol, only x_f is decoded at D_f. Therefore, the received signal-to-interference-plus-noise ratio (SINR) of x_f at D_f by the direct link is expressed as

$$\gamma_{SD_f}^{D_f} = \frac{a_f \rho_{SD_f} \gamma_S}{\left(a_n + \kappa_{SD_f}^2\right) \rho_{SD_f} \gamma_S + 1}, \tag{4}$$

where we have $\gamma_S = P_S / N_0$.

At D_n, x_f is first decoded, and then x_n is decoded by employing SIC. Hence, the received SINRs of x_f and x_n at D_n are respectively expressed as

$$\gamma_{SD_n}^{D_f} = \frac{\alpha a_f \rho_{SD_n} \gamma_S}{\left[\alpha a_n + \kappa_{SD_n}^2 + (1 - \alpha) \xi \varpi_{LI}^N\right] \rho_{SD_n} \gamma_S + 1}, \tag{5}$$

$$\gamma_{SD_n}^{D_n} = \frac{\alpha a_n \rho_{SD_n} \gamma_S}{\left[\kappa_{SD_n}^2 + \varepsilon \alpha a_f + (1 - \alpha) \xi \varpi_{LI}^N\right] \rho_{SD_n} \gamma_S + 1}, \tag{6}$$

where $\varepsilon \in [0, 1]$ denotes the ipSIC coefficient, which is caused by partial signal decoding failure due to the RHIs.

For the cooperative link, D_n forwards x_f to D_f by using the harvested energy. Thus, the observation at D_f by the cooperative link can be expressed as

$$y_{nf} = h_{nf} \left(\sqrt{P_N} x_f + \eta_{nf}\right) + n_{nf}, \tag{7}$$

where $n_{nf} \sim \mathcal{CN}(0, N_0)$ is the AWGN. Furthermore, $\eta_{nf} \sim \mathcal{CN}\left(0, \kappa_{nf}^2 P_N\right)$ denotes the RHI associated with the link of $D_n \to D_f$ and κ_{nf} is the level of RHI. Thus, the received SINR of x_f at D_f by the cooperative link is expressed as

$$\gamma_{nf}^{D_f} = \frac{(1 - \alpha) \xi \rho_{SD_n} \rho_{nf} \gamma_S}{\kappa_{nf}^2 (1 - \alpha) \xi \rho_{SD_n} \gamma_S + 1}. \tag{8}$$

2.2 Sensing Model

In the ISAC system considered, the received signals at the sensing receiver can be divided into four parts: (1) echoes associated with D_n and S; (2) LSI, RHIs, and CEEs; (3) interference from D_n; and (4) AWGN. Thus, the received signals at the BS can be expressed as

$$y_S = \delta \left[(h_{ST} h_{TS} + e_1) (y_{BS} + \eta_{SS}) + (h_{DT} h_{TS} + e_2) \tag{9}\right.$$
$$\left. \times \left(\sqrt{P_N} x_f + \eta_{DS}\right)\right] + h_{DS} \left(\sqrt{P_N} x_f + \eta_{DS}\right) + w_{LI}^S \sqrt{P_S} x_{LI}^S + n_S,$$

where $\eta_{SS} \sim \mathcal{CN}\left(0, \kappa_{SS}^2 P_S\right)$, $\eta_{DS} \sim \mathcal{CN}\left(0, \kappa_{DS}^2 P_N\right)$, $e_1 \sim \mathcal{CN}\left(0, \sigma_1^2\right)$, and $e_2 \sim \mathcal{CN}\left(0, \sigma_2^2\right)$ denote the RHIs and sensing CEEs associated with the links of $S \to T \to S$ and $D_n \to T \to S$, respectively. Furthermore, x_{LI}^S represents the LSI signal at S with $\mathbb{E}\left[|x_{LI}^S|^2\right] = 1$ and $n_S \sim \mathcal{CN}(0, N_0)$ is the AWGN.

Similar to the NO-DLT, SIC is performed at S for detecting the echoes [15]. From Equation (9), we can observe that the signal transmitted from D_n and passively received at S with the direct link of $D_n \to S$ is strong interference for sensing. To this end, the received SINR of echoes at S can be

expressed as

$$\gamma_S^{echo} = \frac{\delta^2 \rho_{ST} \rho_{TS} \gamma_S + \delta^2 \rho_{DT} \rho_{TS} \gamma_N}{\delta^2 \Phi \gamma_S + \Theta \gamma_N + \varpi_{LI}^S \gamma_S + 1}, \quad (10)$$

where $\Phi = \sigma_1^2 \left(1 + \kappa_{SS}^2\right) + \rho_{ST} \rho_{TS} \kappa_{SS}^2 + \rho_{DT} \rho_{TS} \kappa_{DS}^2$, $\Theta = \delta^2 \sigma_2^2 \left(1 + \kappa_{DS}^2\right) + \varepsilon \rho_{DS} \left(1 + \kappa_{DS}^2\right)$, and $\gamma_N = P_N / N_0$.

3 Performance Analysis

In this section, the exact and the asymptotic OPs of users for the communication, the PoD of the BS for the sensing, and the EE are derived for evaluating the ISAC system considered.

3.1 OP Analysis

OP of D_n. If x_n cannot be decoded at D_n by the link of $S \to D_n$, the outage event occurs at D_n. Thus, the OP of D_n can be expressed as

$$\mathrm{P}_{\mathrm{out}}^{D_n} = \mathrm{Pr}\left(\gamma_{SD_n}^{D_n} < \gamma_{\mathrm{thn}}\right), \quad (11)$$

where γ_{thn} denotes the SNR threshold of x_n.

Theorem 1. The exact OP of D_n is expressed as

$$\mathrm{P}_{\mathrm{out}}^{D_n} = 1 - \sum_{g_1=0}^{\alpha_{SD_n}} \frac{1}{g_1!} e^{-\frac{\varphi}{\beta_{SD_n}}} \left(\frac{\varphi}{\beta_{SD_n}}\right)^{g_1}, \quad (12)$$

where $(\cdot)!$ denotes the factorial operation. Furthermore, we have

$$\varphi = \frac{\gamma_{\mathrm{thn}}}{\alpha a_n \gamma_S - \left[\kappa_{SD_n}^2 + \varepsilon \alpha a_f + (1-\alpha) \xi \varpi_{LI}^N\right] \gamma_S \gamma_{\mathrm{thn}}}$$

with $\alpha a_n > \left[\kappa_{SD_n}^2 + \varepsilon \alpha a_f + (1-\alpha) \xi \varpi_{LI}^N\right] \gamma_{\mathrm{thn}}$. Otherwise we have $\mathrm{P}_{\mathrm{out}}^{D_n} = 1$.

Proof. Upon substituting Equation (6) into Equation (11), after some mathematical manipulations, the OP of D_n can be rewritten as

$$\mathrm{P}_{\mathrm{out}}^{D_n} = \mathrm{Pr}\left(\rho_{SD_n} < \varphi\right)$$
$$= \int_0^\varphi f_{\rho_{SD_n}}(x) dx. \quad (13)$$

By using the cumulative distribution function (CDF) of ρ_i, we can obtain Equation (12).

Corollary 1. The asymptotic OP of D_n is expressed as

$$\mathrm{P}_{\mathrm{out},\infty}^{D_n} = \frac{\varphi^{\alpha_{SD_n}}}{\alpha_{SD_n}! \beta_{SD_n}^{\alpha_{SD_n}}}. \quad (14)$$

The CDF of ρ_i in the high SNR regime is expressed as

$$F_{\rho_i}^\infty(x) = \frac{x^{\alpha_i}}{\alpha_i! \beta_i^{\alpha_i}}. \quad (15)$$

By substituting Equation (15) into Equation (13), Equation (14) can be obtained.

OP of D_f. The outage event occurs at D_f when (1) x_f cannot be decoded at D_n and D_f by the links of $S \to D_n$ and $S \to D_f$;

(2) x_f is decoded successfully at D_n by the link of $S \to D_n$, while it cannot be decoded at D_f by the links of $S \to D_f$ and $D_n \to D_f$. Thus, the OP of D_f can be expressed as

$$\mathrm{P}_{\mathrm{out}}^{D_f} = \mathrm{Pr}\left(\max\left(\gamma_{SD_n}^{D_f}, \gamma_{SD_f}^{D_f}\right) < \gamma_{\mathrm{thf}}\right)$$
$$+ \mathrm{Pr}\left(\gamma_{SD_n}^{D_f} \geq \gamma_{\mathrm{thf}}, \max\left(\gamma_{nf}^{D_f}, \gamma_{SD_f}^{D_f}\right) < \gamma_{\mathrm{thf}}\right), \quad (16)$$

where γ_{thf} denotes the SNR threshold of x_f.

Theorem 2. The exact OP of D_f is expressed as

$$\mathrm{P}_{\mathrm{out}}^{D_f} = \left[1 - \sum_{g_1=0}^{\alpha_{SD_f}-1} \frac{1}{g_1!} e^{-\frac{\phi}{\beta_{SD_f}}} \left(\frac{\phi}{\beta_{SD_f}}\right)^{g_1}\right]$$
$$\times \left[1 - \frac{e^{-\frac{\varsigma}{\beta_{nf}}}}{\Gamma(\alpha_{SD_n}) \beta_{SD_n}^{\alpha_{SD_n}}} \sum_{g_3=0}^{\alpha_{nf}-1} \sum_{t=0}^{g_3} \binom{g_3}{t} \frac{\varsigma^{g_3-t} \vartheta^t}{g_3!}\right. \quad (17)$$
$$\left. \times \left(\frac{1}{\beta_{nf}}\right)^{g_3} \int_\lambda^\infty y^{\alpha_{SD_n}-t-1} e^{-\frac{y}{\beta_{SD_n}} - \frac{\vartheta}{\beta_{nf} y}} dy\right],$$

where $\Gamma(\cdot)$ represents the Gamma function, $\vartheta = \frac{\gamma_{\mathrm{thf}}}{(1-\alpha)\xi\gamma_S}$ and $\varsigma = \kappa_{nf}^2 \gamma_{\mathrm{thf}}$. Furthermore, we also have

$$\lambda = \frac{\gamma_{\mathrm{thf}}}{\alpha a_f \gamma_S - \left[\alpha a_n + \kappa_{SD_f}^2 + (1-\alpha) \xi \varpi_{LI}^N\right] \gamma_S \gamma_{\mathrm{thf}}},$$
$$\phi = \frac{\gamma_{\mathrm{thf}}}{a_f \gamma_S - \left(a_n + \kappa_{SD_f}^2\right) \gamma_S \gamma_{\mathrm{thf}}},$$

with $\alpha a_f > \left[\alpha a_n + \kappa_{SD_f}^2 + (1-\alpha) \xi \varpi_{LI}^N\right] \gamma_{\mathrm{thf}}$ and $a_f > \left(a_n + \kappa_{SD_f}^2\right) \gamma_{\mathrm{thf}}$. Otherwise, we have $\mathrm{P}_{\mathrm{out}}^{D_f} = 1$.

Proof. Upon substituting Equations (4), (5), and (8) into Equation (16), after some mathematical manipulations, the OP of D_f can be written as

$$\mathrm{P}_{\mathrm{out}}^{D_f} = \mathrm{Pr}\left(\rho_{SD_f} < \phi\right)$$
$$\times \left[\mathrm{Pr}\left(\rho_{SD_n} < \lambda\right) + \mathrm{Pr}\left(\rho_{SD_n} \geq \lambda, \rho_{nf} < \vartheta/\rho_{SD_n} + \varsigma\right)\right]$$
$$= \int_0^\phi f_{\rho_{SD_f}}(y) \int_0^\lambda f_{\rho_{SD_n}}(x) dx dy$$
$$+ \int_0^\phi f_{\rho_{SD_f}}(z) \int_\lambda^\infty f_{\rho_{SD_n}}(y) \int_0^{\frac{\vartheta}{y}+\varsigma} f_{\rho_{nf}}(x) dx dy dz. \quad (18)$$

By using the probability density function (PDF) and CDF of ρ_i, Equation (17) can be obtained after some mathematical manipulations.

Corollary 2. The asymptotic OP of D_f is expressed as

$$\mathrm{P}_{\mathrm{out},\infty}^{D_f} = \frac{\phi^{\alpha_{SD_f}}}{\alpha_{SD_f}! \beta_{SD_f}^{\alpha_{SD_f}}} \left[\frac{\lambda^{\alpha_{SD_n}}}{\alpha_{SD_n}! \beta_{SD_n}^{\alpha_{SD_n}}} + \frac{1}{\alpha_{nf}! \beta_{nf}^{\alpha_{nf}}}\right.$$
$$\times \frac{1}{\Gamma(\alpha_{SD_n}) \beta_{SD_n}^{\alpha_{SD_n}}} e^{-\frac{\lambda}{\beta_{SD_n}}} \sum_{t_1=0}^{\alpha_{nf}-1} \sum_{t_2=0}^{\Xi} \binom{\alpha_{nf}-1}{t_1} \binom{\Xi}{t_2}$$
$$\left. \times \varsigma^{\alpha_{nf}-t_1} \vartheta^{t_1} \lambda^{\Xi-t_2} \beta_{SD_n}^{t_2+1} \Gamma(t_2+1)\right], \quad (19)$$

where we have $\Xi = \alpha_{SD_n} - t_1 - 1$.

Proof. The asymptotic OP of D_f in the high SNR regime can be written as

$$\mathrm{P}^{D_f}_{\mathrm{out},\infty} = F^{\infty}_{\rho_{SD_f}}(\phi) \left[F^{\infty}_{\rho_{SD_n}}(\lambda) + \int_{\lambda}^{\infty} f_{\rho_{SD_n}}(x) F^{\infty}_{\rho_{nf}}\left(\frac{\vartheta}{x} + \varsigma\right) dx \right]. \quad (20)$$

Upon substituting Equation (15) into Equation (20) and with the aid of [16, Eq. (3.326.2.10)], Equation (19) can be obtained after some mathematical manipulations.

3.2 PoD Analysis

Accurate estimating the existence of the target is the fundamental task of sensing equipment. Therefore, the sensing performance is characterized by deriving the PoD of the BS in this subsection. Specifically, the PoFA is caused by the interference and noise, while PoD is related to the echoes. Both PoFA and PoD can be mathematically formulated as $\mathrm{P}_{\mathrm{fa}} = \Pr(\mathcal{H}_1|\mathcal{H}_0)$ and $\mathrm{P}_{\mathrm{d}} = \Pr(\mathcal{H}_1|\mathcal{H}_1)$, where $\mathcal{H}_0$ and $\mathcal{H}_1$ represent the null and alternative hypothesis, respectively. As stated in Section 2.2, SIC is adopted at the sensing receiver, thus the direct-path interference of $D_n \to S$ can be eliminated. Considering ipSIC, according to Equation (9), we can obtain the following expressions

$$\mathcal{H}_0 : y_S = \delta \left[e_1 (y_{BS} + \eta_{SS}) + e_2 \left(\sqrt{P_N} x_f + \eta_{DS} \right) \right]$$
$$+ h_{DS} \left(\sqrt{\varepsilon P_N} x_f + \eta_{DS} \right) + w^S_{LI} x^S_{LI} + n_S,$$
$$\mathcal{H}_1 : y_S = \delta \left[(h_{ST} h_{TS} + e_1)(y_{BS} + \eta_{SS}) + (h_{DT} h_{TS} + e_2) \right.$$
$$\left. \times \left(\sqrt{P_N} x_f + \eta_{DS} \right) \right] + h_{DS} \left(\sqrt{\varepsilon P_N} x_f + \eta_{DS} \right) + w^S_{LI} \sqrt{P_S} x^S_{LI} + n_S. \quad (21)$$

From Equation (21), we can observe that the received power at the BS is the non-central Chi-square random variable with three and four degrees of freedom (DoFs) under $\mathcal{H}_0$ and $\mathcal{H}_1$. Therefore, the PoFA and PoD of the ISAC system considered are respectively expressed as

$$\mathrm{P}_{\mathrm{fa}} = Q_{\frac{3}{2}} \left(\sqrt{2 P^{\mathrm{fa}}_e / P^{\mathrm{fa}}_{i+n}}, \sqrt{2\psi / P^{\mathrm{fa}}_{i+n}} \right), \quad (22)$$

$$\mathrm{P}_{\mathrm{d}} = Q_2 \left(\sqrt{\frac{2\delta^2 \rho_{TS}(P_S \rho_{ST} + P_N \rho_{DT})}{P^{\mathrm{d}}_{i+n}}}, \sqrt{\frac{2\psi}{P^{\mathrm{d}}_{i+n}}} \right), \quad (23)$$

where ψ and $Q(\cdot, \cdot)$ are the detection threshold and Marcum Q-function [17], respectively. Furthermore, P^{fa}_{i+n}, P^{fa}_{i+n}, and P^{d}_{i+n} are respectively given by

$$P^{\mathrm{fa}}_e = \delta^2 \left[(\sigma^2_1 + \varpi^S_{LI}) P_S + (\varepsilon \rho_{DS} + \sigma^2_2) P_N \right], \quad (24)$$

$$P^{\mathrm{fa}}_{i+n} = \rho_{DS} \kappa^2_{DS} P_N + \varpi^S_{LI} P_S + N_0, \quad (25)$$

$$P^{\mathrm{d}}_{i+n} = \delta^2 \left[\kappa^2_{SS} P_S \rho_{ST} \rho_{TS} + \sigma^2_1 P_S (1 + \kappa^2_{SS}) + \kappa^2_{DS} P_N \rho_{DT} \rho_{TS} \right.$$
$$\left. + \sigma^2_2 P_N (1 + \kappa^2_{DS}) \right] + \varepsilon (1 + \kappa^2_{DS}) \rho_{DS} P_N + \varpi^S_{LI} P_S + N_0. \quad (26)$$

3.3 EE Analysis

In this section, EE is investigated for further evaluating the ISAC system considered, which can be expressed as [18]

$$\eta_{EE} = \frac{R_{\mathrm{sum}}}{Q_E}, \quad (27)$$

where $R_{\mathrm{sum}} = R_n + R_f + R_s$ and $Q_E = P_S + P_N + P_C$ denote the sum rate and total energy consumption of the system. Furthermore, P_C is the fixed energy consumption of the system, R_n and R_f denote the communication rates of D_n and D_f, while R_s is the sensing rate [14], which are respectively given by

$$R_n = \log_2 \left(1 + \gamma^{D_n}_{SD_n} \right), \quad (28)$$

$$R_f = \min \left[\log_2 \left(1 + \gamma^{D_f}_{SD_n} \right), \log_2 \left(1 + \gamma^{D_f}_{SD_f} + \gamma^{D_f}_{nf} \right) \right], \quad (29)$$

$$R_s = \log_2 \left(1 + \gamma^{echo}_S \right). \quad (30)$$

4 Simulation Results

In this section, the accuracy of the theoretical analysis is demonstrated by Monte-Carlo computer simulations. Unless otherwise mentioned, the simulation parameters are set as follows: $\chi = \tau = 2$, $a_f = 0.7$, $a_n = 0.3$, $\gamma_{\mathrm{thf}} = 1$ dB, $\gamma_{\mathrm{thn}} = 1.5$ dB, $\alpha = 0.7$, $\xi = 0.9$, $\delta = 0.5$, $N_0 = 1$. The distance between the nodes is normalized as: $d_{SD_n} = 0.7$, $d_{SD_f} = 1$, $d_{nf} = 0.5$, $d_{ST} = d_{TS} = 1$, $d_{DT} = 0.8$, $d_{DS} = 0.6$.

Figure 2 plots the OP of the communication users versus the SNR both in ideal ($\kappa = \varepsilon = 0$) and non-ideal ($\kappa = \varepsilon = 0.1$) conditions. The OP of the users is also presented without the assistance of SWIPT. Obviously, with the aid of SWIPT, the outage performance of D_f can be greatly improved, while the OP of D_n increases which is due to that D_n forwards part of the harvested energy to D_f. Compared with the scheme without SWIPT, the proposed scheme can further enhance the fairness of communication users. Additionally, it can also be observed that RHIs and ipSIC weaken the outage performance.
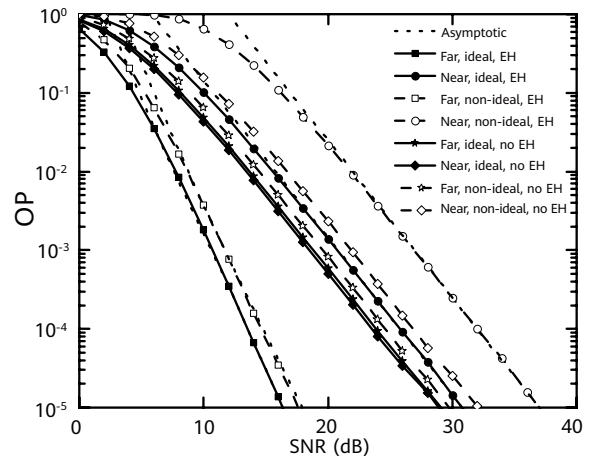

Figure 2 OP of the communication users vs. SNR

Figure 3 shows the detection performance of the BS versus the SNR both in ideal and non-ideal conditions with $P_{fa} = 10^{-6}$. For the convenience of comparison, unique variables are controlled. From Figure 3, we can observe that both RHIs and CEEs as well as ipSIC destroy the detection performance of the ISAC system considered, which is because the detection threshold is increased by the non-ideal factors at a fixed PoFA. Specifically, both RHIs and CEEs intuitively increase the received power at sensing receiver under $\mathcal{H}_0$. While for ipSIC, the interference caused by the direct link $D_n \rightarrow S$ cannot be completely eliminated, which further increases the received power at the BS indirectly under $\mathcal{H}_0$. Moreover, it can be concluded that the PoD of the system considered is more sensitive to RHIs compared with CEEs and ipSIC.

Figure 4 EE of the system vs. SNR

5 Conclusion

In this paper, the performance of the SWIPT NO-DLT ISAC system was developed and analyzed in the presence of RHIs, CEEs, and ipSIC. Specifically, the exact and asymptotic OP of the users, the PoD and PoFA of the BS were derived for characterizing the C&S performance, respectively. Furthermore, the EE of the system considered was presented as well. The simulation results show that RHIs, CEEs, and ipSIC limit both C&S performance. By contrast, owing to SWIPT, where the near user harvests energy to assist the far user communication and the BS sensing, the fairness of the communication users, PoD of the BS, and EE of the system are improved over that of the pure ISAC scheme.

Figure 3 PoD of the BS vs. SNR

Figure 4 illustrates the EE of the SWIPT ISAC system considered versus the SNR in both ideal and non-ideal conditions with $P_C = 5$ dBm. From Figure 4, we can observe that with the increase of SNR, the EE of the system first increases and then decreases. This phenomenon can be attributed to the fact that in the low and medium SNR regimes, the sum rate of the system considered increases rapidly with the increase of SNR. Nevertheless, as SNR continuous increasing, the increase of sum rate gradually slows down and eventually tends to be stable in the high SNR regime. In addition, the EE of the ISAC system considered is enhanced with the assistance of SWIPT, which is due to that the energy harvested at D_n is adopted for target detection as well.

References

[1] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3834–3862, Jun. 2020.

[2] Z. Wei, F. Liu, C. Masouros, N. Su, and A. P. Petropulu, "Toward multi-functional 6G wireless networks: Integrating sensing, communication, and security," *IEEE Commun. Mag.*, vol. 60, no. 4, pp. 65–71, Apr. 2022.

[3] F. Liu, C. Masouros, A. Li, H. Sun, and L. Hanzo, "MU-MIMO communications with MIMO radar: From co-existence to joint transmission," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2755–2770, Feb. 2018.

[4] B. Chang, W. Tang, X. Yan, X. Tong, and Z. Chen, "Integrated scheduling of sensing, communication, and control for mmWave/THz communications in cellular connected UAV networks," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2103–2113, Jul. 2022.

[5] J. Mu, Y. Gong, F. Zhang, Y. Cui, F. Zheng, and X. Jing, "Integrated sensing and communication-enabled predictive beamforming with deep learning in vehicular networks," *IEEE Commun. Lett.*, vol. 25, no. 10, pp. 3301–3304, Oct. 2021.

[6] Y. Liu, S. Zhang, X. Mu, Z. Ding, R. Schober, N. Al-Dhahir, E. Hossain, and X. Shen, "Evolution of NOMA toward next generation multiple access (NGMA) for 6G," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1037–1071, Apr. 2022.

[7] C. Zhang, W. Yi, Y. Liu, and L. Hanzo, "Semi-integrated sensing-and-communication (Semi-ISaC): From OMA to NOMA," *IEEE Trans. Commun.*, *Early Access*, pp. 1–16, Feb. 2023.

[8] C. Ouyang, Y. Liu, and H. Yang, "On the performance of uplink ISAC systems," *IEEE Commun. Lett.*, vol. 26, no. 8, pp. 1769–1773, Aug. 2022.

[9] Y. Yuan, Y. Xu, Z. Yang, P. Xu, and Z. Ding, "Energy efficiency optimization in full-duplex user-aided cooperative SWIPT NOMA systems," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5753–5767, Aug. 2019.

[10] R. R. Kurup and A. V. Babu, "Power adaptation for improving the performance of time switching SWIPT-based full-duplex cooperative NOMA network," *IEEE Commun. Lett.*, vol. 24, no. 12, pp. 2956–2960, Dec. 2020.

[11] L. Dai, B. Wang, M. Peng, and S. Chen, "Hybrid precoding-based millimeter-wave massive MIMO-NOMA with simultaneous wireless information and power transfer," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 1, pp. 131–141, Jan. 2019.

[12] X. Li, X. Yi, Z. Zhou, K. Han, Z. Han, and Y. Gong, "Multi-user beamforming design for integrating sensing, communications, and power transfer," in 2023 *IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2023, pp. 1–6.

[13] T. M. Hoang, L. T. Dung, B. C. Nguyen, X. N. Tran, and T. Kim, "Secrecy outage performance of FD-NOMA relay system with multiple non-colluding eavesdroppers," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 12 985–12 997, Dec. 2021.

[14] C. Zhang, W. Yi, and Y. Liu, "Semi-integrated-sensing-and-communication (Semi-ISaC) networks assisted by NOMA," in *ICC 2022 - IEEE International Conf. Commun.*, May 2022, pp. 3790–3795.

[15] A. R. Chiriyath, B. Paul, and D. W. Bliss, "Simultaneous radar detection and communications performance with clutter mitigation," in *2017 IEEE Radar Conf. (RadarConf)*, Jun. 2017, pp. 0279–0284.

[16] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals*, *Series*, *and Products*, *7th ed*. New York, NY, USA: Academic, 2007.

[17] F. Wang, H. Li, and M. A. Govoni, "Power allocation and codesign of multicarrier communication and radar systems for spectral coexistence," *IEEE Trans. Signal Process.*, vol. 67, no. 14, pp. 3818–3831, Jul. 2019.

[18] X. Li, J. Li, Y. Liu, Z. Ding, and A. Nallanathan, "Residual transceiver hardware impairments on cooperative NOMA networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 680–695, Jan. 2020.
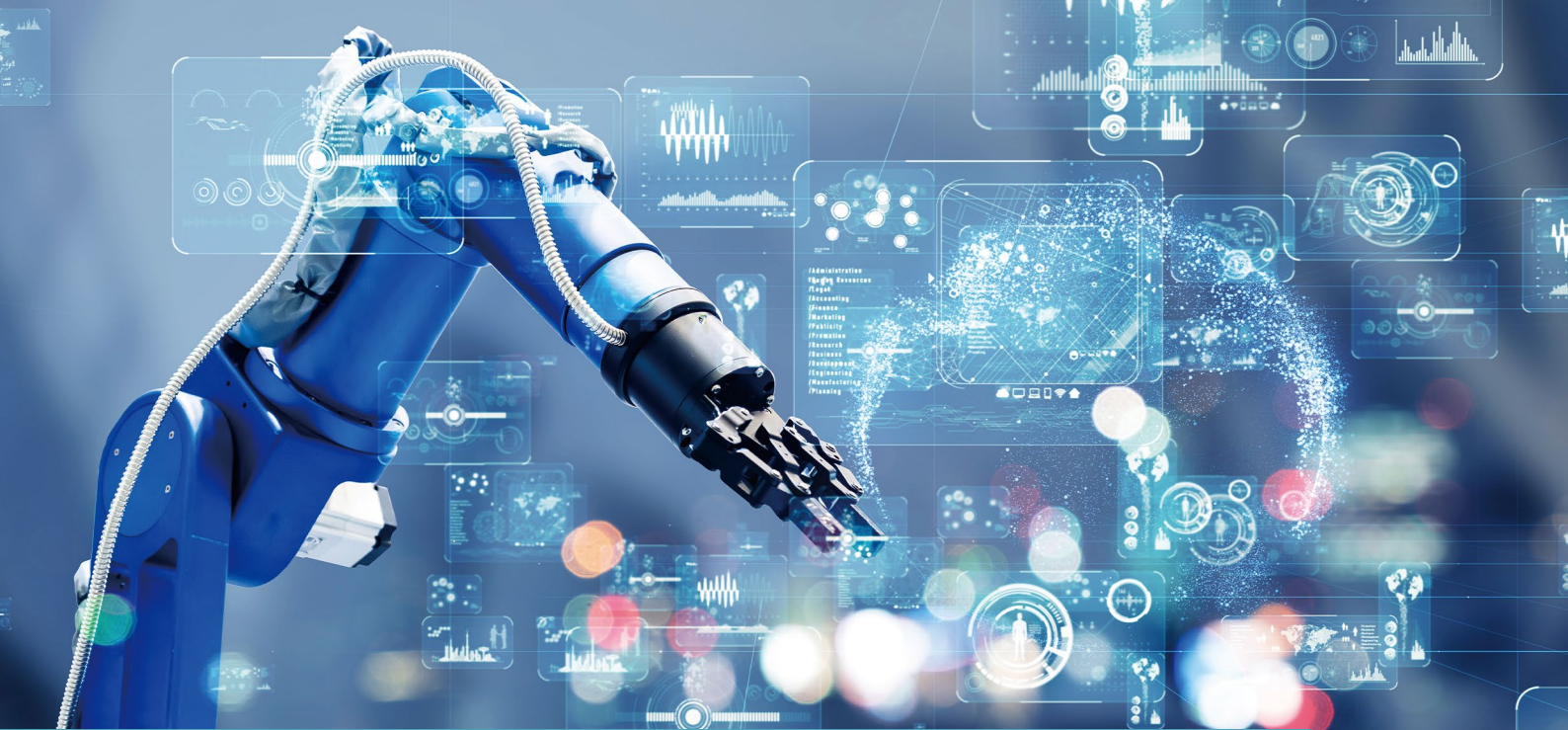
Integrated Sensing and Communication for Vertical Industry and Healthcare Applications

Guangjian Wang [1], Qiao Liu [1], Junwei Zang [1], Jia He [1], Ziming Yu [1], Shibin Ge [1], Jianan Zhang [1], Tao Wan [1], Oupeng Li [1], Xianfeng Du [1], Yao Liu [1], Jianhang Cui [1], Guolong Huang [1], Junping Zhang [1], Lei Lu [2], Yan Chen [3]

[1] Wireless Technology Lab
[2] Munich Research Center
[3] Ottawa Advanced Wireless Technology Lab

Abstract

This paper discusses the evolution of integrated sensing and communication (ISAC) technologies for vertical industry and healthcare applications. We look at the services and applications that ISAC technologies have the potential to empower, including smart factories, transportation, homes, and health monitoring and treatment. We also discuss the potential impact of ISAC technologies in these vertical industries. There are several key challenges to the practical commercial implementation of ISAC technologies. For example, to evaluate the performance of ISAC solutions, a theoretical framework is necessary so that tradeoffs and metrics can be applied in design. In addition to theories, the coverage requirements of the target service, air interface, joint waveform optimization and design, and hardware imperfections also require careful consideration. We propose multiple corresponding solutions, demonstrating prototypes through extensive trials and analysis from theoretical and practical perspectives. Our research concludes that ISAC technologies will become a crucial component in the 6G system; they will significantly boost applications in vertical industries and healthcare applications, providing society with connected intelligence.

1 Introduction

Along with the evolution of integrated sensing and communication (ISAC) technologies, this paper looks at how ISAC technologies can boost vertical industry and healthcare applications, elevating them toward the "connected intelligence" horizon [1] in the 6G era. ISAC could empower various industries through intelligence, including but not limited to smart factories and transportation, advanced healthcare applications, intelligent offices and homes, etc.

Despite the promising visions, we are still inching toward a commercially available intelligent society due to several technical challenges in ISAC for vertical applications, such as integrated signal processing and sensing capabilities. However, we have noted several inspiring achievements over the years. Breakthroughs have been made in theories and practice, and multiple prototypes have been developed and tested for different vertical applications in 6G systems. The results of these prototypes have demonstrated their ability to provide versatile services with enhanced performance.

The sections of this paper are arranged as follows. In Section 2, we discuss vertical services and applications. Section 3 covers the key challenges and technologies. Section 4 analyzes the verification of several ISAC capabilities for versatile vertical applications, multiple prototypes, and test results. Lastly, this paper is concluded in Section 5.

2 Vertical Services and Applications

As introduced in Section 1, sensing will become a new function integrated into the 6G communication system, making 6G an ISAC system. In this ISAC system, the sensing capability can turn 6G base stations, 6G devices, and even the entire 6G network into sensors. Some of the new sensing capabilities in 6G may include positioning, gesture and activity recognition, imaging, and mapping [1]. These new sensing capabilities will enable new services for vertical services and applications, for example, smart factories, smart transportation, healthcare, smart homes, and environment sensing (Figure 1). This section will introduce the vertical services and applications enabled by the 6G ISAC system.

- **High-Accuracy Positioning for Smart Factories**

  A high-accuracy 3D positioning system enables fully automated robots. This high-accuracy 3D positioning system provides both communication and sensing,
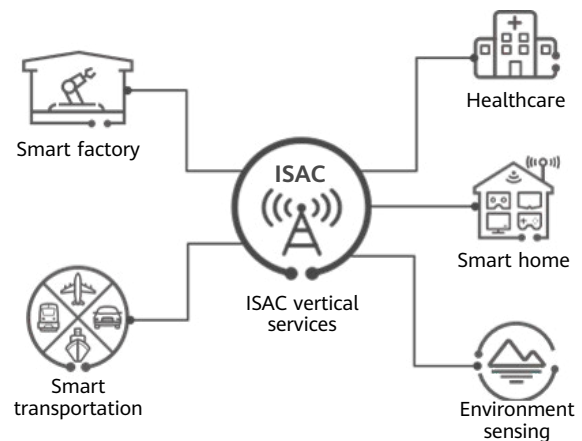
Figure 1 Vertical services and applications of ISAC

including advanced positioning and navigation, and mapping services, to fully automate robots in a smart factory [2]. It serves as a cyber-physical bridge for physical facilities, materials, processed parts, and digital manufacturing workflows. This system allows business partners to monitor, diagnose, and optimize the manufacturing process anywhere and anytime. Small parts are stored in a 3D stacked warehouse to save space. Automated guided vehicle (AGV) robots can easily retrieve goods from the warehouse and load them into the production plant with a very high precision soft vertical landing (Figure 2). The system also involves accurate 3D positioning, which prevents parts from being damaged while lifted. AGV robots can efficiently and quickly transport parts; therefore, they need to avoid colliding with each other and obstacles using real-time positioning. Millisecond-level latency and millimeter-level accuracy enable real-time dynamic path planning.

High-accuracy relative positioning enables smart security in factories. Safety is the prerequisite for industrial production. Technologies and general awareness have significantly reduced the number of safety accidents in factories. However, incidents still occur occasionally, such as explosions of dangerous equipment like boilers, operational accidents, wall collapses, and other incidents. This threatens the lives of tens of thousands of people. Detection and pre-maintenance are key to preventing security incidents. Current detection and pre-maintenance mainly depend on internal sensors and security

Figure 2 High-accuracy positioning for robots

personnel. Comprehensive sensor detection is costly because it needs many sensors to cover the detected object. Manual real-time detection is also costly. An ISAC system facilitates real-time, comprehensive detection with millimeter relative positioning for the detection/ pre-maintenance of hazardous equipment facilities, anti-drop detection of articles, warning of personnel injuries, and so on. This enables smart security factories with zero safety accidents. Relative positioning does not need to calculate the absolute coordinate location of the target and only needs to detect the relative position movement of the target.

- **Simultaneous Imaging, Mapping, and Positioning for Smart Transportation**

Cars are evolving from assisted to autonomous driving [3]. An important requirement of smart transportation is real-time environment reconstruction, where real-time environmental data can be obtained to achieve virtual roads and smart transportation. With this technology, vehicles can observe minute environmental details in real time, improving the efficiency and safety of autonomous driving.

For instance, traffic lights, road signs that are being replaced, and crowd flows can all be sensed and imaged during urban driving. Based on the sensed data, vehicles can avoid crowded roads and improve driving efficiency. This technology is also important for autonomous driving in suburban areas. For example, a fully autonomous vehicle moves along a street with minimal lighting at night and a deer suddenly dashes across the road. The autonomous vehicle immediately senses the situation and acts accordingly to avoid injuring the deer and the passenger.

- **Augmented Human Sensing for Healthcare**

Ultra-high resolution supports assisted surgeries. In the 6G system, sensing techniques will support high-resolution imaging and detection. Coupled with the 6G network, they will enable numerous healthcare applications, such as remote surgery, cancer diagnostics, etc. For this type of application, millimeter-level, ultra-high, cross-range resolution is required, which relies on higher bandwidth and increased antenna array aperture. These augmented human sensing functions can be integrated or installed on portable devices using 6G communication technologies with a high THz frequency and corresponding short wavelength of less than 1 mm.

For example, let's discuss telesurgery (Figure 3). A

Figure 3 Examples of ISAC for telemedicine

surgeon can conduct surgery remotely using an ultra-high-resolution imaging monitoring system and a remote operation platform system. During the surgery, the surgeon wears sensing glasses with ultra-high resolution. This sensing equipment helps the surgeon examine the patient with millimeter-level resolution, reducing the possibility of permanent damage during the surgery.

None-line-of-sight (NLOS) imaging supports health monitoring. Conventional optical imaging technology can only provide line-of-sight (LOS) imaging at a level like the human eye observation. However, 6G-based sensing can achieve NLOS imaging capability. With NLOS imaging, 6G sensing can support health monitoring. The number of people living with chronic diseases is rapidly increasing worldwide and burdening healthcare systems. Chronic diseases (asthma, arrhythmia, hypoglycemia, hyperglycemia, chronic pain, etc.) must be constantly monitored using wearable devices in the long term. For example, the heartbeats of a patient suffering from heart disease need to be monitored 24 hours a day, and the blood sugar level of a patient suffering from hypoglycemia or hyperglycemia needs to be monitored after every meal. Fortunately, 6G sensing technology enables atraumatic medical detection. It detects the heartbeat and can obtain a real-time image of the heart's bloodstream. This technology can help regular consumers monitor their health in an affordable and convenient way. The atraumatic detection technology will not harm or cause discomfort to patients and features ultra-high reliability and accuracy.

- **Gesture and Activity Recognition for Smart Homes**

Gesture and activity recognition based on the 6G ISAC system supports interactive applications and intelligent systems in smart homes. For example, NLOS sensing is possible thanks to the propagation characteristics of 6G signals, enabling nighttime applications. Another important benefit of gesture and activity sensing in 6G is that it does not involve any underlying privacy risks (regarding video surveillance), which can be unsuitable in certain use cases (within homes).

Macro activity recognition facilitates elderly care. Macro activity recognition by the 6G ISAC system ensures that the movements of an older adult are effortlessly always supervised. For example, the health condition of an older adult is continuously monitored by detecting if they fall, have any irregular movements, or sneeze and cough. All this is critical to ensuring their safety. A fall would trigger a stronger sensing response than everyday gestures. With powerful macro activity recognition (2 cm range resolution, 0.05 m/s velocity resolution, and > 99.9% detection probability), the 6G ISAC system can easily detect this abnormal state and issue an alarm to a remote family doctor or the person's legal guardian.

Micro activity recognition supports smart home appliances. Future smart homes will have an advanced hand gesture capturing and recognition system. The system will allow us to track a hand's 3D position, rotation, and gestures. As a result, by simply waving our hands, we will be able to turn the lights on or off anywhere in the house without having to hit a switch. Besides lights, we will also be able to control smart TVs through hand gestures using 6G signals. The main interaction with the smart TV system will be performed with facial expressions and natural hand gestures. Put differently, a facial recognition system identifies the user, and the hand gesture recognition system controls the TV (channels and volume). Since 6G signals capture the hand gestures, the user does not need to be in front of the TV. To precisely capture a person's gestures, a range resolution of 1 cm, velocity resolution of 0.05 m/s, detection probability over 99%, and a coverage range of up to eight meters (the size of a large living room) are required. In addition, to separate the gestures of multiple people in front of the TV, a cross-range resolution of 5 cm or less is ideal.

- **Spectrogram Recognition for Smart Environment Sensing**

   Spectrogram recognition is based on identifying targets through the spectrogram, sensing their electromagnetic or photonic characteristics. This includes the analysis of absorption, reflectivity, and permittivity parameters, which supports environment sensing (including evaluating air quality and detecting pollution). For example, the PM2.5 transmission spectrum has two distinct absorption bands between 2.5 and 7.5 THz [4]. Correlations between the absorption bands and cross-peaks in the synchronous and asynchronous plots indicate that

metallic oxides absorption falls within the 2.5–7.5 THz range. These results verify that the THz spectral analysis of PM2.5 using the 6G ISAC system will be a promising tool that can be utilized to understand the composition and mass of pollutants.

The vertical services and applications described above are only some of the use cases enabled by the 6G ISAC system. For the 6G network, higher frequency bands, wider bandwidths, and massive antenna arrays will help sensing and imaging solutions with a very high resolution and accuracy. This will play a vital role in providing enhanced solutions in many fields, such as public safety and critical asset protection, health and wellness monitoring, intelligent transportation, smart homes and factories, gesture and activity recognition, air quality measurement, and gas/toxicity sensing.

3 Key Challenges and Technologies

3.1 Key ISAC Challenges for Vertical Applications

ISAC system design is in its infancy, and major challenges still exist in the practical implementation of this system. Despite this, ISAC is identified as a core technology in future communication systems, and a significant amount of new research over the next few years will enable an optimized, native design for the 6G system. Some of the significant challenges and key ISAC research directions are discussed below.

- **Fundamental tradeoffs and new evaluation metrics**

   A theoretical framework is necessary to analyze and evaluate the performance of current ISAC solutions to identify the benefits and detect and address any shortcomings. The first step is to define ISAC as a general framework in which any radio frequency (RF) signal can convey both communication and sensing data. Since both types of data are communicated using a single RF source, a fundamental tradeoff seems to exist
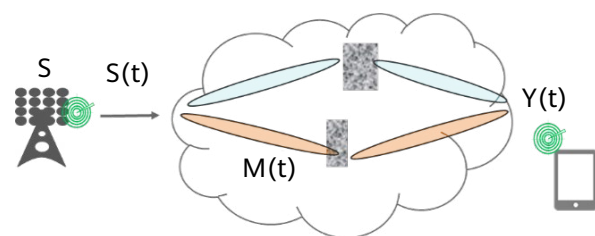

Figure 4 Theoretical framework for ISAC

between the communication and sensing performance. Therefore, a unified upper bound and its achievability will help characterize this tradeoff, which will provide the design criteria for the ISAC network.

- **Coverage requirements:** A traditional communication channel adopts bistatic transmission, in which the path loss is inversely proportional to the square of the distance between the transmitter and receiver. For sensing, the transmitter-target-receiver propagation path is critical, which contains two-hop spatial propagation paths and the target scattering characteristics. In this situation, the path loss is inversely proportional to the square of the distance of each two paths. Therefore, the sensing system suffers a larger path loss at the same coverage distance than the communication system. This means that the ISAC system has different link budget requirements. In other words, the base station needs to consider a new transmit power strategy and high dynamic hardware architecture to achieve the same sensing and communication coverage.

- **Hardware imperfections:** The current design of the ISAC system requires baseband and RF hardware to be functionally shared since the integrated hardware solution reduces the overall power consumption, system size, and information exchange latency between the two systems. In addition, converging the hardware would facilitate the mutually aided functions of sensing and communication in distortion calibration and compensation. Considering the cost and size difference between conventional communication and radar systems, the ISAC system hardware design would be closer to the traditional communication architecture [5]. As a tradeoff, we need to consider the impact of distortion parameters on sensing performance. Proper design of integrated RF architecture with self-interference cancellation is important for the ISAC system.

- **Joint waveform design and optimization:** The main challenge for joint waveform design is that communication and sensing have very different KPIs. Communication performance is based on maximizing the spectral efficiency, while the optimum waveform design for sensing hinges on improving target estimation resolution and accuracy, including MIMO and large-bandwidth friendly waveform OFDM [6] as well as the more traditional radar waveform frequency-modulated continuous wave (FMCW). Current state of the art suggests that there is still room for waveform design to balance good communication and sensing performance.

- **Integrated signal processing:** In a traditional communication system, typical signal processing technologies, such as channel estimation and demodulation, are only implemented at the receiver side. However, self-interference cancellation needs to be considered for monostatic sensing. Moreover, integrated transmission beamforming, signal fusion for collaborative sensing, high-accuracy positioning, and tracking, imaging, and reconstruction technologies are critical for industry ISAC applications.

3.2 ISAC Air Interface Design

In the design of the ISAC system, baseband and RF hardware sharing are crucial. Integrating hardware reduces the power consumption, system size, and information exchange latency between the two systems. The integration solution reduces overheads such as time, frequency, space, and code resources. On the other hand, this approach complicates the joint design of the air interface due to hardware distortion such as phase noise and self-interference mentioned in the previous section.

Most of the works on ISAC's joint air interface design focus on joint waveform design [7, 8]. The main challenge
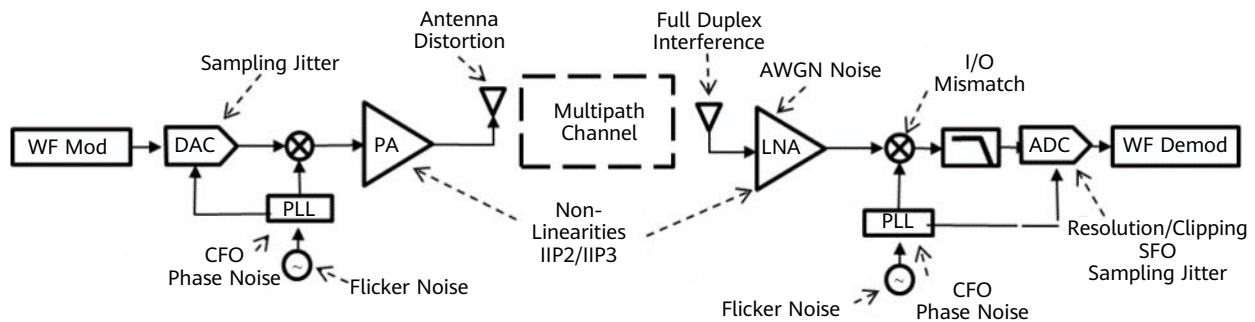

Figure 5 Hardware imperfections of the ISAC system

in joint waveform design is the contradicting KPIs for communication and sensing. In particular, the main target for communication is maximizing spectral efficiency, whereas the optimum waveform design for sensing is focused on estimation resolution and accuracy. Because cyclic prefix orthogonal frequency division multiplexing (CP-OFDM) has been proven to be a favorable option for communication, many researchers have also considered this waveform for sensing. However, CP-OFDM's large peak-to-average power ratio is another major issue for extreme long-range sensing where power efficiency is very important. Even so, FMCW, which has traditionally been used for radar, cannot carry data at the transmission rates desirable for communication services. Some researchers proposed to modify FMCW to make it more communication friendly. Nevertheless, these waveforms still suffer from low spectral efficiency.

Furthermore, the ISAC air interface design should consider multiplexing pilot and sequence design with dual functions to save overheads. Generally, a pilot for communication is designed to target a relatively higher signal-to-noise ratio. However, due to the limited transmit power of the ISAC system and radar cross section (RCS) of the interested target, the sensing pilot usually experiences a very low signal-to-noise ratio (SNR). Sensing detection needs to accumulate long-term coherent signals to ensure performance at a lower SNR, which makes the system more sensitive to sampling jitter, frequency offset, and phase noise [9]. This leads to higher requirements for the synchronization and stability of the system. In short, we need to consider these challenges in the selection of ISAC waveforms, which further effect the sensing algorithms and impairment compensation schemes. An air interface design that carefully combines the two functions can strike a balance between good communication and sensing performance.

3.3 ISAC High Resolution Design

The challenge of ISAC design is to obtain high accuracy sensing performance in terms of the communication architecture, cost, and power consumption. Especially, thousands of antenna elements are required to create a large aperture for high cross-range resolution. However, physically packing thousands of antenna elements into the portable device is infeasible due to the size and power constraints of the device. To solve this problem, virtual aperture techniques are applied in the prototype system. In particular, we propose using a virtual MIMO antenna

array in the hardware transceiver architecture with a sparse sampling design in the scanning process.

A major challenge for the virtual aperture technique is the irregular trajectory caused by the user moving the ISAC imaging module to scan an object. As a result, the sensing signal samples are sparse, causing high and non-uniform sidelobe effects, giving rise to false artifacts, which may lead to imaging failure.

On the other hand, to speed up the formation of virtual apertures, a multi-channel imaging algorithm based on MIMO is applied. The multi-channel imaging process can be treated as a time-domain coherent combination of electromagnetic signals from multiple receiving channels. Theoretically, n receivers can reduce the sampling time to 1/n compared with one receiver with the same imaging quality. Faster sampling will reduce the difficulty of motion error compensation, which, in turn, will improve the imaging quality. However, due to the inconsistency of multi-channel hardware, the gain and delay of different receiver channels are unbalanced. To solve this problem, we use an autofocus-based optimization algorithm to search for and compensate for multi-channel imbalances.

4 Technology Prototype and Demonstration

4.1 Melisac System for Smart Factories, Homes, and Healthcare

Huawei developed the ISAC concept to enable a network-assisted general-purpose mobile robot called "Melisac", showcased at the Mobile World Congress 2023 (MWC 2023). Melisac is a general-purpose robot and can be a form of future user terminal. We expect it to be widely used in Industry 5.0, smart hospitals, and smart homes.

Melisac integrates optical wireless and terahertz wireless communication.

The integrated sensing and communication with optical wireless (ISAC-OW) implements high-speed optical wireless communication, centimeter-level indoor positioning precision, and contactless health monitoring using an integrated waveform, a single hardware architecture, and our signal processing algorithms [10]. In the positioning process, the terminal uses an enhanced reflective surface to independently reflect lights coming from each base
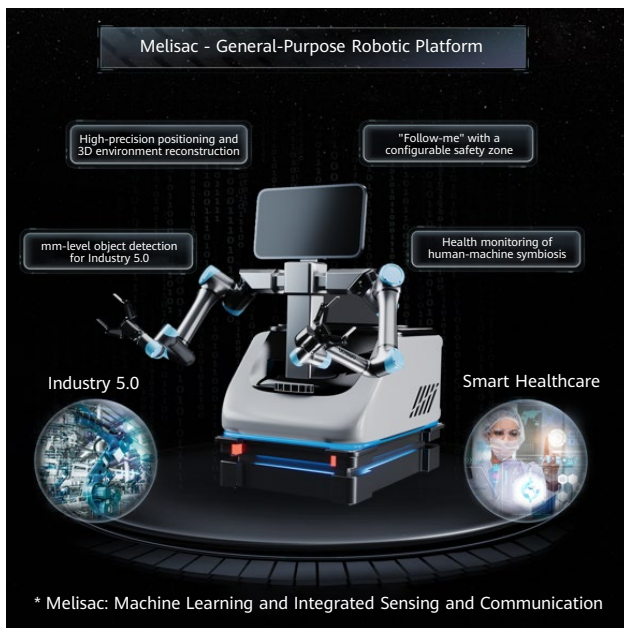
Figure 6 Melisac: ISAC prototype

Figure 7 Remote health monitoring: ISAC prototype

station, avoiding the interference from the reflected light to other base stations. Then, multiple base stations can measure phase differences of the reflected light without synchronization, achieving high positioning precision [11].

The contactless health monitoring prototype uses deep learning algorithms and digital signal processing technologies [12]. Based on the impact of a person's heartbeat on the blood flow in their face, the prototype can accurately measure the subtle changes in the light intensity reflected by a person's in different frequency bands, and then use the noise reduction algorithm to further smooth out the light signals. The monitoring accuracy of the pulse, blood oxygen, and respiratory rate (the respiratory rate can also be measured by the vibrations of the user's stomach during breathing) is similar to a contact health detection device (such as smart watches and fingertip blood oxygen meters).

The prototype uses a photodiode (PD) array at the receiver side and an infrared LED array at the transmitter side. The optical wireless communication and positioning system mainly uses the near infrared frequency band with a wavelength of 825–850 nm, and the contactless health monitoring system uses the visible and near infrared frequency bands with a wavelength of 400–850 nm. The PD array in the prototype senses four different bands of light: near-infrared (825–850 nm), red light (620–700 nm), green light (500–570 nm), and blue light (450–475 nm). Combining these four different frequency bands of light, digital signal processing estimates the user's pulse and

respiratory rate without contact. In addition, by monitoring the intensity of blue light and near-infrared light reflected by the bare skin of the user, the user's blood oxygen can also be monitored without contact. In the prototype, the optical wireless communication and positioning system mainly uses infrared light of 825–850 nm. The signal for communication is modulated on the infrared LED array of the base station with a 60 MHz bandwidth, and the PD array of the terminal can receive the communication signal, thus establishing high-speed optical wireless communication links. In addition, an enhanced light reflection surface is deployed on the terminal, and reflects a part of LED signals. As a result, the PD array on the base station side obtains the orientation and distance of the terminal, accurately measuring the 3D position of the terminal.

At the center of the base station's coverage area, the communication rate is higher than 60 Mbit/s. In the edge area, the communication rate is higher than 10 Mbit/s. The ISAC-OW prototype system aims to accurately measure the health status of users and the location of robots through wireless optical links (visible light and infrared spectrum) in intelligent healthcare, smart home, and intelligent cockpit scenarios to remotely control mobile robots to pick up and transport objects. At the same time, the optical link in the prototype carries the wireless high-speed transmission

Table 1 Performance of the ISAC-OW prototype

Functionality	Distance	Performance
Optical wireless	2.5 m	Communication speed: 20 Mbit/s
3-D positioning	2.5 m	Average error: 3.2 cm
Heart rate monitoring (duration: 30 seconds)	1.5 m	Average error: 1.6 beats/min

of real-time HD video between the mobile robot and the controller, integrating sensing and communication. In the future, our ISAC-OW prototype will evolve to support more functions in intelligent factories, intelligent healthcare, smart homes, and intelligent cockpit use cases, such as material detection, liveness detection, as well as motion and gesture recognition.

4.2 THz-based Terminal-Side Smart Home and Healthcare ISAC

A comprehensive ultra-broadband THz-ISAC prototype can provide high-performance sensing resolution (Figure 8). Our architecture supports unified hardware and waveform, so communication and imaging simultaneously use the same hardware and OFDM waveform, causing zero overhead to communication.

Figure 9 shows the schematic diagram of the THz RF front end, which is developed on a super-heterodyne transceiver architecture. The Tx uses a sub- harmonic mixer to carry the 2–15.6 GHz baseband signal from the output of the digital-to-analogue converter board to a frequency range of 211–224.6 GHz. A filter is used to remove the image signal, which is then amplified by a drive amplifier (DA) and a power amplifier (PA). The last stage of the transmitter is a high gain lens antenna. At the Rx side, the received signal is amplified by a low-noise amplifier (LNA), and then sent to a mixer to down-convert the signal to the baseband. Next, the signal is amplified to an appropriate level by an RF broadband amplifier and sent back to the analogue-to-digital converter board. Moreover, the transmitting and receiving mixers are both driven by a 104.5 GHz signal that is generated by a 6× multiplier. The mixer and multiplier are integrated into a module, namely the frequency converter module.

Figure 9 shows the core of the frequency converter module, which is a balanced sub-harmonic mixer chip that is made of a commercial indium phosphide (InP) high electron mobility transition (HEMT). The THz output is transitioned to the WR04 waveguide using a quartz probe. The packaged device has a frequency conversion loss of 13–15 dB at the 220 GHz frequency, with a bandwidth covering the 200–230 GHz frequency range. In the same module, a 6× multiplier chip is also packaged as the local oscillator (LO) multiplication chain, which takes in a 17.5 GHz LO signal and outputs a 104.5 GHz, 7 dBm multiplied signal to drive the mixer. The mixer and multiplier are connected by using the WR08 waveguide, which suppresses the low-frequency parasitic spectrum of the multiplier due to the high-pass characteristics of the waveguide. The size of the frequency converter module is 28 mm × 30 mm × 20 mm, which has also reduced the overall size of the front end. The single-balanced mixer will output two spectra that are symmetrical with respect to the LO, and an image suppression filter is needed to filter out the lower sideband. The adopted filter is a 7th-order Chebyshev waveguide filter, which is fabricated with a high precision CNC machine. Considering that our baseband signal is in the range of 2–15.6 GHz, and the lowest frequency of the useful signal is very close to the LO frequency, a transmission zero is designed outside of the pass band to achieve a steep transition band, and to make the most of the baseband signal bandwidth.
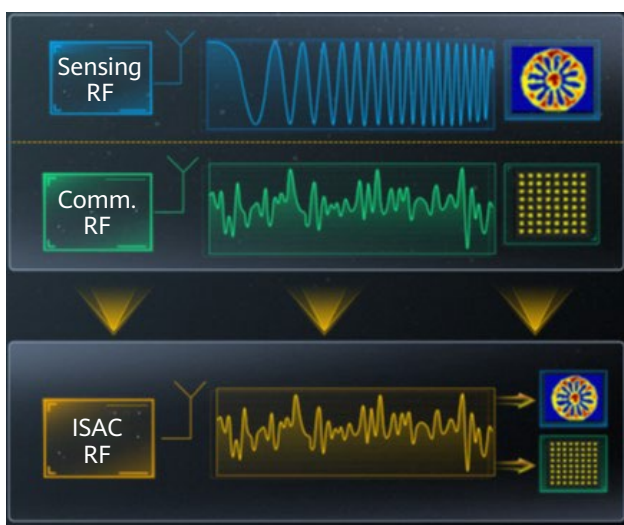

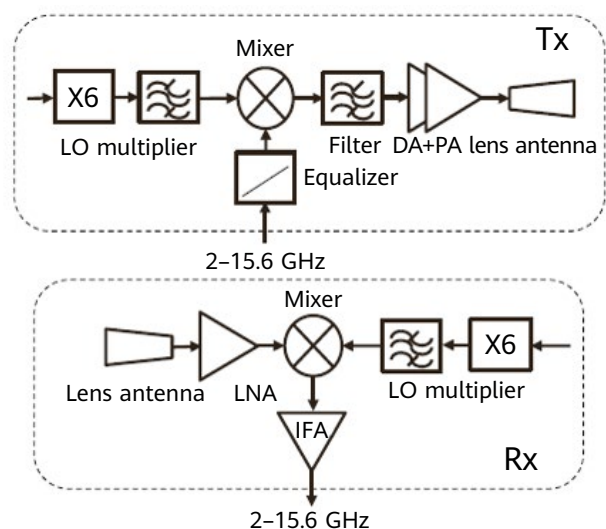**Figure 8** ISAC: Hardware and waveform sharing

Figure 9 RF front end architecture of the THz prototype

As depicted in Figure 10, simultaneous sensing functionality, i.e., the ISAC scenario, is believed to be another key feature of 6G systems, which is verified as well. Note that a 0.6 m distance is selected for a more convenient demonstration, and other distances are also supported.
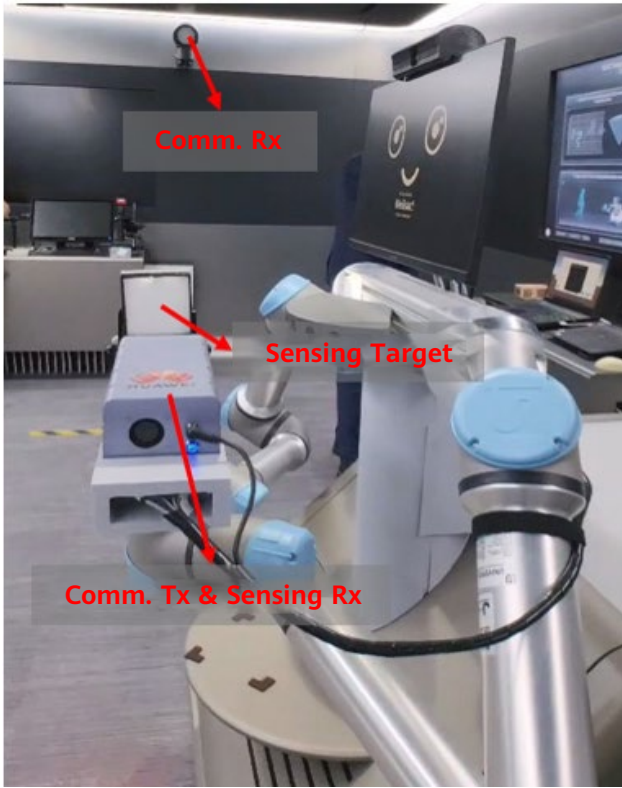

Figure 10 ISAC demonstration system scenarios
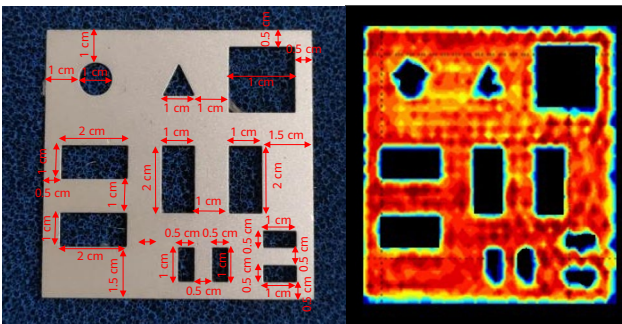

Figure 11 System communication rate

Figure 12 ISAC imaging and real target comparison

Figures 11 and 12 illustrate real-time test results. Compared to the communication system, the ISAC system has used a 2 GHz bandwidth. Although it is different from the communication system, the resolution of the orientation is not related to the bandwidth but the synthetic aperture size. In other words, even though bandwidth is reduced, millimeter-level resolution can still be achieved. At the same time, it is possible to guarantee an effective bandwidth for communication performance, supporting 64-QAM. Without the pilot and other overhead, the spectral efficiency is approximately 5.2 bit/s/Hz, and the throughput can exceed 10 Gbit/s.

4.3 E-Band ISAC RAV Testing for Smart Transportation

4.3.1 E-Band ISAC Architecture and Algorithm

Based on the E-band platform and using OFDM, this paper verifies the feasibility of the root multiple signal classification (MUSIC) algorithm in a millimeter wave frequency band. Figure 13 shows the overall architecture of the ISAC system [13]. The binary user data is divided into parallel streams and mapped onto complex-valued QAM symbols, resulting in the modulation symbol sequence $\mathbf{F}_{Tx}(m, n)$. $\mathbf{F}(m, n)$ is the transmitted data at subcarrier n and symbol m [14].

To verify ISAC's range and velocity measurement capabilities, the root-MUSIC algorithm is used. The power delay spectrum (PDP) of OFDM signals can be expressed as

$$(\mathbf{F})_{k,l} = \sum_{h=0}^{H-1} b_h \cdot e^{j2\pi T_0 f_{D,h} l} \cdot e^{-j2\pi \tau_h \Delta f k} \cdot e^{j\varphi_h} + (\mathbf{Z})_{k,l} \quad (1)$$

where b_h is the attenuation, T_0 is the symbol duration, $f_{D,h}$ is the Doppler shift of the OFDM signal, τ_h is the signal delay caused by the time it takes for it to travel to the object and back, and $(\mathbf{Z})_{k,l}$ is the additive white Gaussian noise (AWGN) with zero mean and variance σ_n^2.
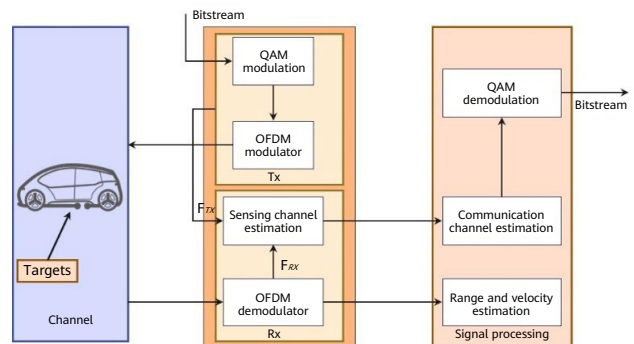

Figure 13 E-band ISAC architecture

We use the constant false alarm rate (CFAR) to detect signals and set an appropriate threshold. If the threshold is greater than the average energy, then the target exists. It is worth mentioning that there is only one target within a certain threshold.

Then, the estimation of range d and velocity v require two autocorrelation matrices R_d and R_v, which can be expressed as follows:

$$\boldsymbol{R}_d = \frac{1}{M}(\mathbf{F})_{k,l}(\mathbf{F})_{k,l}{}^H = \boldsymbol{U}_d \boldsymbol{\Lambda}_d \boldsymbol{U}_d{}^*, \quad (2)$$

$$\boldsymbol{R}_v = \frac{1}{N}(\mathbf{F})_{k,l}{}^H(\mathbf{F})_{k,l} = \boldsymbol{U}_v \boldsymbol{\Lambda}_v \boldsymbol{U}_v{}^* \quad (3)$$

Next, we construct a signal subspace and a noise subspace for range d and velocity v according to a feature root corresponding to the eigenvalues. If the eigenvalues are sorted by size ($\lambda_{i+1} \geq \lambda_i$, $i = 0, ..., N-1$), the first H columns of $\boldsymbol{U}_d$ and $\boldsymbol{U}_v$ (i.e. the eigenvectors corresponding to the largest eigenvalues) span the signal subspace of $\boldsymbol{X}$, and the remaining $N-H$ eigenvectors span the noise subspace $\boldsymbol{U}_d$ and $\boldsymbol{U}_v$. Due to the nature of the noise space, any sinusoid vector $s(e^{j\Omega_d})$ can be expressed as:

$$\boldsymbol{s}\left(e^{j\Omega_{d_k}}\right) = \left[\exp\left(-j2\pi\Omega_{d_k}\frac{1}{N}\right), ..., \exp\left(-j2\pi\Omega_{d_k}\frac{N}{N}\right)\right]^T \quad (4)$$

where $s(e^{j\Omega_d})$ can be rewritten as $\boldsymbol{p}(z)$ and shown as follows:

$$\boldsymbol{p}(z) = [1, z, ..., z^{N-1}]^T \quad (5)$$

To solve the range of the target, we can solve the root of $\boldsymbol{p}^H(z)\boldsymbol{V}_d\boldsymbol{V}_d^H\boldsymbol{p}(z) = 0$. Since only the roots on the unit circle are interested, on the unit circle, $\boldsymbol{p}^H(z) = \boldsymbol{p}^T(z^{-1})$, the polynomial that finds the root-MUSIC algorithm can be written as:

$$P(z) = z^{N-1}\boldsymbol{p}^T(z^{-1})\boldsymbol{V}_d\boldsymbol{V}_d^H\boldsymbol{p}(z) = 0 \quad (6)$$

Since the root of the polynomial is mirror-symmetric about the unit circle, we only need to pay attention to the root within the unit circle, which can be expressed as:

$$z_k = \exp(-j2\pi\Omega_{d_k}\frac{1}{N}) \quad (7)$$

Then, the range of the hth target can be expressed as:

$$\Omega_{d_k} = \arg(z_k)\frac{N}{2\pi} \quad (8)$$

Similarly, the velocity of targets can also be estimated by root-MUSIC.

4.3.2 Parameter Settings and Scenarios

To verify the application of mmWave in smart transportation scenarios, the E-band frequency band is used to test the RAV performance in short-range small-block targets and long-distance vehicle targets. The OFDM signal is implemented using 64-QAM modulation (Table 2).

Figure 14 shows the indoor scenario and Figure 15 shows the outdoor scenario. The test verifies the RAV and communication capabilities of the E-band equipment for ISAC. In the actual

Table 2 Parameter settings

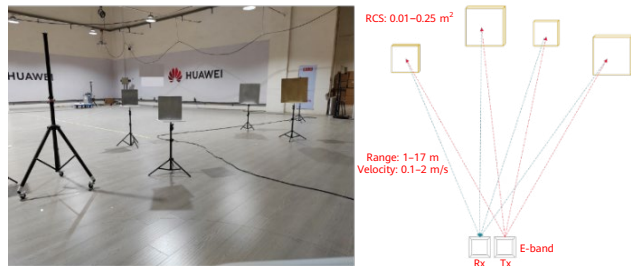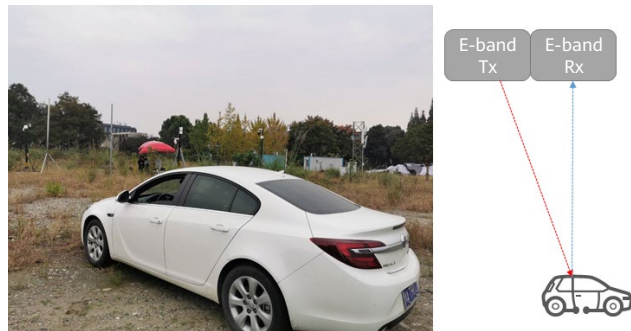| Symbol	Parameter	Value
f_c	Carrier frequency	72 GHz
N	Number of subcarriers	3312
M	Number of symbols	28
BW	Bandwidth	2 GHz
Δf	Subcarrier spacing	600 kHz
T_{Total}	Transmitted per symbol duration	1.78 μs
T_{CP}	Cyclic prefix duration	117.2 ns
Δr	Range resolution	7.5 cm
Δv	Velocity resolution	41.7 m/s

Figure 14 E-band ISAC in indoor scenarios

Figure 15 E-band ISAC in outdoor scenarios

test environment, the E-band uses separate Tx and Rx modes. The OFDM signal is transmitted by the Tx module. After the target is reflected, the Rx module receives the OFDM signal. Then, the range and velocity of the target are estimated by radar signal processing. In the actual test, the communication part of OFDM is in progress all the time.

4.3.3 Test Results and Analysis

Figure 16 shows the OFDM-based sensing test results in this round; range and velocity indicate an indoor scenario while range and angle indicate an outdoor scenario. The results show that the range accuracy is 0.0177 m and the velocity accuracy is 0.047 m/s. The range performance is due to the super-resolution capability of the root-MUSIC algorithm. The algorithm of velocity depends on the ability of the range to solve the derivative of time. The E-band beam width is 12°, and the target angle estimation error accuracy is about 3°. The single-user throughput can reach 9.31 Gbit/s.
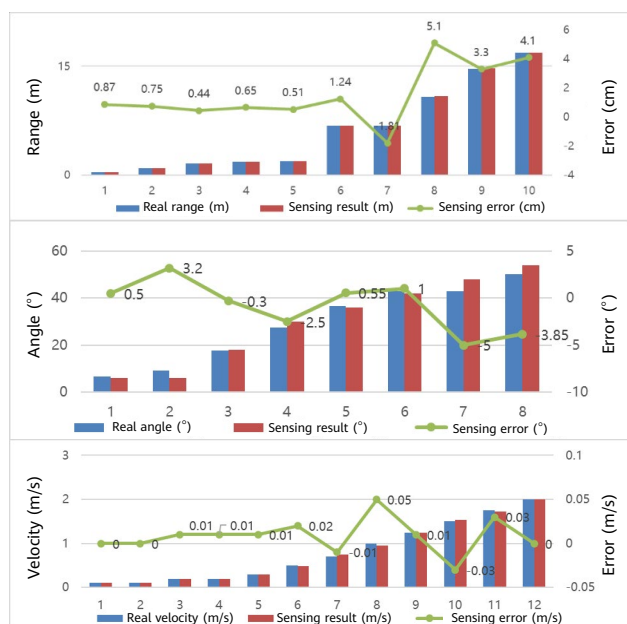
Figure 16 E-band ISAC sensing results

5 Conclusion and Outlook

In this paper, we presented ISAC technologies as a novel paradigm for 6G vertical services and healthcare applications. We covered smart factories, transportation, homes, and health monitoring and treatment applications, which cover a wide variety of major demands. We predict that ISAC will boost applications in various industries in multiple ways. To commercialize 6G connected intelligence, we emphasized certain technical challenges that require careful design considerations from both theoretical and practical perspectives. We then described the corresponding solutions through comprehensive prototypes, where trial results illustrate the feasibility of a number of services. These include ultra-low error rate 3D positioning for smart factories, heart rate monitoring in healthcare applications, ultra-high speed wireless transmission used to carry multiple types of data, and finally, accurate range, angle and velocity sensing and ultra-high resolution imaging for the transportation industry and many other applications. While we took the first step toward ISAC-assisted vertical industrial applications, there are key challenges remaining to be overcome, providing several research directions:

- ISAC system performance with the RF impairment constraint
- Cost and performance tradeoffs for communication and sensing
- Interference hindering communication and sensing integration performance
- ISAC channel modeling and evaluation method

In addition to the above concerns, ISAC also faces other challenges in system protocol and process design, hardware collaborative design, data fusion, privacy and security, power consumption, size, and cost. We believe that as technology evolves, these problems can be solved.

References

[1] Wen Tong and Peiying Zhu, "6G: The next horizon – From connected people and things to connected intelligence," England: Cambridge University Press, 2021.

[2] S. Kumar, A. Majumder, S. Dutta, R. Raja, S. Jotawar, A. Kumar, M. Soni, V. Raju, O. Kundu, E. H. L. Behera *et al.*, "Design and development of an automated robotic pick & stow system for an e-commerce warehouse," arXiv preprint arXiv:1703.02340, 2017.

[3] SAE, "Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles," 30. April 2021.

[4] H. Zhan, Q. Li, K. Zhao, L. Zhang, Z. Zhang, C. Zhang, and L. Xiao, "Evaluating PM2. 5 at a construction site using terahertz radiation," IEEE Transactions on Terahertz Science and Technology, vol. 5, no. 6, pp. 1028–1034, 2015.

[5] D. K. Pin Tan *et al.*, "Integrated sensing and communication in 6G: Motivations, use cases, requirements, challenges and future directions," 2021 1st IEEE International Online Symposium on Joint Communications & Sensing (JC&S), pp. 1–6, 2021.

[6] M. Braun, C. Strum, and F. K. Jondral, "Maximum likelihood speed and distance estimation for OFDM radar," Proc. 2010 IEEE Radar Conf., 2010.

[7] B. Paul, A. R. Chiriyath, and D. W. Bliss, "Survey of RF communications and sensing convergence research," IEEE Access, pp. 252–270, 2016.

[8] K. Wu and L. Han, "Joint wireless communication and radar sensing systems - state of the art and future prospect," IET Microwaves Antennas & Propagation, vol. 7, no. 11, pp. 876–885, 2013.

[9] K. Siddiq, R. J. Watson, S. R. Pennock, P. Avery, R. Poulton, and B. Dakin-Norris, "Phase noise analysis in FMCW radar systems," 2015 European Radar Conference (EuRAD), pp. 501–504, 2015.

[10] Wei, Zixian, Zhaoming Wang, Jianan Zhang, Qian Li, Junping Zhang, and H. Y. Fu, "Evolution of optical wireless communication for B5G/6G," Progress in Quantum Electronics 83 (2022): 100398.

[11] Zhang, Jianan, and Timothy J. Kane, "Acquisition, tracking, and pointing for reconfigurable free space optical communication systems in RF challenged environments," in Free-Space Laser Communications XXXII, vol. 11272, pp. 359–368. SPIE, 2020.

[12] Wang, Wenjin, Albertus C. Den Brinker, Sander Stuijk, and Gerard De Haan, "Algorithmic principles of remote PPG," IEEE Transactions on Biomedical Engineering 64, no. 7 (2016): 1479–1491.

[13] T. Wan, Q. Liu, X. Du, and G. Wang, "Performance analysis of joint range and velocity estimator for E-Band ISAC," 2023 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–5, 2023.

[14] J. B. Sanson, P. M. Tome, D. Castanheira, A. Gameiro, and P. P. Monteiro, "High-resolution delay-Doppler estimation using received communication signals for OFDM radar-communication system," IEEE Trans. Veh. Technol., vol. 69, no. 11, pp. 13112–13123, 2020.
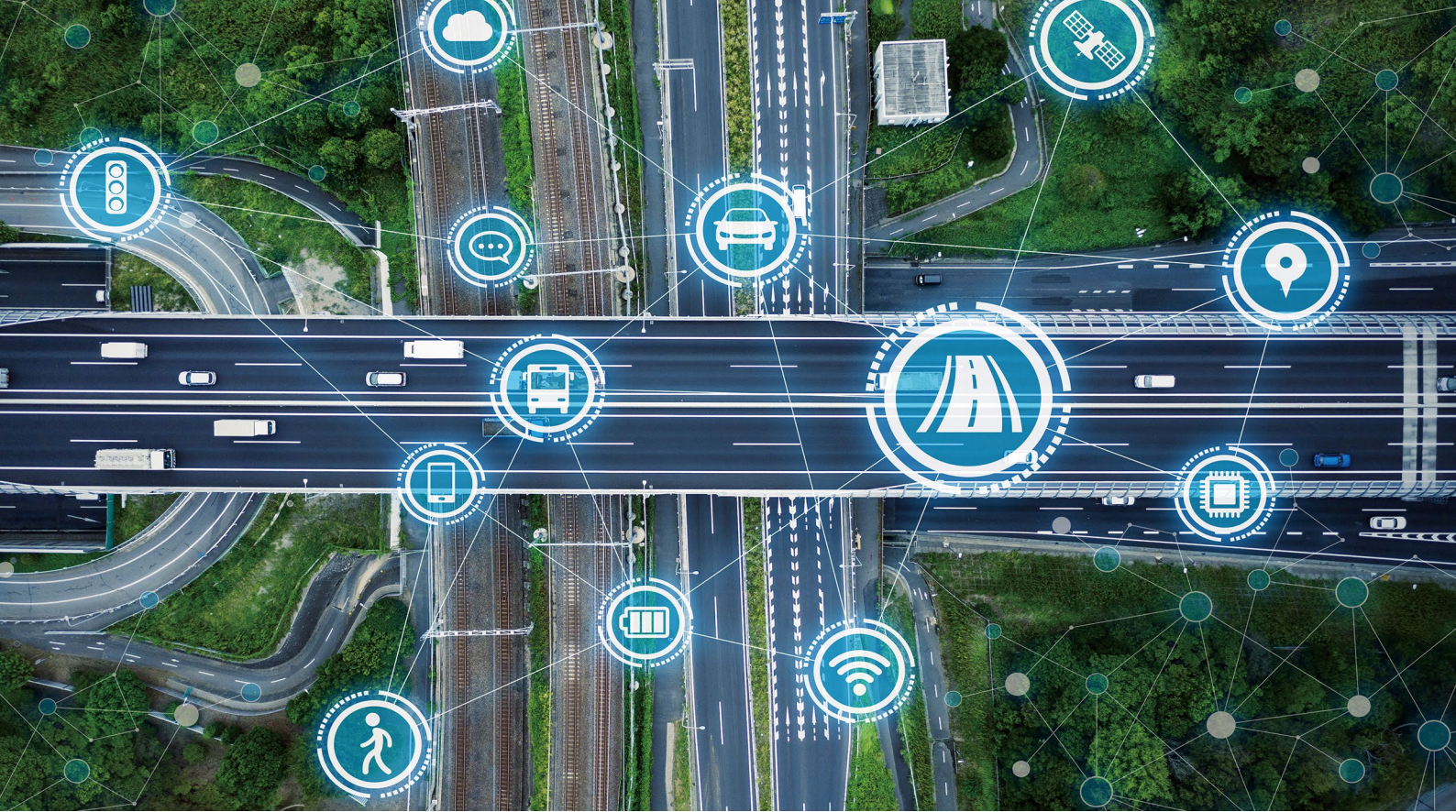
Key Technologies and Applications of Integrated Sensing and Communication in the Internet of Vehicles

Yingpei Lin [1], Lei Lu [2], Chan Zhou [3], Hongjia Su [1], Richard StirlingGalacher [3], Shibin Ge [3], Hong Qi [1], Qi Wang [3]

[1] Research Dept, WN

[2] Munich Research Center

[3] Europe Standardization & Industry Development Dept

Abstract

In this paper, we explore the role of integrated sensing and communication (ISAC) as an emerging fundamental feature in wireless communication systems, specifically targeting Internet of vehicles (IoV) applications for both assisted and autonomous driving. We trace the development of the IoV to underscore the critical importance of environment sensing in assisted driving and autonomous driving, and highlight the advantages offered by ISAC technologies in enhancing environment perception. Furthermore, we delve into real-world use cases that exemplify the technical challenges faced in IoV applications and discuss key ISAC technologies, including waveforms, device-pipe synergy, sidelink enhancement, fusion imaging for extended targets, and micro-Doppler-based detection, as solutions to these challenges.

Keywords

ISAC, IoV, environment sensing, autonomous driving

1 Introduction

Wireless sensing is a natural attribute of radio waves and has now become an emerging fundamental feature of wireless communication systems where radio waves are transmitted and received via wireless sensing to perceive the physical world [1]. The networked sensing capability will transform wireless communication systems into integrated sensing and communication (ISAC) systems with base stations, devices, and the entire network serving as sensors to provide functions such as detection, positioning, identification, imaging, and mapping, as well as advanced solutions for smart transportation, smart factories, smart healthcare, and environment monitoring.

In recent years, the academic community and industry players have conducted extensive research on the application of ISAC in vertical industries, especially in the Internet of Vehicles (IoV). 3GPP SA1 [2] defines ISAC and describes its application cases, more than 70% of which are related to IoV scenarios.

IoV-oriented ISAC has been widely researched and certain significant conclusions have been drawn. However, with the advancement of automatic vehicle operation and digital human-vehicle interaction, it faces higher challenges. Therefore, it is crucial to develop IoV-oriented ISAC in line with the key trends in IoV evolution. We must explore high-value application scenarios for ISAC and pursue focused research on critical ISAC technologies and applications, guided by these trends and scenarios.

Section 2 summarizes the key trends in the evolution of IoV. Section 3 outlines the high-value application scenarios of ISAC. Section 4 analyzes the application requirements and major challenges of IoV-oriented ISAC. Section 5 presents key ISAC technologies in the IoV field. This paper concludes with a summary and prospects for IoV-oriented ISAC.

2 Evolution of the Internet of Vehicles

The automobile industry is experiencing a significant structural shift toward intelligence and electrification. This transformation is primarily driven by the integration of digitalized vehicle control systems, human-vehicle interaction, and automated vehicle operation. The primary objective of this shift toward automotive intelligence is to enhance passenger comfort and safety. Autonomous driving is a core technology that will enable the realization of automotive intelligence, and it will serve as the catalyst for the technological transformation of the automobile industry.

Currently, the automated driving systems (ADSs) available in the market are only at L2 (partial driving automation) and L3 (conditional driving automation) [3]. This means that vehicles still require human control. An advanced driver-assistance system (ADAS) can provide features like adaptive cruise control (ACC), lane centering, and traffic jam chauffeur during low-speed driving or stop-and-go traffic. However, the driver must stay focused during the drive and take control of the vehicle in case of complex or unexpected issues. A true ADS refers to an L4 (high driving automation) or L5 (full driving automation system), where the driver is completely relieved from driving tasks and can enjoy a comfortable and safe trip as a passenger.

According to estimates, by 2025, around 15% of the world's vehicles will have L3 autonomous driving capabilities, while only about 5% will have L4 or higher capabilities [4]. By 2030, it is expected that the proportion of L4 autonomous vehicles worldwide will increase to 10%. Google's Waymo has announced that their autonomous vehicles can operate at L4 automation, but they are not yet mass-produced. Mercedes-Benz has identified L4 and L5 ADSs as major components of the future urban transport solution in its 2030 vision [5]. Volkswagen plans to launch commercial L4 autonomous driving vehicles in 2026 [6], while all mainstream Hyundai automobiles will be equipped with the AutoMode autonomous driving technology to achieve L4 automation by 2026 [7]. Baidu's Apollo RT6 vehicles have been put into operation on its robotaxi platform Apollo Go since 2023, and L4 fully autonomous driving pilots have been established in Beijing, Wuhan, Chongqing, and Shenzhen. The model is expected to be mass-produced between 2026 and 2028 [8].
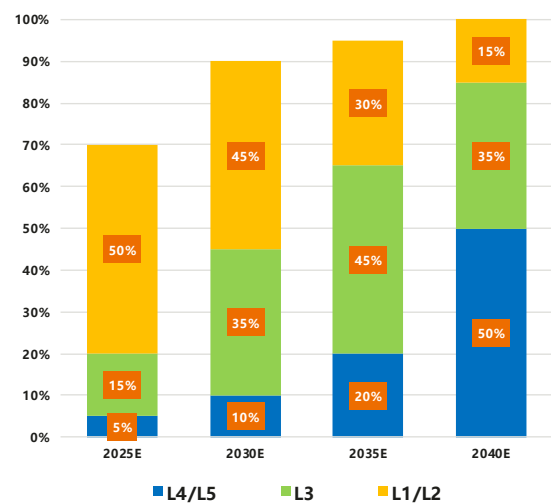

Figure 1 Penetration rate of global autonomous vehicles by level (2025–2040)

The biggest challenges facing the mass production and widespread adoption of L4 or L5 autonomous vehicles is the rigorous reliability and safety standards required by the public for autonomous driving. While the current L4 or L5 automation technology has achieved a safe driving range of one million miles [9] and has superior accident response capabilities compared to humans [10], the public remains cautious about autonomous driving and AI. They will not be fully assured until almost zero errors and accidents are found in autonomous driving practices. To accurately judge the driving environment, ADSs rely heavily on their environment sensing capabilities within the operational design domain (ODD). The ODD defines specific operating conditions for ADSs, such as environmental, geographical, and time of day constraints, or traffic and roadway characteristics. Serious autonomous vehicle accidents have occurred due to misjudgments by ADSs that were unable to sense their environment because it was beyond the coverage of the ODD [11]. To ensure the safety of autonomous driving in various environments, the coverage of the ODD should tend toward infinity, which poses a significant challenge to vehicle-based sensing solutions.

Vehicle sensing faces major challenges that can impact its detection capability and accuracy.

- The detection capability or accuracy toward static environments can be influenced by changes in traffic rules (such as traffic lights), and lane changes due to road repair or traffic control.
- The sensing capability may be limited in long-tail scenarios. For example, non-line-of-sight (NLOS) sensing is limited due to the position shading of in-vehicle sensing components.

These challenges are difficult to solve with technical solutions due to limited hardware capabilities, high costs, and other factors. As a result, achieving near-infinite ODD requirements for high-level autonomous driving solely with vehicle sensing is challenging.

In actual cases, even the driver-assistance systems below L4 highly depend on the environment sensing capability. In harsh conditions such as severe weather and blind spots, if the driver-assistance systems do not receive effective and accurate environment sensing information, they may fail to provide timely warnings and solutions to potential dangers, resulting in lower safety levels or even system shutdown (see Table 1).

Table 1 Sensing and decision-making capabilities of ADSs by level

	L2	L3	L4/L5
Policy	Balancing the detection accuracy and confidence rate	High detection accuracy and range	High detection accuracy and range
Safety guarantee	Driver (at any time)	Vehicle + driver	Vehicle (automatic stops in safe areas)
ODD limitation	Frequent driver interventions or high safety risks (delayed takeovers)	Frequent driver interventions, leading to poor autonomous driving experience	Frequent speed reductions or stops, leading to poor autonomous driving experience

3 Value of Integrated Sensing and Communication

In Section 3, we will examine the necessity and benefits of combining roadside sensing with in-vehicle sensing (Section 3.1). Additionally, we will explore the advantages of the ISAC solution over other roadside sensing solutions (Section 3.2).

3.1 Necessity of Roadside Sensing

Currently, vehicles rely on a variety of in-vehicle sensors to provide the environment sensing capability. These sensors include millimeter-wave wireless rangefinders, laser rangefinders, and cameras. They are strategically placed throughout the vehicle to detect objects, measure ranges, and determine relative speeds. Each sensor type serves a specific purpose and supports different driver-assistance functions. A long-range wireless rangefinder positioned at the front of the vehicle can detect the range and speed of the vehicle ahead within a range of approximately 250 meters, enabling adaptive cruise control. A short-range front wireless rangefinder with a wider field of view is responsible for supporting emergency braking. It is important to note that each sensor type has its strengths and weaknesses. Millimeter-wave wireless rangefinders, for example, can operate effectively in extreme weather conditions, whereas laser rangefinders and cameras may struggle. Cameras, however, excel at image-based detection. Since no single sensor can fulfill all requirements in every scenario, integrating information from multiple sensors simultaneously enhances the overall environment sensing capabilities [12].
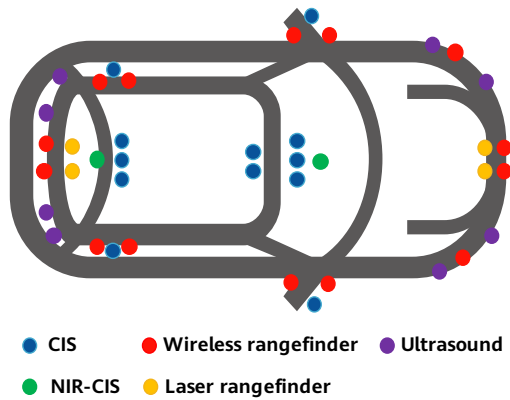
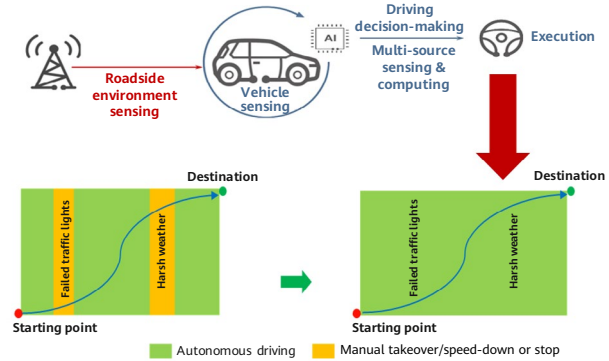**Figure 2** Sensors on L4/L5 vehicles

Figure 3 Reliable and comprehensive environment information provided by the V2X network

With the application of vehicle-to-everything (V2X) systems, the wireless communication capability of the ADAS has been significantly enhanced. Integrating the communication function into vehicles can be viewed as an enhancement of existing sensor systems. Through the exchange of information, vehicles can obtain environmental data that cannot be detected by sensors on a single vehicle. This enables safe and efficient driving through inter-vehicle cooperation. This cooperation enhances the long-range sensing capability of autonomous vehicles. It facilitates the integration of inter-vehicle information, allowing vehicles to detect and identify objects in higher resolution without any blind spots in all scenarios. Furthermore, with the assistance of communication functions, information sensed by roadside devices can also be downloaded to vehicles. Key advantages of roadside sensing devices include:

- High sensing positions for overhead perspectives
  - Roadside sensing: Expand the vertical sensing angle to enhance NLOS sensing capabilities.
  - Roadside cameras: Offer shading protection against extreme weather conditions such as rainstorms, with virtually no interference affecting their performance.
- Large size
  - Roadside sensing: Large-sized antennas (with more antenna array elements) offer enhanced angular resolution for azimuth and pitch angles.
  - Roadside cameras: Various large-sized cameras equipped with lenses of different focal lengths (such as wide-angle and telephoto lenses) can capture distortion-free images due to their super-short exposure times.

3.2 Advantages of Integrated Sensing and Communication

Traditional solutions for vehicle sensing and communication involve deploying two separate systems: an in-vehicle sensing system and a V2X communication system. However, these systems are disjointed and result in several challenges. Firstly, they are hungry for power, require a lot of space for sensors, and increase the vehicle's weight. Additionally, a large number of antennas are required, and complex cabling is needed within the vehicle. As the number of sensors and communication devices increases, it becomes increasingly challenging to design the interior of the vehicle.

Furthermore, independent in-vehicle wireless rangefinders and V2X communication devices must achieve frequency spectrum compatibility to avoid interference between wireless rangefinders or between wireless rangefinders and communication devices. Interference between wireless rangefinders is a significant bottleneck, as the fields of view of multiple co-frequency wireless rangefinders from adjacent vehicles may overlap and cause interference to the overall detection performance. In addition, the use of both sensing and communication systems doubles the frequency spectrum demands.

Therefore, deploying an ISAC system for vehicles is becoming an inevitable technical direction. This system is designed to reuse the frequency spectrum and some components while combining signal processing and adopting the protocol stack. This transformation from loose coupling to high integration allows for a more efficient and effective system.

The ISAC system offers several potential benefits due to its ability to simultaneously transmit data and sense the environment. These benefits include reduced hardware size, fewer antennas, lower costs, reduced latency, improved spectrum utilization, and increased energy efficiency. Additionally, the highly integrated design of the system allows for efficient sharing of information and resources between the sensing and communication systems. The multi-user resource management and anti-interference mechanisms of the communication system also provide an advantage, as they effectively solve interference between wireless rangefinders when multiple ISAC systems are in use.

The information obtained from the ISAC system can also be used to enhance the communication performance, such as using location information for wave beam prediction and management, blocking prediction, channel state information (CSI) compression, and adaptive transmission.
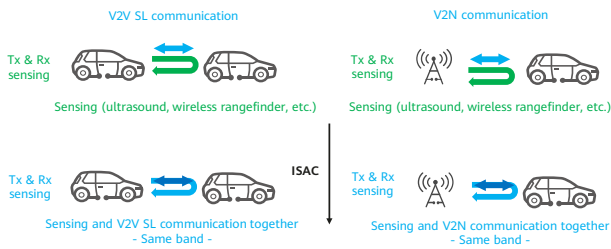
Figure 4 Diagram of ISAC

4 Application Cases of Integrated Sensing and Communication in the Internet of Vehicles

As previously mentioned, network-based sensing can greatly improve the system's environment sensing capabilities. This is particularly useful in situations where sensors' field of view is limited due to their installation positions. By supplementing the sensors with network-based sensing, additional information can be obtained, especially in complex road situations such as at crossroads where blind spots may exist. Network-based sensing can provide clearer information about the road or highway, including the presence of moving objects such as vehicles, animals, and pedestrians, as well as traffic density.

4.1 Scenario and Use Case Analysis

3GPP has defined and described a range of ISAC use cases. Many of these use cases are related to vehicles and adopt the following implementation solutions:

- Integrated communication and sensing functions on base stations and terminal devices
- Exchange of sensing information between networked terminal devices and network entities
- Sensing results of base stations and terminal devices as additional inputs for environment sensing

Table 2 V2X use cases in 3GPP SA1

Use Case	Significance of Network-based Sensing
Pedestrian/Animal intrusion detection on a highway	Improving the sensing coverage along highways on the basis of traditional sensors
Sensing at crossroads	Covering blind spots under complex road conditions
Sensing-assisted automotive maneuvering and navigation	Coordinating network-based sensing information among different user equipment (UE) in a network
Vehicle sensing for ADAS	Complementing existing ADAS sensors with network-based sensing

Table 2 describes some use cases. According to the functional requirements of these use cases, network-based ISAC needs to provide the following mechanisms:

- Select and configure base stations and terminal devices to perform sensing operations.
- Share sensed information between a radio access network (RAN) and a core network (CN).
- Authorize network nodes to perform sensing operations.
- Send sensed information in the network to third-party service providers.
- Charge for network-based sensing services.

The performance requirements for ISAC depend on the service objectives of different use cases. In IoV-related use cases, the target environments and objects mainly exist outdoors. The specific performance requirements for each use case primarily vary in terms of accuracy, range, reliability, delay, and refresh interval of location and speed measurements.

Pedestrian/Animal intrusion detection on a highway: When it comes to highways, base stations need to be able to effectively detect pedestrians and animals from a long range. This is due to the width and length of the roads. Additionally, vehicles on highways tend to move at high speeds, which creates a highly dynamic environment. Therefore, 3GPP has strict requirements for the sensing refresh interval and accuracy [2]:

- Positioning accuracy (horizontal): < 1 meter
- Refreshing interval: < 0.1 seconds

Sensing at crossroads: The sensing performance requirements are influenced by the density of vehicles at crossroads. For instance, during peak hours, in addition to high sensing accuracy and short refresh intervals, the presence of dense pedestrian and vehicle flows necessitates a high sensing resolution. Scenario-based missed detection rate and false alarm rate are also crucial. 3GPP has outlined the following requirements [2]:

- Positioning accuracy (horizontal): < 1 meter

- Refreshing interval: < 0.1 seconds

- Missed detection rate: < 5%

- False alarm rate: < 5%

Sensing-assisted automotive maneuvering and navigation: In this scenario, it is crucial to coordinate and provide information from multiple in-vehicle sensors. This is necessary to ensure accurate and timely vehicle maneuvering and navigation, especially considering the short distances between moving vehicles. The information gathered by these sensors must be highly accurate and updated frequently. Furthermore, since the information is collected from various in-vehicle sensors, it is important to prioritize low-latency information synchronization.

Vehicle sensing for ADAS: The ADAS in vehicles comprises both long-range sensors (250–300 meters) and short-range sensors (30–100 meters). The ISAC technology will supplement or replace the original sensors on the vehicle. To ensure the safety of ADAS, 3GPP has outlined specific requirements for error rate and error margin [2]:

- Positioning accuracy (horizontal): < 40 centimeters for long-range sensors; < 10 centimeters for short-range sensors

- Refreshing interval: < 0.2 seconds for long-range sensors; < 0.05 seconds for short-range sensors

- False alarm rate: < 1%

4.2 Key Requirements and Challenges

Technical challenges must be addressed to implement the preceding use cases. First, the IoV encompasses various target objects, such as vehicles, pedestrians, animals, and temporary barriers like roadblocks for road construction. Therefore, the sensing system needs to not only identify different types of objects but also accurately locate them. Second, not all target objects are capable of wireless communication. Hence, the ISAC system must be able to

simultaneously detect active targets (vehicles equipped with V2X communication capabilities) and passive targets (pedestrians or animals) without communication capabilities. In addition, the mobility of target objects presents additional requirements for ISAC. For instance, in a multi-vehicle coordination use case, continuous sensing of multiple nearby vehicles and objects is necessary, rather than a momentary one-time detection.

In high mobility scenarios, continuous sensing services often require the deployment of multiple sensing signal transmitters and receivers due to the limited sensing range of a single sensor. When these transmitters and receivers are located at different network nodes, such as in a bidirectional sensing scenario, it becomes necessary to coordinate and align their measurement results to minimize or eliminate false negatives.

While the introduction of ISAC is expected to enhance the environment sensing capability in V2X use cases, traditional sensors such as wireless rangefinders will still play a crucial role, at least for a certain period of time. Therefore, wireless rangefinders and cellular systems may co-exist in the same frequency band or adjacent bands (for example, 77 GHz). However, the transmission of data by vehicles may interfere with wireless rangefinder receivers, and similarly, wireless rangefinder signals may interfere with vehicles or base stations. To effectively handle interference in ISAC systems, technical solutions involving frequency spectrum sharing, waveforms, resource allocation, and artificial intelligence technologies are required.

5 Key Technologies of Integrated Sensing and Communication for the Internet of Vehicles

Despite demonstrating significant potential, the business-driven and scenario-based ISAC solution still faces many technical challenges. Section 5 analyzes several key technologies of ISAC for IoV scenarios.

5.1 Basic Introduction to the Integrated Sensing and Communication Solution

ISAC aims to enable sensing in a communication system, allowing for the sharing of network resources, frequency spectrum, and processing capabilities. This can be achieved at three different levels:

- Service level: Sensing and communication are implemented using separate hardware operating on their own frequency bands. The sensing and communication systems interact at the application layer to realize synergistic benefits. For example, a base station may utilize sensors to gather environment information, enabling more effective wave beam prediction and management. Through the coordination of communication units, the interference between wireless rangefinders from different base stations can be minimized.

- Spectrum level: The sensing and communication modules operate within the same system and share frequency spectrum resources through techniques such as time division, frequency division, or space division. Additionally, the two functional modules can share certain hardware resources, reducing the need for additional hardware and implementation overheads.

- System level: A single system supports both sensing and communication functions, sharing air interface technologies and frequency spectrum resources. The system dynamically optimizes resource scheduling and even air interface design based on service requirements and performance indicators, maximizing the efficiency of network resources.

5.2 ISAC Waveforms

Extensive research has been conducted by the industry on ISAC waveforms. This paper provides an overview of several waveforms including orthogonal frequency division multiplexing (OFDM), OFDM-chirp, orthogonal chirp division multiplexing (OCDM), and orthogonal time frequency space (OTFS).

OFDM-chirp has shown excellent sensing performance, but its communication performance in the Rayleigh channel is inferior to that of OFDM. OFDM is a widely used communication multi-carrier modulation technology. Under different parameters, the integrated sidelobe ratio (ISLR) and peak to sidelobe ratio (PSLR) of OFDM are slightly lower than those of linear frequency modulation (LFM). Currently, academic researchers are actively exploring methods to improve the sensing performance of OFDM.

The basic principle of OFDM-chirp waveform is to modulate chirp signals to different subcarriers. Because subcarriers of OFDM signals are orthogonal, signals on different subcarriers are orthogonal. The OFDM-chirp signal model [13] generated in the time domain is represented as

$$s(t) = \sum_{n=0}^{N-1} u(t-T) \exp\left[j2\pi \left(f_n t + k_n t^2 / 2 \right) \right]$$

where t ($0 \le t \le T$) is the time sampling of the signal; $u(t) = 1$, ($0 \le t \le T$) is the rectangular window function; f_n and k_n are the start frequency and slope of the nth subcarrier of signal $s(t)$, respectively. The multi-carrier feature of OFDM-chirp can address the problem of low transmission rate of LFM single-carrier signals. In terms of compatibility with existing systems and implementation complexity, OFDM-chirp shows potential as an ISAC waveform, and performs exceptionally well in sensing compared to other multi-carrier waveforms.

Under the same parameters, the range and speed resolution of OFDM-chirp signals are equivalent to those of OFDM integrated signals. In a Gaussian channel, performance is minimally affected by varying signal-to-noise ratios as depicted in Figure 5 which illustrates the bit error rates under different signal-to-noise ratios. Since OFDM-chirp signals are derived from OFDM signals through LFM spread spectrum, the OFDM-chirp signal model shares many similarities with the OFDM signal model. Analyzing the peak-to-average power ratio (PAPR) curves of OFDM-chirp and OFDM signals reveals that the PAPR performance of OFDM-chirp integrated signals is nearly identical to that of OFDM signals.
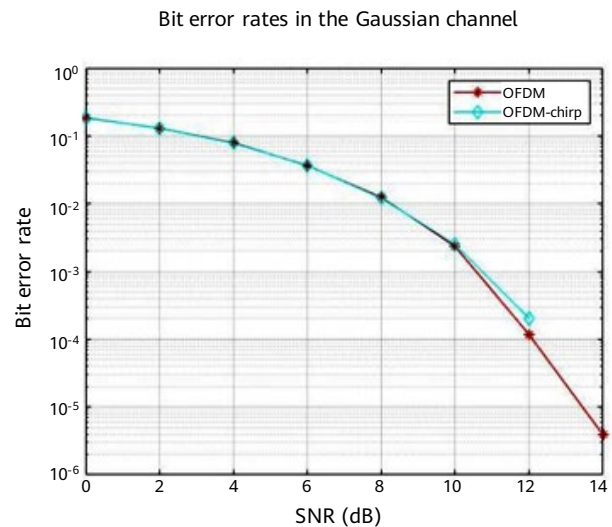

Figure 5 Performance comparison of OFDM and OFDM-chirp waveforms

The basic principle of OCDM involves performing a Fresnel transform within the same bandwidth to form a group of orthogonal linear chirp signals, and modulate the communication information to the amplitude and phase of these chirp signals [14]. This allows for efficient transmission of communication data. The OCDM signal is composed

of multiple chirp signals. Therefore, the OCDM signal is insensitive to the Doppler shift and can effectively resist time-selective fading in case of insufficient guard interval. Compared to an OFDM signal, an OCDM signal requires two additional matrix multiplications at the transmitter, and is generated in a more complex manner. To make OCDM commercially viable use, we need to address technical difficulties, such as building superior optical address codes, producing high-performance optical encoders or decoders, improving system performance, increasing system capacity, and improving system resource utilization.

OTFS modulation can be regarded as a special type of OFDM modulation — adding an encoder before the input end and a decoder after the output end of OFDM modulation [15]. Figure 6 shows the modulation and demodulation processes of OTFS signals.

The effective channel matrix of OTFS modulation is in the delay-Doppler domain. In this case, most elements in the matrix are zero, and locations of non-zero elements are related to the delay and Doppler shift of the path. Therefore, we only need to estimate the delay, Doppler shift, and effective multiplexing gain of each path in the channel. In this way, the channel estimation can be significantly simplified. Because OTFS modulation uses rectangular transmitter and receiver pulses, and the input-output relation is not a two-dimensional convolution operation, the frequency-time domain equalization algorithm cannot be directly used at the receiver end, and other symbol demodulation algorithms are more complex than that of OFDM modulation.
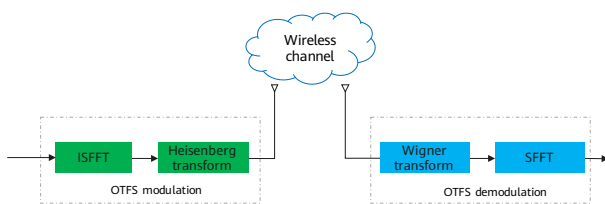
Figure 6 OTFS signal modulation and demodulation processes

vehicle poses challenges to integrating information obtained from distributed in-vehicle and roadside sensors. These sensors are unable to obtain the vehicle's accurate position and speed because it is moving fast. In addition, the physical characteristics of the same object perceived by in-vehicle and roadside sensors may be different. For example, due to different sensing angles, a static object perceived by the roadside sensors is a moving object from the perspective of the in-vehicle sensors.

To solve this challenge, the vehicle can substitute the uncertain location and mobility parameters as unknown variables into the target detection problem, forming a joint detection problem of unknown target and in-vehicle parameters. Then the vehicle optimizes the algorithm, such as finding the optimal solution with the Newton method or simulated annealing optimization algorithm, to improve the accuracy of target detection.

As shown in Figure 7, when different vehicles and roadside units (RSUs) or gNodeBs (gNBs) perform joint sensing, the estimated locations of vehicles are different from the actual locations because the vehicles are moving fast. In addition, for the same object, each sensing source will have errors, which may be magnified when information is integrated. To solve this problem, each vehicle can report the estimation errors to the integration center, and build an optimization problem to improve the joint sensing accuracy [16].
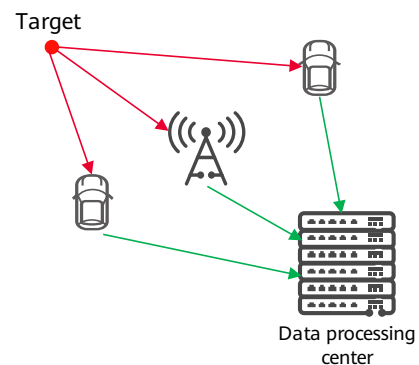

Figure 7 Uncertainty in device-pipe sensing

5.3 Device-Pipe Synergy for Distributed Integrated Sensing and Communication

In IoV scenarios, due to the last 5% long-tail issues in fragmented scenarios and unpredictable human behavior, autonomous vehicles must obtain information from roadside sensors to supplement the information from in-vehicle sensors to make judgments when driving. However, the mobility of a

5.4 Sidelink-Enhanced Distributed Integrated Sensing and Communication

In IoV scenarios, considering that terminal devices may be located in areas without cellular networks, each device should be capable of communicating through PC5 interfaces (sidelink) in addition to basic uplink and downlink communication capability through Uu interfaces. Multiple

devices should communicate through PC5 interfaces to form a distributed network, improving communication and sensing capabilities of vehicles.

However, the design of the existing distributed sidelink-based air interface technology pursues only an optimal communication capability, regardless of the sensing capability. Therefore, the technical bottleneck is how to enable a high-performance sidelink-based sensing capability on a PC5 interface without affecting the communication performance.

One solution is to enhance the sidelink frame structure. In distributed scenarios, the relative positions of devices are not fixed. For reliable information transmission, the automatic gain control (AGC) symbol is introduced into an existing sidelink frame structure, as shown in Figure 8. Before actual data transmission, AGC needs to be performed. Generally, information of the first AGC symbol is copied from that of the second symbol, which means that the AGC symbol does not transmit valid information. Therefore, the AGC symbol also provides space for building sidelink-based sensing capability.
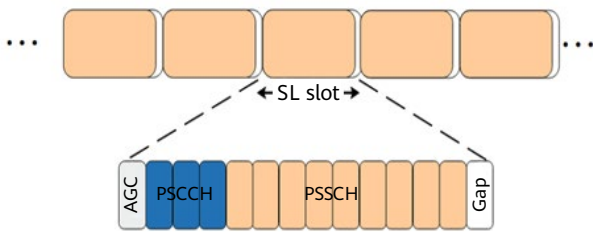
Figure 8 New radio (NR) sidelink frame structure design

A direct method is to fill the AGC symbol with a sensing sequence. However, the AGC symbol may occupy just a small frequency-domain bandwidth, but the signal bandwidth is strongly coupled with the range resolution of sensing. To achieve higher sensing accuracy, we can build a wideband AGC sequence — the AGC symbol occupies as wide a bandwidth as possible in the frequency domain. As shown in Figure 9, an AGC symbol can be divided into two parts in the time domain: The former part sends the wideband sensing sequence; while the latter part receives the sensed echo signal. However, when the AGC symbol occupies a relatively large bandwidth in the frequency domain, as multiple devices all send sensing sequences in the bandwidth, the sequences are superimposed, causing mutual interference. To solve this problem, the sensing sequences sent by different devices are constructed as

orthogonal wideband sequences. For example, the ID of a device is used to scramble a sequence. When the device receives echo signals from the all sequences, the interference signals of other devices are filtered out, and the device receives only its own sensed echo signal. In this way, the sensing capability is enabled without affecting communication performance.
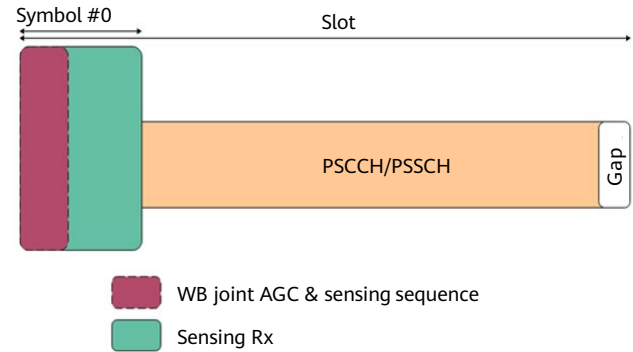

Figure 9 Design of a wideband AGC sensing signal

5.5 Fusion Imaging for Extended Target Detection

ISAC aims to detect, locate, and classify key targets in various IoV applications. These targets include vehicles, vulnerable road users (VRUs) (pedestrians, two-wheeled motorcycle riders, and other non-motor vehicle drivers), and other static or moving obstacles. A traditional 5G NR positioning solution only locates active targets that can send or receive 5G NR signals. In contrast, the next-generation ISAC technology can detect, locate, and identify passive targets by parsing radio signals reflected by the transmission environment.

However, in IoV applications, many targets are bulky vehicles with external surfaces that typically extend to several meters, covering the front, rear, and sides of the vehicles. These sensing targets will generate multiple reflections with a range of several meters and are thus referred to as extended targets. Therefore, by observing multiple reflections and scattered signals generated by an extended target, the position, state, shape, and outline of the target can be estimated and inferred to image the target. When the distributed base stations observe a target vehicle from different angles, they form estimates for different positions on the vehicle. These estimates correspond to a particular side (or portion) of the target vehicle as a set of scattered

points associated with a particular base station. The scattered points seen by each base station are different and may be several meters apart. Therefore, the following steps are required for data association, clustering, and location:

(1) Different base stations estimate the locations of different scattered points. These estimated locations need to be associated with the same target (in this case, a vehicle) to form a set of scattered points corresponding to the target. This process is referred to as clustering.

(2) After obtaining the set of scattered points of the extended target (usually an object outline is formed), the location of the target needs to be determined. Generally, the positioning reference center may be a geometric center of the scattered point set. Alternatively, other reference points may be used when the observed set of scattered points corresponding to the target surface is incomplete, resulting in an inaccurate geometric center.
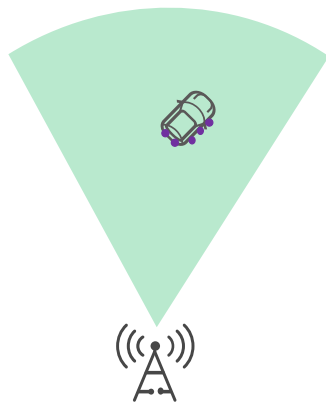
Figure 10 Extended target detection

We can perform steps (1) and (2) by applying different extended target processing algorithms. These algorithms include associating and clustering targets with the movements of the set of scattered points [17, 18], and combining post-processing with target classification [19].

Note that the preceding problem of extended targets does not exist in active target positioning because the transmit and receive antenna ports of active targets are positioning reference points fixed on the targets. Furthermore, conventional sensing systems for V2X collision detection or adaptive cruise control do not have the preceding problem, because they only need to detect the most protruding portion of a target or vehicle in a single angle of view to avoid collision with the target or specific part of the vehicle.

5.6 Micro-Doppler-based Target Detection and Recognition

As described in Section 5.5, target recognition and classification are critical for IoV applications, including various security applications for coordinated manipulation and protection of VRUs. Target recognition and classification can be implemented in a variety of ways. One direct approach is to use scattered points obtained from sensing and observation to determine the shape or outline of a target (such as the vehicle shown in Section 4.1), i.e., imaging. However, to obtain such an image, the system needs to observe the target from multiple angles (especially high angles) with high resolutions to obtain multiple sets of scattered points corresponding to each surface of the target.

Another approach that does not require a complete image of the target involves using Doppler parameters based on the sensed signals, especially those based on the micro-motions of the internal components of the target, such as the swinging legs when a person walks, spinning wheels of a moving car or bicycle, or blades turning at high speeds on a drone. Because the moving speed of an internal component is different from that of the target as a whole, the Doppler component is called micro-Doppler. Figure 11 shows micro-Doppler signatures of a flying drone on the Doppler range [20].
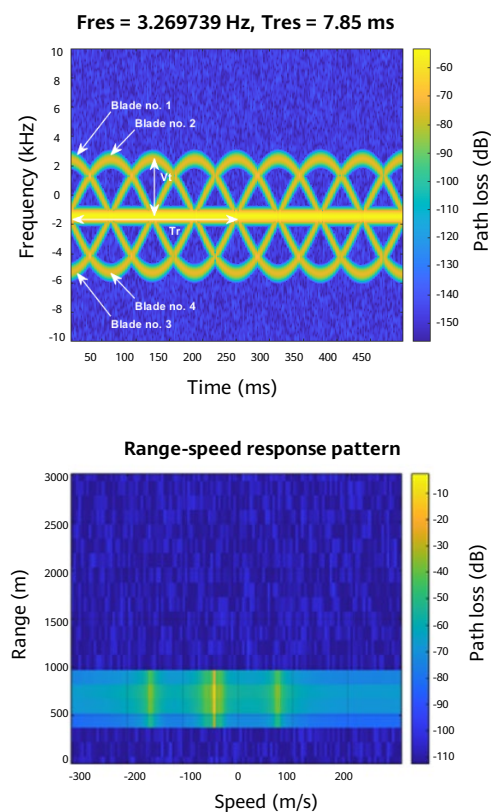

Figure 11 Micro-Doppler example of a target

Different targets such as pedestrians, cars, or bicycles, have different types of internal motions, resulting in target-specific Doppler characteristics. The characteristics are used for target recognition and classification, eliminating the need for detailed high-resolution target imaging from multiple angles, but achieving higher target accuracy (> 90%). [21] demonstrates more examples. Micro-Doppler-based target recognition and classification can be used together with low-resolution imaging to improve the reliability of target classification.

6 Conclusion and Prospect

We can envision a future in which ISAC will bring functional and performance improvement to the IoV. To make it a reality, it is necessary to solve key technical challenges in multiple fields from the air interfaces to the upper-layer applications, and finally meet the expected sensing and communication requirements. The overall design of ISAC also needs to consider complexity of the wireless networks and IoV, such as the mobility of in-vehicle devices, available frequency bands (licensed and unlicensed frequency bands), base station deployment, and collaboration of PC5 and Uu interfaces.

In addition to technical challenges, compliance issue is also a key success factor for ISAC. ISAC systems extend the use of licensed frequency bands which were originally allocated for sensing or intelligent transportation system (ITS) communication and for mobile/V2N communication networks. Implementing simultaneous communication and sensing functions on these bands may encounter compliance issues. When communication or sensing requires a new frequency spectrum, the in-vehicle ISAC devices will require upgrades, which will impact their lifespan. In addition, because wireless rangefinder sensing is a safety-critical function, ISAC solutions must meet the automotive safety integrity level (ASIL) requirements.

ISAC systems must also support networks and devices from different vendors to ensure seamless integration in IoV sensing services, especially anti-interference services. Standardization of functions and interfaces of different ISAC devices is critical to their wide deployment in the future. Overall, solving these technical and compliance challenges will be essential to realizing the full potential of ISAC in the IoV.

References

[1] W. Tong and P. Zhu, 6G: *The Next Horizon: From Connected People and Things to Connected Intelligence*, Cambridge University Press, April 30, 2021.

[2] 3GPP TR 22.837, "Study on Integrated Sensing and Communication," June 2023.

[3] SAE, "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles," April 30, 2021.

[4] Guosheng Securities, "2022 zhineng che liuda kandian" [Six focuses on intelligent vehicles in 2022], November 21, 2021.

[5] Mercedes-Benz, "An interview about progress in development work," https://group.mercedes-benz.com/innovation/case/autonomous/interview-hafner-2.html

[6] Volkswagen, "Volkswagen launches its first autonomous driving test program in the United States," July 6, 2023.

[7] HYUNDAI, "Kia presents 2030 roadmap to become global sustainable mobility leader," March 3, 2022.

[8] Apollo, "Shoupi huozhun! Luobo kuaipao zai Shenzhen kaiqi quanwuren shangyehua yunying" [License approved! Apollo Go starts all-unmanned commercial operations in Shenzhen], June 17, 2023.

[9] Waymo, "Safety Performance of the Waymo Rider-Only Automated Driving System at One Million Miles."

[10] Waymo, "Collision Avoidance Effectiveness of an Automated Driving System Using a Human Driver Behavior Reference Model in Reconstructed Fatal Collisions."

[11] NHTSA, "Incident Reporting for Automated Driving Systems (ADS) and Level 2 Advanced Driver Assistance Systems (ADAS)."

[12] GSA, "The Challenges to Achieve Level 4/Level 5 Autonomous Driving," https://www.gsaglobal.org/forums/the-challenges-to-achieve-level-4-level-5-autonomous-driving/

[13] D. Dash, K. D. Sa, and V. Jayaraman, "Time Frequency Analysis of OFDM-LFM Waveforms for Multistatic Airborne Radar," *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp. 865–870, April 2018, doi: 10.1109/ICICCT.2018.8473233.

[14] L. G. de Oliveira, M. B. Alabd, B. Nuss, *et al.*, "An OCDM radar-communication system," *14th European Conference on Antennas and Propagation (EuCAP)*. IEEE, pp. 1–5, 2020.

[15] L. Gaudio, M. Kobayashi, G. Caire, and G. Colavolpe, "On the effectiveness of OTFS for joint radar parameter estimation and communication," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5951–5965, 2020.

[16] K. Gu, Y. Wang, and Y. Shen, "Cooperative detection by multi-agent networks in the presence of position uncertainty," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5411–5426, 2020.

[17] E. Favarelli *et al.*, "Map Fusion and Heterogenous Object Tracking in Joint Sensing and Communications Networks," *20th European radar conference (EuRad 2023)*, September 20–22, 2023.

[18] H. Kaulbersch *et al.*, "EM-based Extended Target Tracking with Automotive Radar using Learned Spatial Distribution Models," *2019 22nd International Conference on Information Fusion (FUSION)*, pp. 1–8, 2019.

[19] W. Cao, J. Lan, and X. R. Li, "Extended Object Tracking and Classification Using Radar and ESM Sensor Data," *IEEE Signal Processing Letters*, vol. 25, no. 1, pp. 90–94, January 2018.

[20] Mathworks, "Introduction to Micro-Doppler Effects," https://de.mathworks.com/help/radar/ug/introduction-to-micro-doppler-effects.html

[21] J. Pegoraro, J. O. Lacruz, F. Meneghello, E. Bashirov, M. Rossi, and J. Widmer, "RAPID: Retrofitting IEEE 802.11ay Access Points for Indoor Human Detection and Sensing," *IEEE Transactions on Mobile Computing*, doi: 10.1109/TMC.2023.3291882.

Multi-Target Detection and Location in an Integrated Sensing and Communication System

Qi Yang [1], Zhiqiang Han [2], Chuangxin Jiang [3], Junpeng Lou, Juan Liu

State Key Laboratory of Mobile Network and Mobile Multimedia Technology

ZTE Corporation

[1] yang.qi7@zte.com.cn

[2] han.zhiqiang1@zte.com.cn

[3] jiang.chuangxin1@zte.com.cn

Abstract

As a key technology of 6G, integrated sensing and communication (ISAC) coordinates communication and sensing functions through joint signal design. An important application of ISAC is multi-target detection and location, which is a fundamental prerequisite for many sensing tasks. However, because ISAC systems have complex channels, multi-target detection and location is often intractable. In this paper, by setting two thresholds, we propose a heuristic multi-target detection and location algorithm for ISAC applications. This algorithm successfully detects and estimates the location information of multiple targets in a dense clutter environment. Simulation results for an unmanned aerial vehicle (UAV) scenario show that, using ISAC technology, the proposed multi-target detection algorithm is effective.

Keywords

ISAC, multi-target, detection, location, double-threshold

1 Introduction

The pace of mobile communication innovation has never stopped. At present, 6G research is advancing rapidly. 6G will usher in a new era of deeper integration between the physical world and digital world, ultimately realizing the vision of building a world where all things are connected through their digital twins [1].

Integrated sensing and communication (ISAC) is a key technology that supports the 6G vision mentioned above, which leverages radio signals to coordinate communication and sensing functions in a unified design [2, 3]. Since the signal processing modules and the surrounding environment information are shared for both communication and sensing purposes, ISAC technology has certain advantages in improving spectral efficiency, reducing processing cost, and enhancing performance. This has recently motivated significant research interests in ISAC technology [2–8]. Present sensing use cases can be divided into three categories according to the sensing features:

- Detection and location. For intrusion detection use cases with low requirements for sensing KPIs (e.g., UAV, pedestrian, and animal intrusion into specific areas), the major KPIs include probability of successful detection, probability of false alarms, and location accuracy of any successfully detected targets [9, 10]. These use cases require low accuracy requirements. For example, the acceptable distance accuracy is usually 2 m or above. Further use cases based on positioning and location, such as UAV flight path management, assisted driving, and automated guided vehicle (AGV) path management, require higher accuracy, such as 1 m (or better) distance accuracy.

- Motion recognition. Such use cases include gesture recognition, health monitoring, motion monitoring, etc. They are mainly indoor use cases.

- Environment monitoring. These use cases monitor the environment and extract environmental information, typically for weather monitoring, environmental reconstruction, environmental imaging, etc. [11].

ISAC is envisioned to play a crucial role in 5G and 6G standardization and applications. Sensing use cases based on detection and location backed by 5G technology lay a solid foundation for deeper integration between sensing and communication, and more use cases (from demanding detection and location to motion recognition and environment monitoring) will emerge in the 6G era, further bridging the digital world with the physical one.

For sensing use cases based on detection and location, the first problem to be solved is channel modeling. A relatively mature communication channel is defined in 3GPP [12]. [13] has proposed a channel model for ISAC systems, with both the propagation impact and the reflection impact considered. However, it remains challenging to detect and locate multiple sensing targets in ISAC applications. As far as we know, for multi-target detection and location methods based on ISAC channel modeling, there is no comprehensive research so far.

As an enhancement to the research result submitted to the RAN#100 meeting [14], this paper proposes a heuristic multi-target detection and location algorithm to address the multi-target detection and location challenges presented by complex multi-target characteristics in ISAC systems. Based on the two thresholds, the proposed algorithm successfully detects and estimates the location information of multiple targets in a dense clutter environment, and delivers effective simulation results in a UAV scenario where ISAC technology is used.

2 Background and Problem Formulation

2.1 Channel Modeling of ISAC

In an ISAC system, the radio signal is transmitted by a transmission wireless device and reflected on the sensing target before the signal is received by the wireless receiver. Specifically, the sensing target in an ISAC system is not equipped with antennas and therefore not capable of receiving wireless signals. Thus, an ISAC channel model should consider the impact of both the direct wave and the reflection wave. A cascaded channel model by combining the propagation path and reflection path is described in [12]. Specifically, the channel model of ISAC systems consists of path loss calculation and fast fading modeling.

Path loss calculation:

Path loss refers to the amount of power loss introduced by the propagation environment or channel fading between the transmitter and receiver. The path loss value is highly correlated with the scenario and also significantly affected by whether the direct path from the transmitter to sensing target and the reflection path from the sensing target

to receiver are line-of-sight (LOS). Distance from the transmitter to the receiver passing the sensing target is considered an important factor for path loss calculation. In addition, the radio wave reflected on the sensing target can cause power loss, which is characterized using the radar cross section (RCS) [15–17]. In summary, the channel path loss of an ISAC system is calculated by

$$PL = PL^C(d_1) + 10 \log_{10} \left(\frac{c}{f}\right)^2 - 10 \log_{10}(4\pi) \quad (1)$$

$$+ PL^C(d_2) - 10 \log_{10} \sigma.$$

Where, PL^C denotes the path loss value calculated by the path loss calculation formula of the wireless communication system [14], d_1 denotes the distance between the transmitter and the sensing target, d_2 denotes the distance between the sensing target and the receiver, c denotes the speed of light, f denotes the carrier frequency, and σ denotes the RCS of the sensing target.

Fast fading modeling:

Fast fading modeling of the wireless communication system is based on clusters and rays. In the wireless communication system, different clusters and rays have different physical properties. Consider a channel from the transmission device to the receiving target containing N clusters, where each cluster contains M rays. In general, the extra LOS ray is added to the first cluster. The LOS ray embodies the transmission time, power, angle of arrival (AOA), and angle of departure (AOD) of the LOS direction from the transmission device to the receiving target. Then, the channel coefficient can be generated by merging the LOS ray parameters.

By exploiting the fast fading channel model of the wireless communication system, the fast fading channel model for ISAC systems is developed based on clusters and rays. In an ISAC system, the fast fading model between the transmitter and the receiver is considered as a whole. The

characteristics of sensing factors are mainly embodied in parameters including the propagation time, AOA, and AOD in the fast fading model. In detail, the propagation time of the LOS ray equals $(d_1 + d_2) / c$. The LOS direction of AOD and zenith angle of departure (ZOD) refers to the direction of the direct link from the transmitter to the sensing target. The LOS direction of AOA and zenith angle of arrival (ZOA) is the direction of the reflection link from the sensing target to the receiver. This ISAC fast fading model is detailed in [12].

For clutter in an ISAC system, its channel is modeled between the transmitter and the receiver based on the wireless communication system [14]. In real situation, the receiver always receives reflection echos from multiple sensing targets and clutter simultaneously.

2.2 MUSIC Spectrum

Multiple signal classification (MUSIC) is a super-resolution algorithm [18, 19], which is based on subspace decomposition. Based on orthogonality of the signal subspace and noise subspace, the spatial MUSIC spectrum can be constructed. By searching for the peaks in the MUSIC spectrum, signal-related parameters can be estimated.

2.3 Problem Formulation

For multi-target detection in an ISAC system, the receiver can receive multiple echos from multiple targets and clutter. In general, multiple distances between the different sensing targets and the transmitter and the receiver are always different. As well, the reflecting powers reflected by different sensing targets are different, causing different receiving powers of different sensing targets. Due to different distances between and different reflecting powers of the sensing targets, building a multi-target channel by merging multiple sensing targets is complicated. Figure 1
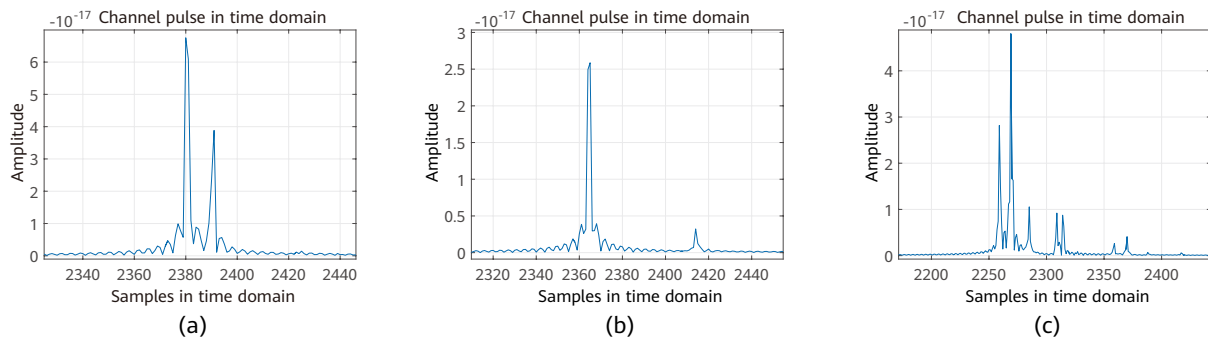

Figure 1 Multi-target channel pulse by merging two targets

provides three cases of multi-target channel pulses when two sensing targets are merged in the time domain, with different distance between two sensing targets and different reflecting powers in three cases. In Figure 1a, the receiving powers of the two sensing targets are consistent, and these two sensing target can be detected distinctly. In Figure 1b, the receiving powers of the two sensing targets are dramatically different, which leads to the detection of these two sensing targets intractable. In Figure 1c, due to the interference of multiple non-line-of-sight (NLOS) clusters, there are many glitches in the temporal channel, which cause a complex channel coefficient.

Therefore, our goal is to justify and detect multiple sensing targets though a complex multi-target channel in ISAC system, and then to obtain the location estimations of multiple sensing targets.

3 Multi-Target Detection and Location in an ISAC System

In this section, by analyzing the characteristics of the multi-target channel, we propose a double-threshold detection algorithm. The procedure of the proposed algorithm is described as follows.

Step 1: Setting the first threshold for the time channel coefficient

In this step, by setting a threshold for the multi-target time channel coefficient, we aim to obtain a time channel coefficient window containing multiple sensing targets. First, find the maximum amplitude value of the time channel coefficient and the corresponding index in time sampling domain. Then, according to the characteristics of the multi-target channel in the ISAC system, we can set a proportion X_1. X_1 is multiplied by the maximum amplitude

value of the time channel coefficient to obtain the first time channel coefficient threshold. Find amplitude values of the time channel coefficient larger than the threshold, and store the corresponding indexes. Let the minimum index of these saved indexes be the start of the time channel coefficient window, and let the maximum index of these saved indexes be the end of the time channel coefficient window. Figure 2 shows an example for setting the first time channel coefficient threshold and obtaining the corresponding time channel coefficient window, where the information about the two sensing targets is embodied within the time channel coefficient window.

Step 2: Setting the second threshold for MUSIC spectrum

In this step, the second threshold is set for the MUSIC spectrum corresponding to the time channel coefficient window. First, calculating the MUSIC spectrum of the multi-target channel within the time channel coefficient window. Find the maximum MUSIC spectrum value within the window and the corresponding index in time sampling domain. Then, we can set the other proportion X_2. X_2 is multiplied by the maximum MUSIC spectrum value to obtain the second MUSIC spectrum threshold. Find and store MUSIC spectrum values larger than the second MUSIC spectrum threshold, store the corresponding indexes as well. Figure 3 gives the MUSIC spectrum of the channel coefficient in Figure 2. For clarity, Figure 3a shows the complete MUSIC spectrum of the channel coefficient, and Figure 3b provides the MUSIC spectrum of the channel coefficient within the time channel coefficient window.

Step 3: Clustering multiple targets

In this step, we expect to classify the saved MUSIC spectrum values and their indexes. The MUSIC spectrum values with consecutive indexes are regarded from the same sensing
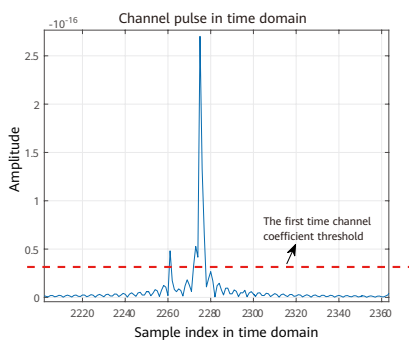

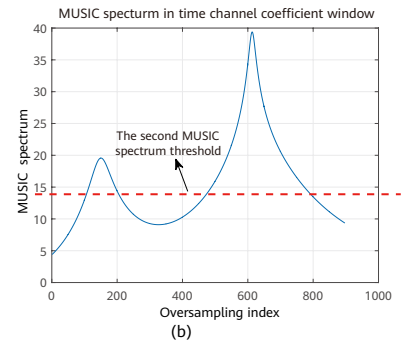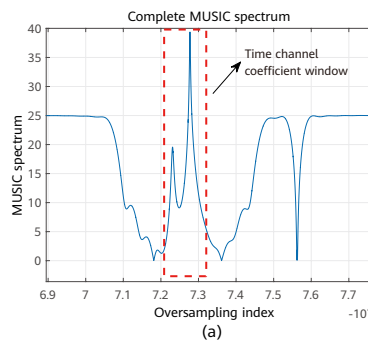**Figure 2** An example for setting the first time channel coefficient threshold

Figure 3 MUSIC spectrum of the channel coefficient in Figure 1, where (a) is the complete MUSIC spectrum, and (b) is the MUSIC spectrum within the time channel coefficient window
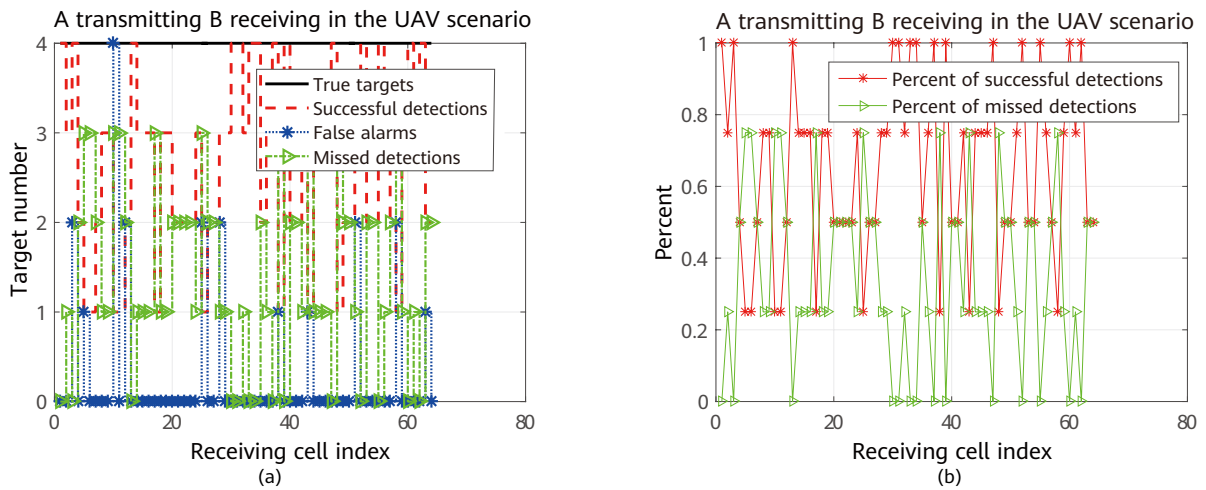
Figure 4 Detection number and detection probability per receiver

target. Therefore, we can obtain multiple clusters, each has several MUSIC spectrum values with consecutive indexes. For each cluster, search for the peak MUSIC spectrum value and the corresponding peak index. In general, the index of the MUSIC spectrum peak of a cluster represents the receiving time index of the reflection echo through a sensing target. That is, each cluster corresponds to a sensing target.

Step 4: Estimating locations of multiple targets

In this step, based on the MUSIC spectrum peaks of multiple clusters, we intend to estimate the location information of multiple sensing targets. For each cluster, a time of arrival (TOA) can be obtained by transforming the time sample index of the MUSIC spectrum peak to absolute time. Then, by multiplying TOA and speed of light, the relative distance between the transmitter and each sensing target and the receiver can be obtained. In addition, the AOA and ZOA can be estimated based on phase difference of multiple receiving antennas.

As described above, the heuristic multi-target detection and location algorithm is implemented by setting two thresholds.

4 Simulation Results

In this section, we demonstrate the detection and location performance of the proposed detection and location algorithm in a UAV scenario. To evaluate performance, the following metrics are defined:

- Number of successful detections: target number of successful detections for each receiver, which is per receiver.

- Probability of successful detections: target number of successful detections divided by total target number for each receiver, which is per receiver.

- Number of missed detections: target number of missed detections for each receiver, which is per receiver.

- Probability of missed detections: target number of missed detections divided by total target number for each receiver, which is per receiver.

- Number of false alarms: error number of determining non-targets as sensing targets, which is per receiver.

- Location accuracy: location accuracy of successful detected targets, involving distance accuracy, horizontal AOA accuracy, and ZOA accuracy.

In UAV evaluation scenarios, seven base stations, each including three cells, are positioned on a circumference. Eight UAVs are dropped in each cell evenly. The channel model between the transmitting and receiving cells is the RMa channel model, which is specified in [14]. The carrier
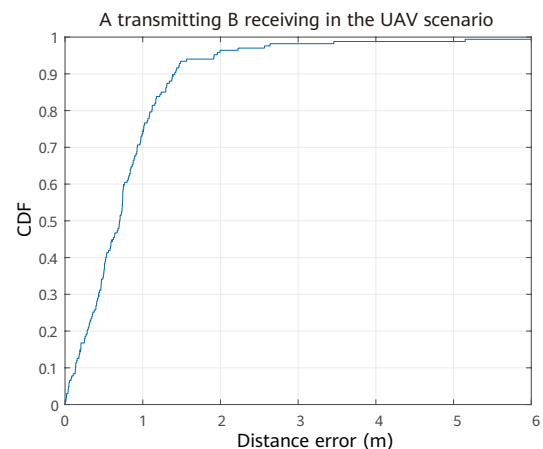

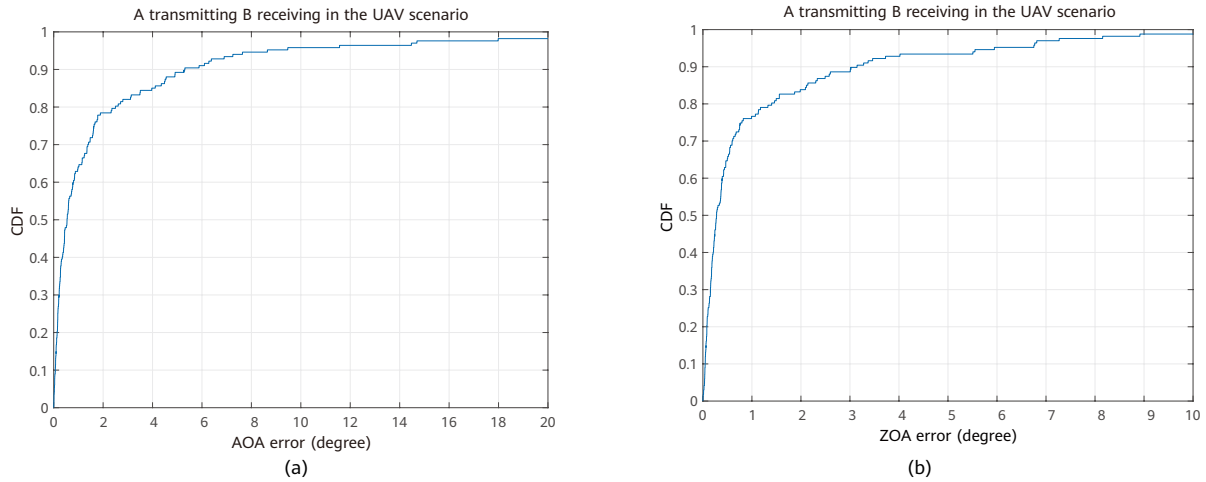**Figure 5** CDF curve of the distance error

Figure 6 CDF curves of angle errors, where (a) is AOA accuracy, and (b) is ZOA accuracy

frequency and subcarrier spacing of base stations are 6 GHz and 30 kHz, respectively. The system bandwidth is set to 100 MHz. The antenna height of base stations is 35 m. The height of the UAVs is 100 m. The antenna configurations of base stations are all [4,4,2,1,1]. The RCS of UAV is set to 0.2. The MUSIC algorithm [18, 19] is employed to obtain sensing estimation.

Figure 4 gives the detection number and detection probability per receiver. Figure 5 gives the distance accuracy. Figure 6 provides the AOA accuracy and ZOA accuracy. As shown in Figure 5, the distance accuracy is less than 1.5 m for 90% of sensing targets. In Figure 6, the AOA accuracy and ZOA accuracy are 6 degrees and 3 degrees for 90% of sensing targets, respectively. This shows that the proposed algorithm can yield satisfactory location performance, which demonstrates the effectiveness of the proposed algorithm.

5 Conclusion

This paper addresses the multi-target detection and location problem in ISAC systems. Considering the complex multi-target characteristics in an ISAC system, we propose a heuristic multi-target detection and location algorithm by setting two thresholds, which successfully detects multiple targets and estimates the multi-target location from a dense clutter environment. The ISAC simulation results in a UAV scenario prove that the proposed multi-target detection and location algorithm is effective.

6 Future Outlook

In the future, we call for further research on the ISAC channel model and a more accurate ISAC channel that combines the measured data, so as to lay a solid foundation for simulated evaluation. Additionally, it is necessary to further optimize the multi-target detection algorithm to extend it to more scenarios, such as strong and weak targets, multi-target sensing fusion in the network, and multi-mode multi-target sensing fusion.

References

[1] IMT-2030 (6G) Promotion Group: "6G vision and candidate technologies", 2021.

[2] Y. Wu, F. Lemic, C. Han *et al*, "Sensing Integrated DFT-Spread OFDM Waveform and Deep Learning-Powered Receiver Design for Terahertz Integrated Sensing and Communication Systems," *IEEE Transactions on Communications*, vol. 71, no. 1, pp. 595-610, 2023.

[3] J. A. Zhang, *et al*, "Enabling Joint Communication and Radar Sensing in Mobile Networks—A Survey," *IEEE Communications Surveys and Tutorials*, vol. 24, no. 1, pp. 306-345, 2022.

[4] W. Jiang, *et al*, "Improve Sensing and Communication Performance of UAV via Integrated Sensing and Communication," 2021 *IEEE 21st International Conference on Communication Technology*, pp. 644-648, 2021.

[5] J. Han, Z. Wei, L. Ma, *et al*, "A Multiple Access Method For Integrated Sensing and Communication Enabled UAV Ad Hoc Network," *IEEE Wireless Communications and Networking Conference*, 2022.

[6] Z. Wang, K. Han, J. Jiang, *et al*, "Multi-Vehicle Tracking and ID Association Based on Integrated Sensing and Communication Signaling," *IEEE Wireless Communications Letters*, vol. 11, no. 9, pp. 1960-1964, 2022.

[7] Z. Fang, *et al*, "A Silicon-Based Radio Platform for Integrated Edge Sensing and Communication Toward Sustainable Healthcare," *IEEE Transactions on Microwave Theory and Techniques*, vol. 71, no. 3, pp. 1296-1311, 2023.

[8] D. K. Pin, *et al*, "Integrated Sensing and Communication in 6G: Motivations, Use Cases, Requirements, Challenges and Future Directions," *1st IEEE International Online Symposium on Joint Communications and Sensing*, pp. 1-6, 2021.

[9] 3GPP TR 22.837: "Feasibility Study on Integrated Sensing and Communication".

[10] IMT-2020 (5G) Promotion Group: "Research Report on the Requirements for Integrated Sensing and Communication of 5G-Advanced", 2022.

[11] IMT-2030 (6G) Promotion Group: "Research Report on integrated sensing and communication technology", 2022.

[12] 3GPP TR 38.901: "Study on channel model for frequencies from 0.5 to 100GHz".

[13] IMT-2020 (5G) Promotion Group: "Research report on simulation evaluation method for integrated sensing and communication of 5GAdvanced", 2023.

[14] RWS-230283: "Discussion on sensing in NR Rel-19".

[15] A. Ioffe, W. Doerr, H. Yan, *et al*, "RCS characteristics of street curbs and the applications in automotive radar classification," *2016 European Radar Conference*, pp. 241-244, 2016.

[16] C. Meng, X. Wang, C. Ma, *et al*, "Fast and multi-band RCS measurement based on a microwave photonic inverse synthetic aperture radar," *20th International Conference on Optical Communications and Networks*, pp. 1-3, 2022.

[17] T. M. How, and Y. H. Lun, "Radar detector performance analysis using EM simulations of targets' RCS," *IEEE Radar Conference*, pp. 1-4, 2016.

[18] Xinrong Li, and K. Pahlavan, "Super-resolution TOA estimation with diversity for indoor geolocation," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 224-234, 2004.

[19] K. V. Rangarao and S. Venkatanarasimhan, "Gold-MUSIC: A Variation on MUSIC to Accurately Determine Peaks of the Spectrum," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 2263-2268, 2013.

Software and Hardware Verification Platforms for Cooperative Integrated Sensing and Communication

Dongsheng Xue, Kejia Ji, Jiachen Wei, Dingyou Ma, Qixun Zhang, Zhiyong Feng

Key Laboratory of Universal Wireless Communication (Ministry of Education), Beijing University of Posts and Telecommunication

{xuedongsheng, jkj2020, weijiachen, dingyouma, zhangqixun, fengzy}@bupt.edu.cn

Abstract

Integrated sensing and communication (ISAC) is one of the key technologies involved in Fifth Generation-Advanced (5G-A) and Sixth Generation (6G) wireless communication. It integrates communication and sensing software and hardware devices to sense various devices and analyze information on wireless networks. Centered on the communication functionality, the ISAC roadmap establishes the sensing functionality based on the Fifth Generation New Radio (5G NR) live network with low-cost networking and fast development. This paper describes the application scenarios of cooperative ISAC, and proposes the 5G NR-based ISAC frame structure design and signal processing method. This method uses reference signals that are flexibly mapped in the time and frequency domains to implement the sensing functionality, and develops the ISAC software and hardware function verification platforms for network cooperation, aiming to provide key technology verification in addition to macro and comprehensive performance statistics and evaluation for network cooperative ISAC. The software platform is designed to evaluate the communication and sensing system performance, and the hardware platform is designed to verify the sensing performance gain under the network cooperation mechanism. The multi-band multi-node ISAC performance evaluation capability is thus achieved.

Keywords

network cooperation, ISAC, software and hardware verification platforms

1 Introduction

With the rapid development of new use cases and services such as self-driving, smart industry, and smart city, future wireless communication networks require ubiquitous sensing capabilities in addition to high-performance communication capabilities. Aiming to meet such requirements, the integrated sensing and communication (ISAC) technology has attracted significant attention and research throughout the industry. By integrating communication and sensing software and hardware devices, ISAC uses unified spectra and integrated signals to implement both communication and sensing functions. In this way, it improves spectrum efficiency and reduces hardware costs in wireless device sensing and information analysis. Currently, ISAC is considered one of the key technologies involved in Fifth Generation-Advanced (5G-A) and Sixth Generation (6G) wireless communication [1]. Furthermore, it is able to promote the evolution of future communication networks to multi-functional networks.

The communication functionality is at the core of the ISAC technology roadmap, , in which sensing is native in communication networks. This technology enables the communication system to comprehensively understand the physical environment, assists in enhancing communication performance, and thus forms perceptive mobile networks (PMNs) [2]. The Fifth Generation New Radio (5G NR) cellular network is the first choice for implementing PMNs. It provides the sensing functionality based on the 5G NR live network, with the inherent advantages of low-cost networking and fast deployment [3–5]. Monostatic, bistatic, and distributed sensing modes can be implemented between base stations and user equipment (UE). By deploying bases stations at a high density, it is possible to ensure high sensing coverage. Sensing performance can be improved by using many sensing nodes, which can sense a target from multiple angles and in multiple frequency bands.

In terms of the communication system and the radar system, long-term research has been conducted on network cooperation methods [6]. The distributed radar technology [6] leverages the coherent phase difference of distributed node carriers to provide high resolution positioning under the premise of phase synchronization. In cooperative communication, a plurality of wireless devices in a mesh subnet may discover and connect to adjacent devices in order to offload network traffic and enhance radio coverage. In 6G networks, the large-scale distributed multi-input multi-output (MIMO) technology leverages scalable methods and low-cost densification to achieve coordination and cooperation between multiple cells or multiple transmission receiving points.

However, research on ISAC-oriented network cooperation is still in its infancy. This has created a series of key issues that need to be solved urgently (involving network interference management methods, multi-node clock synchronization, high dynamic environment awareness, low-complexity and high-precision sensing and estimation algorithms, and efficient cooperation mechanisms between network nodes) [6]. Currently, cooperative ISAC is limited to architecture design, lacking key technology verification and complex, fundamental, and comprehensive performance statistics and evaluation. In order to address the preceding challenges, the system-level ISAC hardware and software simulation platforms need to be developed for deep study of the network cooperative ISAC performance. The software platform must be able to support the simulation of a large-scale cellular sensing network, provide fast and efficient algorithm verification, and guide system R&D and implementation, and the hardware platform must be able to verify the key technologies in the network cooperative ISAC.

This paper first describes the application scenarios of network cooperative ISAC, proposes the 5G NR-based ISAC frame structure design and signal processing method, and uses the reference signals that are flexibly mapped in the time and frequency domains to implement the sensing functionality. The paper then proposes the module design and simulation process of the software platform of the cooperative ISAC, and evaluates the system-level communication and sensing performance tradeoff in the 3.5 GHz frequency band. This process designs and develops an ISAC hardware function verification platform for network cooperation, proposes a sensing network cooperation mechanism based on clustering algorithms, and verifies the cooperative sensing performance gain of this mechanism in the 26–29 GHz millimeter wave (mmWave) frequency band. The multi-band and multi-node ISAC performance evaluation is therefore realized through the capability of the software and hardware platforms.

2 Network Cooperative ISAC Scenarios and Signal Processing

This section presents the application scenarios of network

cooperative ISAC, and briefly describes the signal processing of single-site proactive sensing.

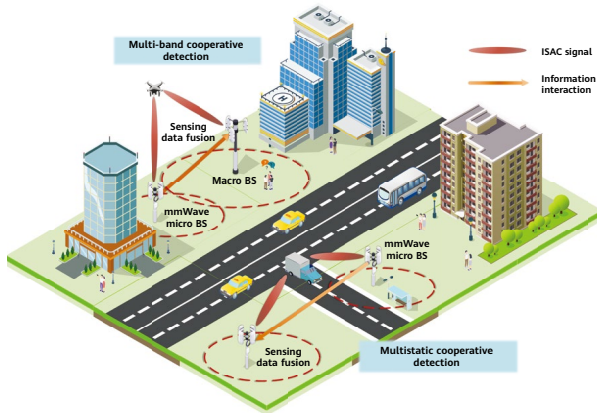## 2.1 Cooperative ISAC Scenarios

Figure 1 Cooperative ISAC scenarios

Future 5G-A/6G mobile networks are expected to integrate new functionalities such as positioning, sensing, and imaging, as well as enhancing the existing communication capabilities, in order to implement real-time mapping between the physical and digital worlds. To serve diversified application scenarios such as connected automated vehicle (CAV) and unmanned aerial vehicle (UAV) detection, as well as Internet of Things (IOT) equipment monitoring, multistatic and multi-band coordination is required to collect sensing parameters in an all-round and multi-dimensional manner.

Figure 1 shows a network cooperative ISAC scenario, in which, high- and low-frequency cooperative detection aims to collaborate between a macro base station in the 3.5 GHz band and an mmWave micro base station in the 26–29 GHz band to continuously detect the target. The macro base station preprocesses sensing information of the detected target, and transmits coarse positioning information of the target to the micro base station. The micro base station further performs precise positioning and tracking of the target, expanding the sensing range and enhancing the detection precision.

Multistatic coordinated detection is native in network-based coordination and in interference coordination capabilities. It has advantages in sensing coverage holes, reducing deployment costs, and coordinating interferences. Multistatic detection leverages monostatic and bistatic sensing to realize signal-level fusion. In monostatic sensing mode, a base station receives its own downlink sensing echoes to

proactively detect the target. In bistatic sensing mode, a base station receives its neighbor's downlink sensing echoes to passively detect the target. Multistatic preprocessed sensing data is fused through algorithms and then shared to the sensing center, so as to improve the detection precision, narrow sensing coverage holes, continuously track targets, and precisely detect objects that are small or difficult to detect.

2.2 ISAC Signal Frame Structure Design and Signal Processing

ISAC signals use the existing 5G NR cyclic prefix orthogonal frequency division multiplexing (CP-OFDM) waveform and signal frame structure. Specifically, sensing signals are added to the communication frame structure to compose an integrated frame structure for time division (TD)-based ISAC [7]. The ratio of communication signals to sensing signals can be changed in the frame structure to flexibly regulate and control the communication and sensing functions. Base stations may flexibly allocate resources based on sensing and communication service requirements, improving spectrum utilization efficiency, and implementing both communication and sensing functions without changing the existing communication protocol framework.

Figure 2 shows a schematic diagram of TD-based ISAC frame structure design. The frame structure is based on a 2.5 ms dual-period frame structure in 5G NR time division duplex (TDD) mode. It has three types of timeslots (Slot): downlink timeslot, flexible timeslot, and uplink timeslot. A downlink timeslot is divided into two parts: downlink sensing (S) and downlink communication (D), with adjustable symbol positions. An uplink timeslot retains its original uplink communication signal structure. In this way, it is possible to achieve uplink and downlink communication between the base station and UEs and downlink sensing of the base station.
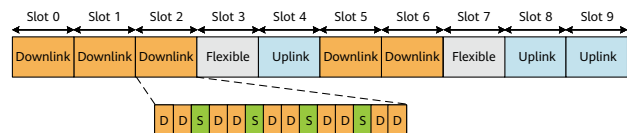

Figure 2 ISAC signal frame structure design

Figure 3 shows the signal processing for single-site proactive sensing in ISAC [8]. The ISAC base station sends an OFDM downlink signal through the communication channel. Upon receiving this signal, a UE performs clock synchronization,
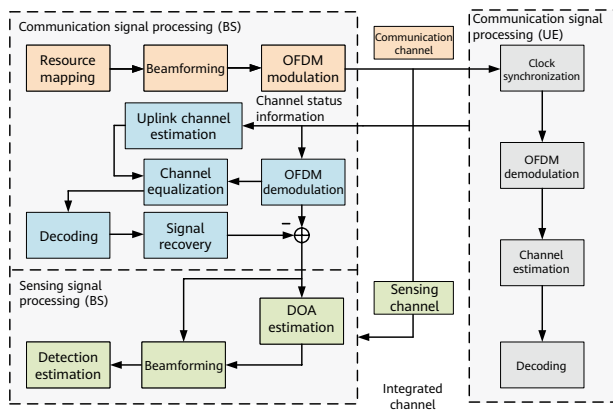
Figure 3 ISAC signal processing

3.1 Software Platform Design and Simulation Process

Figure 4 System-level ISAC simulation scenario

downlink channel estimation, demodulation, and decoding, and then feeds back channel state information (CSI) to the ISAC base station through an OFDM uplink signal. The base station receives the OFDM uplink signal and performs processing such as uplink channel estimation. Meanwhile, the OFDM downlink signal sent by the base station passes through the sensing channel, and generates an OFDM echo. After receiving the echo, the ISAC base station first performs interference cancellation on the OFDM uplink signal. It then estimates the direction of arrival (DOA) of the echo, performs beamforming in this direction to generate a delay Doppler spectrum, and finally detects the constant false alarm rate (CFAR) to estimate the distance and speed of the target, thus completing target positioning.

3 Software Platform Design and Performance Evaluation

This section describes the software platform developed on the basis of the semi-open-source MATLAB simulation platform [9]. The software platform is able to analyze and evaluate system-level communication and sensing network nodes, focusing on multiple typical frequency bands, such as the 2.4 and 3.5 GHz low-frequency bands and the 26–29 and 60 GHz high-frequency bands. It supports the construction of urban road networks and simulation of UE/target movement tracks.

The software platform uses the Monte Carlo simulation method to simulate multiple key performance indicators (KPIs) of communication and sensing. Communication KPIs include data packet delay, throughput, and block error rate (BLER), and sensing KPIs include detection probability and root mean square error (RMSE).
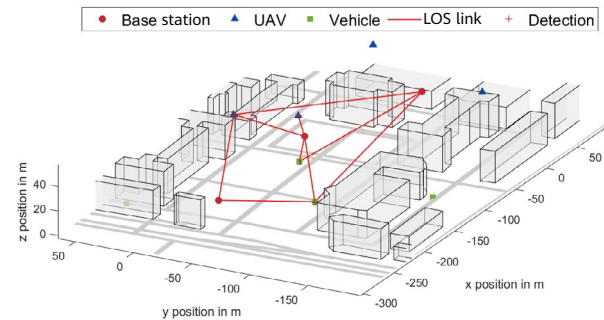
To support sensing simulation, the software platform provides multiple functional modules, such as sensing target modeling, line-of-sight (LOS) link channel modeling, sensing parameter estimation algorithms, and inter-cell interference signal modeling. Figure 4 demonstrates the visualized scenario generation capability of the software platform. In addition to supporting concurrent network simulation of multiple cells and user-defined track topologies of the base station and UEs/targets, the platform can import open-source Open Street Map™ road network information and simulate building blocks. The LOS link determines whether an obstacle exists between the base station and a UE/target. If no LOS link exists between the base station and a target, the target's current state cannot be effectively sensed.

Figure 5a shows the simulation process. First, we configure simulation parameters and sensing modes, define scenarios, and start simulation, which generates the communication and sensing network topology and initializes the communication and sensing channel model. Within a transmission time interval (TTI), the simulation allocates resources, and it respectively manages Rx and Tx beams based on the communication and sensing service models. This is performed to determine the mapping relationship between the base station and the UE/target. Then, the simulation determines whether a LOS channel exists between the base station and the UE/target according to the topology status of the UE/target. If a LOS channel exists, the LOS sensing channel is updated, and various interference modeling instances are added. After receiving the ISAC signal, the base station processes the signal, and obtains the target detection result in single-site or network cooperative sensing mode. Once simulations in all TTIs are complete, we can collect statistics on the performance evaluation indicators of communication and sensing in this scenario.
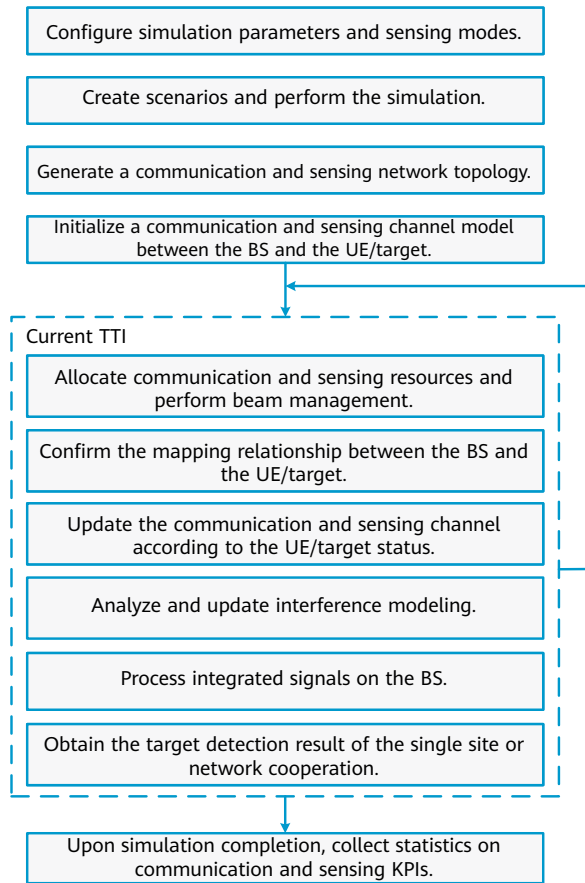
(a) ISAC software simulation process
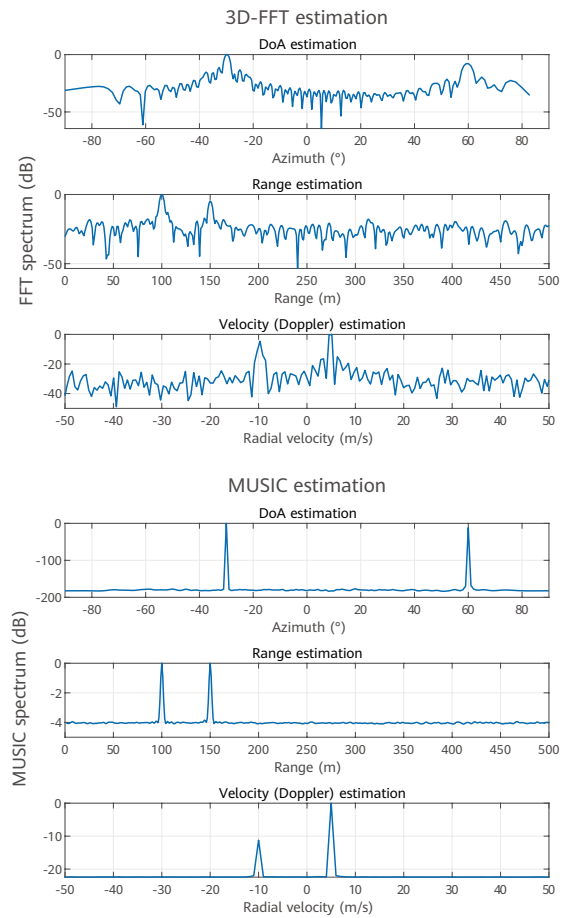
(b) 3D-FFT algorithm and MUSIC algorithm

Figure 5 ISAC software simulation process and sensing parameter estimation algorithm

The platform uses two typical radar signal processing algorithms [5]: three-dimensional fast Fourier transform (3D-FFT) algorithm [10] and multiple signal classification (MUSIC) algorithm [11]. Figure 5b shows the parameter estimation spectrum of the two algorithms. The estimated parameter value is the value on the x axis corresponding to the peak point on the y axis. In this value, the azimuths of the two targets are –30° and 60°, the distances are 100 m and 150 m, and the radial velocities are –10 m/s and 5 m/s, respectively.

3.2 System-Level Communication and Sensing Performance Evaluation

Simulation in this section is based on the 3.5 GHz frequency band on a 5G NR live network, adopting the demodulation reference signal (DMRS) of the physical downlink shared channel (PDSCH) as the sensing signal. Figure 6 shows the DMRS arrangement patterns in the type A resource mapping mode. In the simulation, a cell containing 20
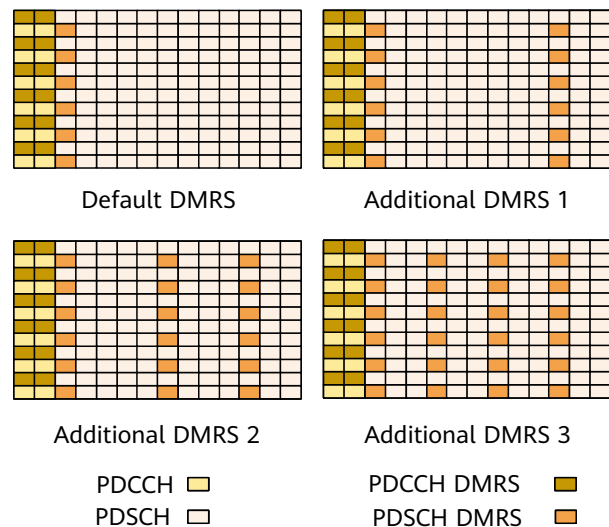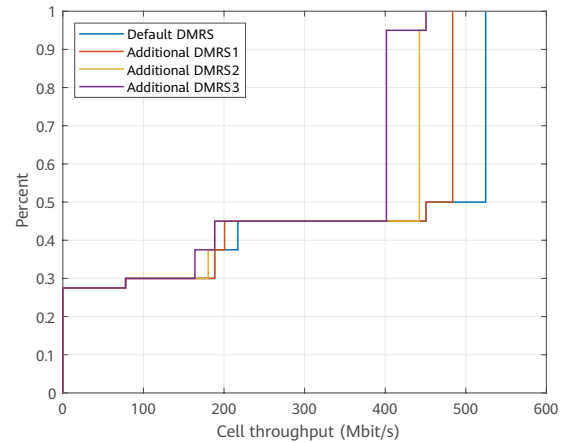
UEs that follow Poisson distribution is generated. The UEs are also detected targets, with an average radar cross section (RCS) of 1 m^2. For other simulation parameters, see Table 1. The measurement criteria for the performance of

Default DMRS   Additional DMRS 1

Additional DMRS 2   Additional DMRS 3

PDCCH □   PDCCH DMRS ■
PDSCH □   PDSCH DMRS ■

Figure 6 DMRS arrangement patterns in type A resource mapping mode

Table 1 Simulation parameters [12–15]

Parameter	Value
Carrier frequency	3.5 GHz
Subcarrier spacing	3 kHz
Bandwidth	100 MHz
Number of frames	1
Transmission waveform	CP-OFDM
Cyclic prefix length	Normal
Modulation scheme	QPSK
Antenna specifications	16 T 16 R
Transmit power	46 dBm
Antenna gain	25.5 dBi
Communication channel model	CDL-D
Communication path loss model	UMa
CFAR algorithm	Cell-averaging
False alarm rate	1×10^{-9}
Parameter estimation algorithm	3D-FFT

(a) ECDF of cell throughput

(b) ECDF of ranging RMSE

Figure 7 Tradeoff between the performance of communication and that of sensing

Communication and sensing are the empirical cumulative distribution function (ECDF) of cell throughput and ranging RMSE, respectively. ECDF is a statistical distribution function related to empirical measurements of samples. In an ECDF graph, the value on the y axis is the corresponding percent of the throughput/RMSE of less than or equal to a value on the x axis.

Figure 7a shows the system-level ISAC performance simulation results. When the physical resources occupied by PDSCH DMRSs increase, the overall cell throughput decreases. Figure 7b shows the ranging RMSE distribution in a cell with different DMRS configurations. A denser DMRS distribution indicates a smaller positioning error. One reason for this is that more PDSCH DMRSs occupy more data bit communication resources, thereby reducing the cell throughput. However, in sensing signal processing, more symbols bring more processing gain, thereby improving ranging precision. Therefore, when 5G NR reference signals are used to implement the sensing functionality, a tradeoff between the performance of communication and that of sensing needs to be considered.

4 Hardware Platform Design and Performance Evaluation

This section designs and builds an ISAC hardware prototype based on the mmWave 5G communication signals. Hardware resources such as basebands, radio frequencies, and antennas are shared in communication and sensing, and sensing information is carried in a TD frame structure in 5G NR, thereby implementing integration of communication and sensing. In the 26–29 GHz frequency band with 800 MHz bandwidth, the communication rate is greater than 2.8 Gbit/s and the target ranging error is less than 0.2 m.

4.1 Hardware Platform Architecture Design

Figure 8 shows the architecture of the ISAC-enabled cooperative sensing hardware platform. The cooperative sensing system includes two ISAC systems based on the
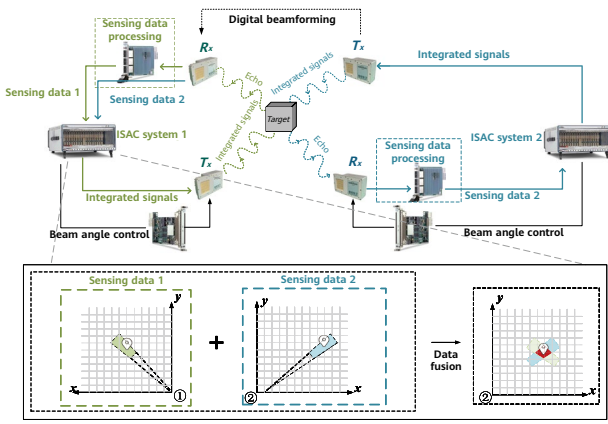
Figure 8 mmWave ISAC hardware platform oriented to cooperative sensing

4.2 Multistatic Cooperative Sensing Performance Evaluation

Sensing data is shared among multiple nodes to maintain data integrity and improve data fusion reliability. The distance between the two ISAC systems is set to 2.7 m, and the sensing target ADV moves at a constant speed of 1 m/s. In the data fusion phase, ISAC system 2 aggregates 50 observation samples generated by each ISAC system, and then uses the clustering algorithms to perform rapid searches and density peak discovery in order to remove unreliable samples and noises from the observation dataset. In this way, accurate target positioning is achieved. Specifically, the distance between any two observation sample points is computed after aggregation, and then the cluster density and the nearest-cluster distance of a random sample point are computed according to the truncated distance. This is performed in order to obtain a cluster decision graph. Then the sample point with the larger cluster density and nearest-cluster distance is marked as a cluster center, and the point with the smaller cluster density and nearest-cluster distance is marked as a noise. Finally the remaining sample points are allocated to adjacent clusters in order to achieve data-level fusion. Figure 9 shows the hardware test results. Fusing the sensing data of the two ISAC systems results in the cooperative ISAC system achieving a positioning RMSE of 0.1319 m, which is 61% lower than that of a single ISAC system. At the same time, the communication throughput is stable at 2.8 Gbit/s.
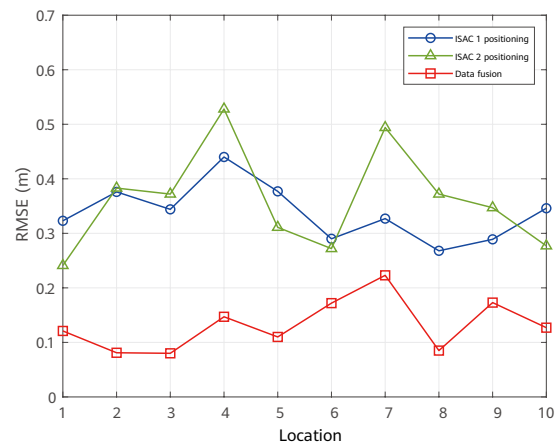
5G NR frame structure, one horn antenna, one corner reflector, and one automated driving vehicle (ADV). Each ISAC system includes two mmWave 5G NR communication platforms and two 64 phased array antennas (PAAs). The equivalent isotropically radiated power (EIRP) of the PAAs is set to 51 dBm, and the numbers of array elements in the horizontal and vertical directions are both 8. Internal clock signals are generated by the oven controlled crystal oscillator with a precision of 50 ns, and are shared to achieve clock synchronization between the two cooperative ISAC systems. The horn antenna on the ADV is set as the receiving end of each ISAC system. The two PAAs are placed on a fixed mmWave communication platform and serve as a transmitter and a receiver, connected to the Tx and Rx ends of the RF head, respectively. In the sensing phase, the ADV moving in a circular trajectory is regarded as the sensing target. The two ISAC systems measure the azimuths of the PAAs to obtain the target angles, send integrated signals, and receive and process echoes to obtain the target distances. Finally, the sensing data of ISAC system 1 and ISAC system 2 are shared and fused, improving the sensing precision.

In the design of the communication-centric ISAC system, PDSCH symbols are employed to implement 5G NR-based TD-ISAC, with the center frequency as 28 GHz. An aggregated bandwidth of 800 MHz is used for communication, among which 100 MHz is used for sensing. A communication frame contains 50 subframes for a total duration of 10 ms. First, the host computer generates an integrated waveform, then sends it to the baseband transmission and processing module for signal processing and digital-to-analog conversion, and finally converts the processed baseband signal to 28 GHz through the IF module and RF head. Subsequently, the PAA sends the RF signal for communication data transmission and receives the echo signal.

Figure 9 Comparison of positioning RMSE between a single ISAC system and cooperative ISAC systems

5 Summary and Outlook

This paper discusses the application scenarios, signal design, and processing of the cooperative ISAC. To address the cooperative ISAC challenges, this paper designs and develops the cooperative ISAC software and hardware verification platforms, which leverage the software part to quantify the system-level performance tradeoff between the communication and sensing functions, and the hardware part to verify that the proposed cooperative ISAC system improves the positioning accuracy by 61%.

In this paper, we do not delve sufficiently deep into the ISAC system in terms of discussing and researching multi-node efficient collaboration mechanism, data-level multistatic cooperative sensing, and signal-level sensing fusion technologies of the integrated network collaboration system. Comprehensive performance evaluation of the cooperative ISAC scenarios requires further research. The future software simulation platform is expected to incorporate the data-level multistatic cooperative sensing function, which will provide a powerful basis for the research on multi-base station and multi-band cooperative sensing methods. The future hardware platform needs to have the capability of verifying the multi-dimensional performance of time, frequency, and space for the ISAC system. For the software and hardware platforms, the function update and iteration will promote the standardization of research on communication and sensing convergence, boost the evolution of the ISAC technology from the point-to-point link level to the large-scale network level, and enable the industrialized application of communication and sensing.

Acknowledgments

References

[1] Z. Feng, Z. Fang, Z. Wei, X. Chen, Z. Quan and D. Ji, "Joint radar and communication: A survey," *China Commun*, vol. 17, no. 1, pp.1–27, Jan. 2020.

[2] J. A. Zhang, M. L. Rahman, K. Wu, *et al.*, "Enabling joint communication and radar sensing in mobile networks—A survey," *IEEE Commun. Surv. Tutor*, vol. 24, no.1, pp. 306–345, 1st Quart. 2021.

[3] D. Ma, N. Shlezinger, T. Huang, Y. Liu and Y. C. Eldar, "Joint radarcommunication strategies for autonomous vehicles: Combining two key automotive technologies," *IEEE Signal Process. Mag*, vol. 37, no. 4, pp. 85–97, Jul. 2020.

[4] F. Liu, Y. Cui, C. Masouros, *et al.*, "Integrated sensing and communications: Towards dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun*, vol. 40, no. 6, pp. 1728–1767, Jun. 2022.

[5] Z. Wei, *et al.*, "Integrated sensing and communication signals towards 5G-A and 6G: A survey," *IEEE Internet Things J*, pp. 1–1, Jan. 2023.

[6] L. Xie, S. Song, Y. C. Eldar and K. B. Letaief, "Collaborative Sensing in Perceptive Mobile Networks: Opportunities and Challenges," in *IEEE Wirel. Commun*, vol. 30, no. 1, pp. 16–23, Feb. 2023.

[7] Q. Zhang, H. Sun, X. Gao, X. Wang and Z. Feng, "Time-division ISAC enabled connected automated vehicles cooperation algorithm design and performance evaluation," *IEEE J. Sel. Areas Commun*, vol. 40, no. 7, pp. 2206–2218, Jul. 2022.

[8] X. Chen, Z. Feng, Z. Wei, J. A. Zhang, X. Yuan and P. Zhang, "Concurrent Downlink and Uplink Joint Communication and Sensing for 6G Networks," in *IEEE Trans. Veh. Technol.(Early Access)*, Feb. 2023.

[9] M. K. Muller, F. Ademaj, T. Dittrich, A. Fastenbauer, B. R. Elbal, A. Nabavi, *et al.*, "Flexible multi-node simulation of cellular mobile communications: The Vienna 5G system level simulator," *Eurasip J. Wireless Commun. Networking*, vol. 2018, no. 1, pp. 227, Sep. 2018.

[10] C. Sturm and W. Wiesbeck, "Waveform design and signal processing aspects for fusion of wireless communications and radar sensing," *Proc. IEEE*, vol. 99, no. 7, pp. 1236–1259, Jul. 2011.

[11] X. Chen, *et al*., "Multiple Signal Classification Based Joint Communication and Sensing System," in *IEEE Trans. Wirel. Commun.(Early Access)*, Feb. 2023.

[12] 3GPP TS 38.211, "Physical channels and modulation," V17.3.0, Sep. 2022.

[13] 3GPP TS 38.213, "Physical layer procedures for control," V17.5.0, Mar. 2023.

[14] 3GPP TS 38.214, "Physical layer procedures for data," V17.5.0, Mar. 2023.

[15] 3GPP TR 38.901, "Study on channel model for frequencies from 0.5 to 100 GHz," V17.0.0, Mar. 2022.

An Exploration on Integrated Sensing and Communication for the Future Smart Internet of Things

Zhaoxin Chang [1, 4], Fusang Zhang [2], Daqing Zhang [3, 4]

[1] Télécom SudParis

[2] Institute of Software of the Chinese Academy of Sciences

[3] Peking University

[4] Polytechnic Institute of Paris

zhaoxin.chang@telecom-sudparis.eu, fusang@iscas.ac.cn, dqzhang@sei.pku.edu.cn

Abstract

Internet of Things (IoT) technologies are the foundation of a fully connected world. Currently, IoT devices (or nodes) primarily use dedicated sensors to sense and collect data at large scales, and then transmit the data to target nodes or gateways through wireless communication for further processing and analytics. In recent years, research efforts have been made to explore the feasibility of using wireless communication for sensing (while assiduously improving the transmission performance of wireless signals), in an attempt to achieve integrated sensing and communication (ISAC) for smart IoT of the future. In this paper, we leverage the capabilities of LoRa — a long-range IoT communication technology, to explore the possibility of using LoRa signals for both sensing and communication. Based on LoRa, we propose ISAC designs in two typical scenarios of smart IoT, and verify the feasibility and effectiveness of our designs in soil moisture monitoring and human presence detection.

1 Introduction

The rapid development of the Internet of Things (IoT) enriches information exchange applications both between people and things and among things themselves. IoT also enables diversified and intelligent management for these applications by leveraging various sensing devices and wireless communication technologies. In today's applications, however, IoT devices mainly rely on dedicated sensors to sense and collect data, and then transmit the data to target nodes or gateways via wireless communication signals. In recent years, research has demonstrated the sensing capability of wireless communication signals used in the IoT. Analyzing the changes in such wireless signals allows us to sense the status information of people and the environment around IoT devices. With this dual capability of wireless signals, we can explore integrated sensing and communication (ISAC) technologies that enable future IoT devices to support both communication and sensing and consequently replace traditional sensor devices. ISAC technologies create new possibilities for the development of smart IoT. Leveraging wireless signals for sensing not only reduces our reliance on dedicated sensors, but also simplifies the processes of device deployment and maintenance. Utilizing the inherent communication function of IoT devices also makes it possible to transmit sensing data to gateways and the cloud in real time, which is crucial for prompt responses and real-time decision-making.

LoRa is a highly promising IoT technology due to its low power consumption and support for long-range transmission. In urban environments, for instance, LoRa devices can communicate with each other over distances of several kilometers. The long-range capabilities of LoRa make it particularly advantageous in IoT applications and are the reason why it is primarily used in outdoor scenarios. In recent years, researchers have utilized LoRa signals to sense human vital signs in long-range scenarios, demonstrating the unique advantages of LoRa over other types of wireless signals in long-range sensing [1–4]. This paper studies the technical feasibility of applying LoRa signals as an ISAC approach for smart IoT. We explore the sensing and communication requirements in two typical scenarios, and propose corresponding LoRa-based ISAC designs. One design is optimized for outdoor scenarios that require low power consumption more than continuous sensing, whereas the other is tailored for indoor scenarios that require continuous sensing more than low power consumption. For the outdoor and indoor scenarios, we use

soil moisture monitoring and human presence detection as separate examples to verify the feasibility of the two designs. Our proposal and verification provide a preliminary solution to the integration of sensing and communication for the future smart IoT.

2 Background

2.1 LoRa Communication

LoRa signals utilize chirp spread spectrum (CSS) modulation. CSS involves a sequence of chirps whose frequencies increase linearly with time. Each chirp serves as the fundamental data encoding unit for LoRa signals, with all frequencies within a given bandwidth being used to encode the same data. This maximizes the use of the available bandwidth, allowing LoRa signals to overcome attenuation and interference during transmission. As a result, LoRa technologies can support long-range communication over significant distances. As shown in Figure 1, a LoRa data packet consists of the preamble, start frame delimiter (SFD), and data payload. The preamble contains several uplink chirps used for synchronization between the transmitter (a LoRa node) and receiver (a LoRa gateway). The SFD contains several downlink chirps that indicate the start of the payload. By changing the start frequency of chirps, we can realize data encoding and transmission of the payload.

A typical large-scale LoRa network consists of multiple LoRa nodes and LoRa gateways [5]. LoRa nodes collect data from sensors and transmit the data to the gateways. Because LoRa nodes usually transmit small data packets at low frequencies, they mostly remain in sleep mode to conserve power. The LoRa protocol stipulates that the duty cycle for transmitting data from a LoRa node should be kept low, usually not more than 1% [6]. A low duty cycle helps reduce data packet collisions or conflicts among different nodes, ultimately improving the reliability and efficiency of data transmission. In addition, LoRa nodes have a long battery
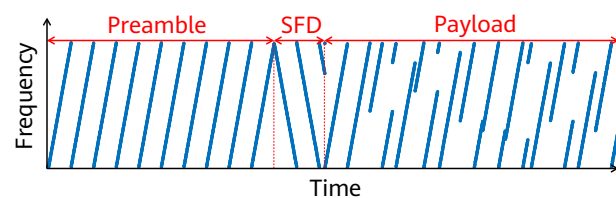

Figure 1 LoRa data packet structure

life, lasting for months or even years, making them an ideal option for remote and long-term deployment in wireless sensor networks.

2.2 LoRa Sensing

Similar to other wireless sensing technologies, LoRa sensing operates by analyzing the relations between received signals and target states. LoRa nodes transmit signals to a LoRa gateway through different paths in an environment. The transmission paths can be straight, reflective, diffractive, or scattering. The signals from these paths are superposed at the receiver to form the received signals. By analyzing the amplitudes and phases of the received signals, we can restore the information of these paths. For instance, we can analyze the changes in the amplitudes and phases of the received signals over time to track the movements of a target person. Additionally, we can analyze the speeds at which the signals are transmitted along straight paths to obtain the properties of the materials that the signals have passed through.

3 LoRa-based ISAC Designs

In this section, we propose two LoRa-based designs for two typical scenarios. The first design is intended for outdoor scenarios that require low power consumption more than continuous sensing. To implement sensing in this design, we propose reusing the existing low-rate data packets used for wireless communication. This achieves low-power ISAC because no additional data transmission is needed for sensing. We demonstrate the feasibility of this design by using soil moisture monitoring as an example. The second design focuses on indoor scenarios that require continuous sensing more than low power consumption. We propose adding sensing-dedicated data packets at the transmitter and employing the multi-antenna division operation at the receiver to eliminate amplitude and phase information irrelevant to sensing. This way, we satisfy both communication and continuous sensing requirements, although at the cost of increased power consumption. We have explored the feasibility of this design through human presence detection, where we can determine the walking and still states of people indoors.

3.1 ISAC Design by Reusing Communication Data Packets, with Soil Moisture Monitoring as an Example

Soil moisture monitoring plays an important role in smart agriculture. Its implementation is based on the principle that changes in soil moisture levels affect the transmission speeds of wireless signals in soil. When LoRa signals are transmitted over a fixed distance in soil, the varying soil moisture levels cause different degrees of phase change. By measuring the phase change, we can calculate the soil moisture level.

In smart IoT applications, ISAC designs should consider both sensing and communication requirements. To address the sensing aspect, we collect original signals from signal receivers (i.e., gateways) and extract phase information from the signals. When monitoring soil moisture, it is crucial to establish a fixed distance for LoRa signal transmission within the soil and then measure the phase change of signals transmitted over this distance to monitor soil moisture levels. Traditionally, a LoRa transmitter (or node) is usually configured with only one antenna, which can result in non-fixed signal transmission distances in the soil because the distance between the antenna and the soil surface is uncertain. To address this issue, we propose a dual-antenna hardware design for LoRa nodes. This design employs a radio frequency (RF) switch to extend the originally single transmit channel of a LoRa node to two channels, each connected to one antenna. The two antennas are placed in the soil with a known distance between them. LoRa signals are sent from both antennas and then transmitted to a LoRa gateway. At the gateway, the phase difference between the signals sent from the two antennas is measured to estimate the speed of signal transmission in the soil, which is then used to calculate the soil moisture.

In outdoor scenarios, an ISAC design must satisfy the low-power requirements of LoRa communication. To reduce power consumption and extend the battery life of LoRa nodes, the LoRa communication protocol typically allows for a very small duty cycle (specifically, smaller than 1%) for a single LoRa node to transmit signals. While adding data packets for sensing is a direct solution, it will inevitably increase the duty cycle and power consumption. Nonetheless, we have observed that soil moisture monitoring has a low requirement for continuous sensing because the soil

moisture often remains stable for a certain period (e.g., one hour). Therefore, when the requirement for low power consumption is high and that for continuous sensing is low, we propose using communication data packets for sensing to avoid the need for adding extra data packets. However, we need to ensure that the newly introduced sensing function does not interfere with the communication function that the data packets are originally intended for. To achieve the dual-antenna design of LoRa nodes, it is necessary to obtain the phase difference between the two antennas placed at different depths in the soil. This can be done by having the two antennas send signals at different times instead of simultaneously. Therefore, during the transmission of a LoRa data packet, we need to switch the antennas in the middle of the transmission process. To minimize the impact of sensing on communication, we should determine an appropriate timestamp for the switching to occur during the transmission of one data packet. As shown in Figure 1, a LoRa data packet includes the preamble and payload, with the data to be transmitted being encoded in the payload. To ensure that the antenna switching will not affect data decoding, we choose to switch the antenna in the preamble part. Figure 2 shows the soil moisture monitoring scenario

and its antenna switching process. Before sending a data packet, we reset the RF switch to restore the connection to antenna 1, and start transmitting the preamble part. In the middle of this transmission process, we trigger the RF switch to use antenna 2 for subsequent transmission, which includes the remainder of the preamble that has not yet been transmitted, SFD, and payload. See [7] for details about the design.

We deployed our prototype system in a real-world scenario, as shown in Figure 3b, where LoRa nodes and antennas were placed in the soil. Figure 3c shows the soil moisture measurement precisions achieved at different distances from a LoRa node to the gateway. Our findings indicate that the system's maximum sensing range, i.e., the maximum distance between a LoRa node and the LoRa gateway, can reach up to 100 meters. At this distance, the soil moisture measurement error is about 13%, demonstrating our ISAC system's effectiveness and accuracy in real-world scenarios. We also evaluated the system's communication performance when the sensing function was in operation. The gateway decoded data packets with 100% accuracy, proving that our sensing design does not affect the original communication function.
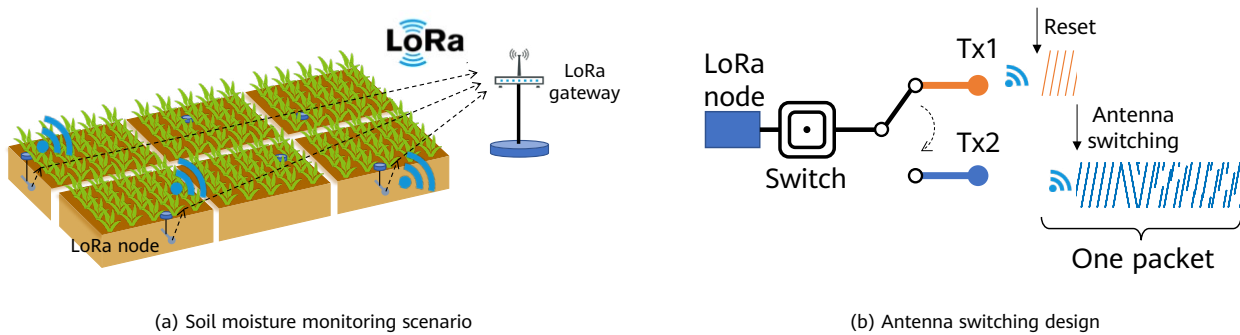

(a) Soil moisture monitoring scenario

(b) Antenna switching design

Figure 2 Soil moisture monitoring and antenna switching

(a) Experiment setup

(b) Deployment of LoRa nodes and antennas

(c) Soil moisture measurement precisions

Figure 3 Experiment setup and results

3.2 ISAC Design for Continuous Sensing, with Human Presence Detection as an Example

Human presence detection is widely used in real-world applications, including smart buildings, security surveillance, and behavior recognition. This detection method primarily relies on the movement of the human body, specifically the walking and still states. The use of wireless signals allows us to detect human presence because the movement of a human body changes the transmission path and distance of the signals reflected by the human body. As a result, the received signals will change with time. By analyzing the changes in the amplitudes and phases in the detection signals, we can determine if human bodies are in the walking or still state in a specific environment. This enables us to achieve human presence detection.

The preceding fundamental principle tells us that continuous sensing, which involves repeatedly transmitting signals, is crucial to ensuring accuracy and preventing missed detections. However, the low duty cycle (typically smaller than 1%) currently stipulated in the LoRa communication protocol does not meet the requirements for continuous sensing. Based on our analysis in the previous section, the limited duty cycle in LoRa communication serves an important purpose in most outdoor application scenarios. It ensures low power consumption, extends battery life, and reduces system maintenance costs. Fortunately, in indoor scenarios, continuous power supply to LoRa nodes is often guaranteed, and there is little requirement for low power

consumption. This allows us to develop an ISAC design for continuous sensing in indoor scenarios. Specifically, we enable LoRa nodes to continuously send null data packets that are specially used for sensing, in addition to the original data packets used for communication. Although null packets do not contain any valid data, we can determine human presence by detecting the amplitude and phase changes of these packets at the receiver. Yet there are difficulties in detecting these changes. First, LoRa signals using CSS modulation experience constant frequency variations, resulting in amplitude and phase changes over time. Second, random phase changes can occur due to carrier and sampling frequency offsets in the received signals, because the clocks are not synchronized between LoRa nodes and the gateway. Therefore, in the received signals, we need to eliminate the changes brought by signal modulation and clock asynchronization, and retain only the amplitude and phase changes that are introduced as a result of the varying signal reflection paths. To solve these two issues, we propose configuring two antennas on the LoRa gateway to receive signals. The two antennas are synchronized by the same clock source, resulting in identical phase offsets and modulation characteristics for the signals they receive. However, due to the two antennas being deployed in slightly different spatial locations, the sensing signals they receive may contain minor amplitude and phase changes. By performing a division operation on the two copies of signals received by the two antennas, we can effectively eliminate information irrelevant to sensing. See [1] for details about the principles and signal processing algorithms. Figure 4a illustrates the scenario of human presence detection in a smart building. We deploy a
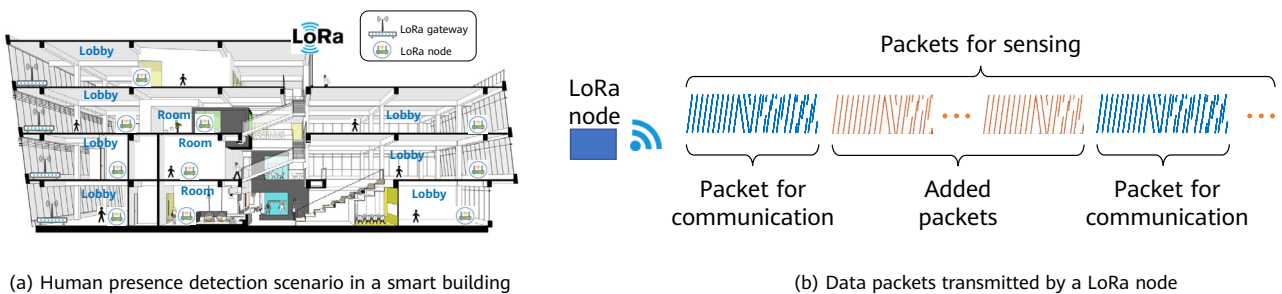

(a) Human presence detection scenario in a smart building

(b) Data packets transmitted by a LoRa node

Figure 4 Human presence detection and related LoRa data packets
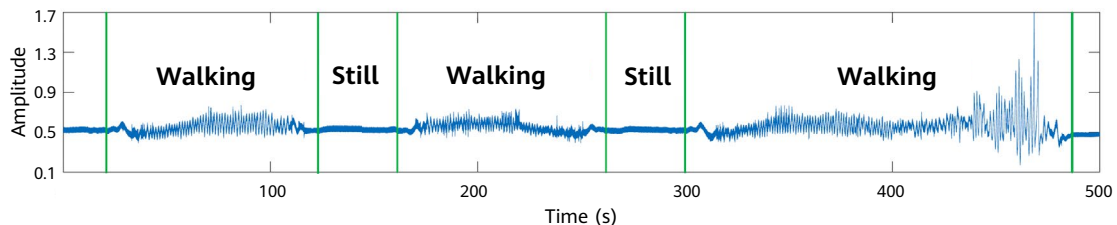

Figure 5 Signal waveforms for the walking and still states of persons

LoRa gateway on each floor and a LoRa node in every room. This setup allows each LoRa node to detect human presence in its respective room. Figure 4b illustrates the transmission of LoRa data packets based on our ISAC design. In our experiment, we deployed a LoRa node and gateway pair in a 30-meter-long corridor. As shown in Figure 5, the signals obtained through the division between the two antennas exhibit significant fluctuations over time when a person walks in a room and remain constant when a person is still. These signal waveforms enable us to detect human presence and determine their walking and still states.

In the preceding ISAC designs, we used either existing communication packets (in which valid data is encoded) or dedicated null packets to implement sensing. It is noteworthy, however, that the large number of LoRa nodes distributed in a smart building poses a major challenge in eliminating interference between the signals transmitted by various nodes. This can be addressed based on our observation that LoRa nodes can work at different frequencies. When we enable LoRa nodes to transmit signals at different frequencies, a LoRa gateway can receive these signals simultaneously without interference. Alternatively, through proper protocol design, we can ensure that the LoRa nodes working at the same frequency transmit data packets at different time slices to prevent signal conflicts. We have also explored using more antennas on a gateway to implement beamforming for received signals. Beamforming can distinguish signals that come from different spatial orientations, preventing interference and supporting multi-target sensing [2].

4 Conclusion

In this paper, we have explored ISAC approaches for the smart IoT of the future, and proposed two LoRa-based designs to meet sensing requirements in two typical scenarios. We have also verified the feasibility of our designs through examples of soil moisture monitoring and human presence detection. We have demonstrated that the wireless signals commonly used in the IoT can be used for sensing the environment and detecting human presence, in addition to their communication purpose. Such ISAC designs based on wireless IoT signals offer several advantages. First, there is no need for extra sensor devices, reducing costs and complexity in system deployment and maintenance when compared to traditional IoT applications. Second, wireless signals can travel long distances and cover large areas, providing both sensing and communication functions with long-range support and wide coverage.

References

[1] F. Zhang, Z. Chang, K. Niu, J. Xiong, B. Jin, Q. Lv, and D. Zhang, "Exploring LoRa for long-range through-wall sensing," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 2, pp. 1–27, 2020.

[2] F. Zhang, Z. Chang, J. Xiong, R. Zheng, J. Ma, K. Niu, B. Jin, and D. Zhang, "Unlocking the beamforming potential of LoRa for long-range multi-target respiration sensing," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 2, pp. 1–25, 2021.

[3] B. Xie, Y. Yin, and J. Xiong, "Pushing the limits of long range wireless sensing with LoRa," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 3, pp. 1–21, 2021.

[4] B. Xie, M. Cui, D. Ganesan, X. Chen, and J. Xiong, "Boosting the long range sensing potential of LoRa," in *Proceedings of the 21st Annual International Conference on Mobile Systems, Applications and Services*, 2023, pp. 177–190.

[5] A. Augustin, J. Yi, T. Clausen, and W. M. Townsley, "A study of LoRa: Long range & low power networks for the internet of things," *Sensors*, vol. 16, no. 9, p. 1466, 2016.

[6] J. C. Liando, A. Gamage, A. W. Tengourtius, and M. Li, "Known and unknown facts of LoRa: Experiences from a large-scale measurement study," *ACM Transactions on Sensor Networks (TOSN)*, vol. 15, no. 2, pp. 1–35, 2019.

[7] Z. Chang, F. Zhang, J. Xiong, J. Ma, B. Jin, and D. Zhang, "Sensor-free soil moisture sensing using LoRa signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies,* vol. 6, no. 2, pp. 1–27, 2022.

Evolution of Balong Modems

Xin Bai, Zhiyong Zhang, Jianghua Liu
Balong Solution Dept, HS

1 Introduction

A modem, short for modulator-demodulator, is an electronic device that modulates digital signals into analog signals for transmission and demodulates received analog signals into digital signals. Modems can be classified into different types based on their intended applications, such as cable modems, digital subscriber line (DSL) modems, and cellular modems (e.g., 2G, 3G, 4G, and 5G modems). This paper describes our technical thoughts on the evolution of cellular modems.

The Balong modem is a cellular mobile communications chipset that provides industry-leading communication solutions for user equipment (UE), making the communication experience a key strength of Huawei mobile phones. After a number of years of the Balong team's efforts, the Balong modem has been playing an important role in the mobile UE industry. Equipped with advanced competitive technologies, the Balong modem leads the continuous evolution of modem specifications in the industry. In the 4G era, Huawei released the industry's first Long Term Evolution (LTE) UE chipsets, including Balong 710 for LTE Category 4, Balong 720 for LTE Category 6, Balong 750 for LTE Categories 12 and 13, and Balong 765, which supports LTE Category 19 and LTE-V. In 2019, Huawei released the Balong 5000, the world's first 7 nm 5G multimode UE chipset. This chipset marked a significant milestone in the industry — it was the first to support 5G non-standalone (NSA) and standalone (SA) mode. Furthermore, it was the first chipset that passed all the UE chipset test items defined in the 5G enhanced technology research and development test activity organized by the IMT-2020 (5G) Promotion Group [1]. Huawei mobile phones equipped with the Balong 5000 chipset emerged as the top performers in China Mobile's initial assessment of 5G mobile phone communication capabilities. The assessment covered various aspects such as data transmission, voice quality, and power consumption, and Huawei's mobile phones powered by Balong 5000 outperformed competitors in all these areas [2]. This significantly contributed to advancing the large-scale commercial use of 5G SA technology.

Following the release of the Balong 5000 chipset, the Balong team continuously optimize and iterate the design of Balong modems in line with the evolution of 5G standards to maintain communication competitiveness. The 3rd Generation Partnership Project (3GPP) has developed specifications for 5G New Radio (NR) that have evolved to the 5G-Advanced phase (also known as 5G-A) starting from Release 18, covering Releases 19 and 20. In addition to covering the three major application scenarios of 5G — enhanced Mobile Broadband (eMBB), massive machine-type communications (mMTC), and ultra-reliable low-latency communication (URLLC) — Huawei proposes that 5G-A also covers three new application scenarios: real-time broadband communication (RTBC), uplink-centric broadband communication (UCBC), and harmonized communication and sensing (HCS). Essentially, 5G-A is a smooth evolution of 5G that builds on the basic framework of 5G standards to support new features. Furthermore, it is a means to explore and implement new services and technologies for commercial use.

The industry is currently discussing the vision of 6G, including the framework, requirements, technical capabilities, candidate technologies, and roadmap. In February 2020, the ITU-R Working Party 5D (WP 5D) initiated the "Future technology trends of terrestrial IMT systems towards 2030 and beyond" activity, which garnered significant interest and active engagement. In December 2022, the working party released the associated research report on technology trends [3]. After a year of discussion, the overall schedule for IMT-2030 was finalized at a workshop at the ITU-R WP 5D meeting #41. The schedule mainly comprises three phases: vision, technical requirements and evaluation methodology, and technical specification certification [4]. In the recent two years, many companies, standards developing organizations (SDOs), and industry experts have published a range of white papers and technical reports on IMT-2030 or 6G [5–15]. The papers illustrate their views on 6G from the perspective of vision, emerging services, technical requirements, and candidate technologies.

ITU-R WP 5D has arrived at a preliminary consensus [16] on the typical scenarios and technical capabilities of IMT-2030: eMBB, mMTC, and URLLC will continue to evolve along with new services. Furthermore, three other new scenarios will be explored, including integration of artificial intelligence (AI) and communications, integration of sensing and communications, and ubiquitous coverage.

The most anticipated new 6G services are real-time broadband services that demand high bandwidth and low latency, including immersive cloud extended reality (XR), holographic communication, and digital twin [3, 6, 9, 11, 15]. These services will considerably impact the network and UEs due to their high requirements for the UEs' processing capabilities and ultra-low end-to-end (E2E) latency. An obvious trend is to offload part of computing capabilities, such as rendering, from UEs to cloud processors, thereby

reducing the computing load and energy consumption at the UE side [3, 6, 9, 14].

During the evolution of mobile cellular communication systems, each generation is expected to have significantly stronger capabilities than the previous one. These technical advancements continue to define the key performance indicators of IMT-2030, which is expected to outperform IMT-2020 by 10 to 100 times in many aspects [6, 11, 12, 14, 16]. Notably, the baseline performance of wireless mobile communications has already reached a superior level, and the advantages that can be obtained from the evolution of semiconductor process is slowing down significantly [17]. Therefore, for modems, we must consider whether to continue pursuing a higher peak data rate in line with the network's peak data rate. As a result, simplifying UEs has become a significant priority for the development of 6G technology [8].

On the basis of the preceding six typical scenarios, another consensus is that green and low-carbon communication and trustworthiness should be considered as the basic requirements for 6G networks [6, 7, 9, 10, 12, 15]. Energy saving for network infrastructure remains a significant industry concern. It is also crucial to prioritize energy saving on the UE side. With the advent of 6G services, the demand for data rates and processing capabilities of UEs is increasing rapidly. While advancements in semiconductor processes have resulted in energy-saving benefits, the power required for these new services will eventually outweigh these benefits [17].

From Huawei's vision for 5G-A and the industry's preliminary viewpoints on 6G, we can observe that 5G-A and 6G have some similarities in terms of the exploration of emerging service and technology, while 6G aims to achieve stronger technical capabilities. Therefore, 5G-A can serve as an intermediary between 5G and the upcoming 6G technology. With the ongoing development of 5G-A and the upcoming 6G, it a good opportunity to consider the evolution of the Balong modem.

In both 5G-A and the upcoming 6G, modems will continue to play a crucial role in delivering a seamless communication experience for users across various new services. To achieve this goal, we need to simplify modem design by considering the constraints on modem implementation, including power consumption, cost, and compatibility. Considering new 5G-A services, potential applications, future technical requirements, and the experience gained from their

implementation and commercial use, this paper provides some thoughts on the evolution of Balong modem from the following aspects:

- Users' communication experience will shift from relative experience and absolute experience in high-value scenarios to deterministic experience. Deterministic networks are essential for delivering a deterministic experience and facilitating network deployment with new features and highlights.

- Green and low-carbon communication is fundamental for advancing cellular mobile communications. It is equally crucial to focus on modem energy saving alongside network energy saving. Therefore, it is necessary to address this requirement to prevent the recurrence of energy consumption issues encountered in 5G UEs during the initial phase of commercial use. The new system design should take into consideration and natively support the modem's low energy consumption at the initial stage.

- In the evolution of communication systems, every new generation is expected to possess stronger capabilities regarding various communication key performance indicators (KPIs), especially the peak data rate. The increase in peak data rate can be several times or even dozens of times greater than the previous generation, leading to significantly higher modem costs. However, UEs cannot perceive a noticeable improvement in their overall experience that justifies these increased costs. Therefore, it is important to note that for modems, the peak data rate should not align with the network's capabilities, and striking a balance between modem complexity and user experience is crucial.

- In the evolution of communication standards, incorporating new features requires modems to support different feature combinations and flexible parameter configurations. Consequently, numerous unnecessary or commercially impractical combinations are generated, which brings increasing challenges for modem compatibilities. Therefore, the standards should transition from being solely network-centric to achieving a balance between networks and UEs to guarantee flexibility and robust compatibility of modems.

- Trustworthiness is a cornerstone of the future information society. To improve E2E trustworthiness, the trustworthiness of modems should be considered in addition to the trustworthiness of core networks and base stations.

2 Deterministic Network and Deterministic Experience

2.1 Background

The Balong modem offers three phases of communication experiences, namely communication experiences 1.0, 2.0, and 3.0, as shown in Figure 1.

In communication experience 1.0, the Balong modem offers a better communication experience than competing products and achieves the best field failure rate (FFR) and third-party KPI evaluation results.

In communication experience 2.0, the Balong modem outperforms competing products and achieves flawless performance in high-value user scenarios. In this phase, there are some breakpoints of UE experience caused by the network like subways or high-speed railways. These problems cannot be fully resolved or eliminated through UE optimization. The comprehensive analysis of the entire network architecture that includes UEs is crucial to finding a final solution to these issues. Therefore, the next phase of the Balong modem aims to guarantee a deterministic experience. Deterministic networks are crucial for deterministic experiences. They enhance user experience and bring new insights into network construction. Traditionally, network construction has been focused on improving coverage. However, the average resource utilization of 5G networks is low. As a result, communications service providers (CSPs) are hesitant to invest more in network construction. By adopting an experience-centric approach to network construction, we can increase the users' willingness to pay more for high-quality services and maximize network value, leading to a mutually beneficial outcome for both CSPs and network vendors.

Figure 1 Overall roadmap for the evolution of Balong's communication experiences

2.2 Analysis of Deterministic Experience in Typical Scenarios

2.2.1 Deterministic Experience of Voice Services

Firstly, theoretical models are available for voice services. Agner Krarup Erlang, a Danish mathematician, developed formulas B and C [20] based on the assumption that random calls follow the Poisson distribution. Formula B provides service assurance for call lost systems, whereas formula C is designed for call delay systems. The Erlang B table (Table 1) is developed using Erlang models and formulas. It illustrates the relationship between the number of channels, user traffic, and call loss rate.

Secondly, the transmission resources for voice services are guaranteed. In the 2G and 3G eras, voice services are provided in the circuit switched (CS) domain. This domain allocates exclusive traffic resources to users, guaranteeing service quality. In the 4G and 5G eras, voice services are provided in the packet switched (PS) domain. In this domain, voice services are transmitted through a dedicated transmission channel with the highest scheduling priority. In addition, the data rate of voice services is not high, making it easy to guarantee resource availability.

Finally, the key quality indicators (KQIs) of voice experience are directly mapped to network KPIs in the network construction standards and measurement systems. This guarantees voice experience KQIs and communication KPIs.

2.2.2 Deterministic Experience of Electric Power Private Networks

An electric power network consists of two planes: the production plane and the management plane. The production plane imposes stringent requirements on deterministics and security. As a result, a private network is established to isolate it from the public network through resource block (RB) reservation, slicing, and exclusive user plane function (UPF). The management plane, although not subject to such strict requirements, still operates within the virtual private network. Therefore, resources are guaranteed, even though they do not necessarily need to be exclusively utilized.

Table 1 Erlang B table

Number of Channels	Call Loss Rate				Number of Channels	Call Loss Rate			
	1%	2%	3%	5%		1%	2%	3%	5%
1	0.01	0.02	0.03	0.05	11	5.16	5.84	6.33	7.08
2	0.15	0.22	0.28	0.38	12	5.88	6.61	7.14	7.95
3	0.46	0.60	0.71	0.90	13	6.61	7.40	7.97	8.83
4	0.87	1.09	1.26	1.52	14	7.35	8.20	8.80	9.73
5	1.36	1.66	1.88	2.22	15	8.11	9.01	9.65	10.6
6	1.91	2.28	2.54	2.96	16	8.88	9.83	10.5	11.5
7	2.50	2.94	3.25	3.74	17	9.65	10.7	11.4	12.5
8	3.13	3.63	3.99	4.54	18	10.4	11.5	12.2	13.4
9	3.78	4.34	4.75	5.37	19	11.2	12.3	13.1	14.3
10	4.46	5.08	5.53	6.22	20	12.0	13.2	14.0	15.2

Table 2 QoS requirements for different phases of an electric power private network

Scenario	Service	Bandwidth	Latency
Power distribution	Precise load control	48.1–1.13 Mbps	20–50 ms
	Differential protection and control for distribution networks	2 Mbps	< 10 ms
	Phasor measurement unit (PMU)	4 kbps	< 10 ms
	Load control based on user responses	1.12 Mbps	20 ms
	Automated power distribution (including telemetry, remote communication, and remote control)	19.2 kbps	50 ms
	Automated power distribution (including telemetry and remote communication)	19.2 kbps	2s
Power transmission and transformation	Substation status monitoring	≥ 2 Mbps	< 300 ms
	Robot/Drone-based inspection	≥ 2 Mbps	120 ms
Power consumption	Power consumption data collection	10 kbps	15 minutes

Furthermore, a power system typically comprises four phases: power distribution, transmission, transformation, and consumption. Each phase has specific quality of service (QoS) requirements on bandwidth and latency (see Table 2). The bandwidth and coverage requirements can be accurately calculated based on the QoS requirements, number of devices, number of reserved resources, and link budget. This ensures a deterministic experience in the electric power scenario.

2.2.3 Deterministic Experience of Flexible Manufacturing

Flexible manufacturing systems have a hierarchical architecture and model according to related standards and protocols. Within these systems, the lower layer has extremely high requirements for latency and reliability. In Huawei's BEM project, the company utilized Balong modem-based chip modules on the network to deliver URLLC services. These services offer a two-way latency of 4 ms and reliability of 99.999% between the programmable logic controllers (PLCs) and controlled devices. The service level agreement (SLA) objectives can be achieved due to two primary factors. Firstly, standard URLLC features and proprietary URLLC features, such as grant-free, mini-slot, complementary time division duplex (TDD), and XSO, are incorporated. Secondly, OS scheduling and hardware accelerators are optimized within the modules to reduce forwarding latency and minimize jitter for specific data packets. For flexible manufacturing, an SLA deterministic experience can be achieved through optimization.

2.3 Deterministic Requirements for 2C Network Data Services

2.3.1 Deterministic Requirements of Conventional Data Services

Table 3 lists the typical types of apps and categories of services used by mobile users. These services have different requirements for latency and bandwidth, and they do not come with the same level of stringent SLA assurance as voice or 2B services. These services cater to consumers, and not all factors have the same impact on user experience. User experience is evaluated based on quality of experience (QoE).

$$QoE = \sum_{i=1}^{N} (Mapping\ Score\ of\ KQI_i * weight_i\%)$$

$$KQI = Formula(KPI_1, KPI_2 \dots, KPI_n)$$

QoE is the weighted statistical value of KQIs, which are calculated based on a series of KPIs. The impact of each KPI on QoE is subject to the mapping and weight. This is why the deterministic requirements for the network vary depending on the types of conventional 2C services.

2.3.2 Increased Deterministic Requirements of Emerging Services

Emerging services are the main driving force of technology evolution. In the era of 5G-A, RTBC services impose significant demands on communication bandwidth and E2E latency. To meet these requirements within the 5G-A framework, it is necessary to enhance technologies such as layered QoS and frame integrity without changing the existing network architecture, transmission protocol, and radio access network (RAN) architecture of 5G. Additionally, the industry has reached a preliminary consensus on the typical emerging services for 6G. Specifically, 6G is expected to provide real-time immersive services that offer the ultimate user experience. These services include immersive XR, holographic communication, and digital twin [3, 6, 9, 14, 15]. Compared to RTBC services of the 5G-A era, these 6G services have much higher data rate, reliability, and E2E latency requirements. See the corresponding white papers [6, 14] for details about these new services.

To accommodate such new services, we need to enhance the processing capabilities of modems and utilize application processing (AP) for processing, rendering, and displaying received image information. These services involve a large amount of real-time computing and processing workload, which poses considerable challenges to the processing capabilities of UEs, including central processing units (CPUs) and graphical processing units (GPUs). They also have demanding power requirements to support high-speed communication signal processing and image processing and display on the UE side. The processing capabilities of UEs will not be enough to meet the requirements of new services due to the limitations in UE size, energy consumption, and cost, as shown in Figure 2. This figure assumes that the computing power will continue to increase by 30% annually, as witnessed in the past decade. The peak computing power

Table 3 Typical app types and service requirements

App Type	Service	Transmission Protocol	KPI	Latency (ms)	Throughput (Mbps)
Short video	Short video playing	TCP	First screen latency/Frame freeze rate	300–500	10
	Live streaming - online streamer	TCP/UDP	Frame freeze rate/Frame rate/Bit rate	200–500	5–20
	Live streaming - audience	TCP/UDP	First screen latency/Freeze rate/Bit rate	200–500	5–20
Instant messaging	Video call	UDP	Freeze rate/Frame rate/Bit rate	100–300	2–5
	Voice call	UDP	Freeze rate	100–300	0.2
	Conversation using texts and images	HTTP	Loopback latency	500+	1
	Voice conversation	HTTP	Loopback latency	500+	0.2
	Transfer and reception of large files	HTTP		500+	60
	Mobile payment	HTTP	Loopback latency	200–500	1
Online video	Playback of video segments	HLS/DASH	First screen latency/Freeze rate/Bit rate	300–500	50
Mobile gaming	Realtime battle games	UDP	Freeze rate/Latency	100–300	0.2
	Non-realtime games	HTTP	Loopback latency	500+	0.2
News	Image and text download	HTTP	Loopback latency	500+	1
	Playback of video segments	HLS/DASH	First screen latency/Freeze rate/Bit rate	300–500	30
Online shopping	Image and text download	HTTP	Loopback latency	500+	1
	Short video playing	TCP	First screen latency/Frame freeze rate	300–500	30
	Live streaming - online streamer	UDP	Freeze rate/Frame rate/Bit rate	200–500	5–20
	Live streaming - audience	UDP	First screen latency/Freeze rate/Bit rate	200–500	5–20
Online reading	Image and text download	HTTP	Loopback latency	500+	1
	Playback of voice segments	UDP	Loopback latency	500+	0.2
Search and download	Image and text download	HTTP	Loopback latency	500+	1
	Large file download	HTTP	Loopback latency	500+	60

of UEs is projected to reach 24.7 teraFLOPS (TFLOPS) by 2028. However, even with such improvement in computing power, the computing power still falls short of meeting the requirements of deep or fully immersive XR experiences.

To better support future services, there is a growing trend to move computing tasks (such as rendering) from UEs to the cloud or edge. In this scenario, latency-sensitive tasks are performed by AP on the UE side, while modems handle data transmission between AP and the cloud. This enables the convergence of communications and computing, achieving device-cloud synergy.

New services that leverage device-cloud synergy demand a more deterministic experience. For instance, cloud-based photography necessitates shorter transmission durations and enhanced reliability, distributed cloud gaming requires lower latency and improved reliability, and immersive XR experiences call for reduced latency and increased bandwidth. They also require significant improvements in the capabilities of communication pipes to perform both downlink and uplink transmission within the E2E latency budget. There is still a need for extensive analysis and research on allocating computing capabilities between UEs and the cloud and establishing communication capabilities between UEs and the network while considering the power requirements of the UEs to achieve these objectives.
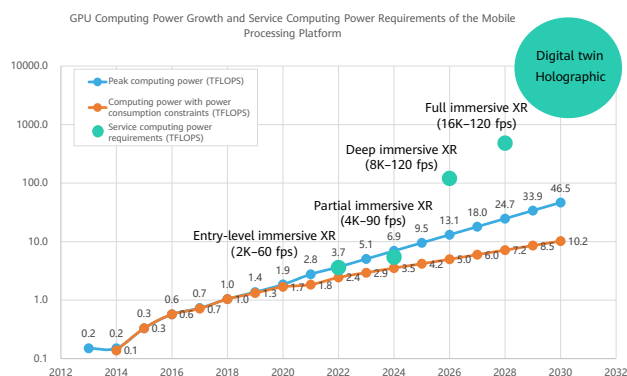
Figure 2 Projected computing capability growth of mobile processing platforms and the service computing capability requirements

2.4 Status Quo of 2C Data Services and Challenges of Deterministic Experience

2.4.1 Status Quo of 2C Data Services

In the realm of 2C data services, the focus of network construction is transitioning from being based on coverage to being driven by user experience. This experience-based approach involves constructing and accepting networks that meet the user experience requirements and standards of each service type. Given the diverse types of 2C data services and their varying QoS requirements, adopting experience-based network construction is challenging. Consequently, resorting to coverage-based network construction becomes necessary. Furthermore, there is a lack of service model support, meaning the network functions as a pipe without knowledge of specific services. This creates a large gap between the network KPIs and experience KQIs. Consequently, 2C networks often rely on a best-effort mechanism to deliver data services.

2.4.2 Challenges of Achieving Deterministic Experience

- **Coverage**

  Coverage is crucial for ensuring a good communication experience. If the coverage is poor, it is impossible to deliver a deterministic experience. However, achieving high-quality coverage in any region or location is challenging due to constraints on the network construction standards and engineering implementation costs.

- **Capacity**

  Coverage is not the primary issue in typical scenarios like high-speed railways and subways. The limited capacity poses the main challenge. Subways and high-speed railways typically experience heavy communications traffic that arrives simultaneously but does not last long in each cell. For 4G private networks on high-speed railways, the instantaneous physical resource block (PRB) usage of cells is almost 100% when a train passes. In comparison, 5G private networks provide larger bandwidth. However, this only alleviates network congestion, but does not address resource congestion issues. During peak hours, the test results show that the number of frame-freezing instances in short videos remains high, negatively affecting user experience. This problem worsens in latency-sensitive services like WeChat-based voice calls and live streaming. Theoretically, this problem can be solved by deploying more base stations or expanding the spectrum, but the cost is substantial.

- **Power consumption**

  Power consumption is another key factor in providing deterministic experience. For instance, when a UE is

used for cloud-based photography, the power consumed during uplink data transmission accounts for the highest portion of its total power consumption. The amount of power consumed is directly proportional to the transmission duration. The uplink transmission duration is subject to various factors, such as the signal strength of the base station at the location of the UE, signal quality, the number of users in the cell, and even the way the UE is held. Due to the uncertainty of uplink signal quality, there is considerable variation in uplink transmission duration, resulting in uncertain power consumption.

- **Jitter**

  In the 5G network architecture, the base station, core network, and the transmission network between them, and connections between the core network and the Internet are connected via cables. The bandwidth is large enough to enable deterministic transmission. However, the air interface between a UE and a base station is an open interface, making it susceptible to various factors and resulting in extreme QoS fluctuations. Figure 3 shows the actual uplink data rate of the 5G network. It can be observed that each test, despite being conducted in the same location and for the same time period, yields a significant fluctuation in the actual uplink data rate. Numerous variables affect the jitter over the air interface. Some variables result from the characteristics of radio signal transmission, and some variables are due to the design limitations of the product. Eliminating these jitters remains a significant challenge. However, in deterministic networks, it is necessary to contain jitters within a small range.
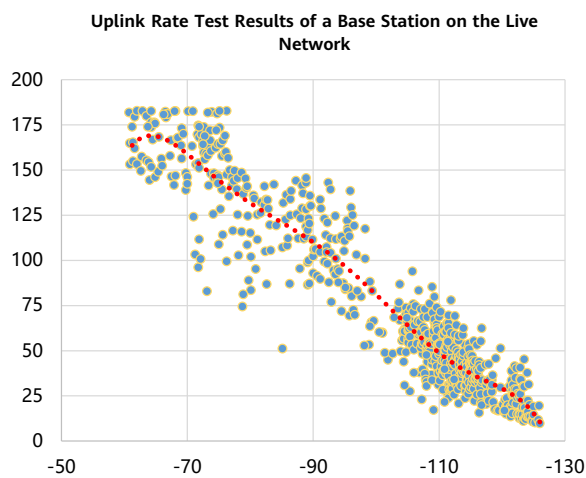
2.5 Technical Directions of Deterministic Network and Experience

2.5.1 Service Awareness

Service awareness is essential for ensuring a deterministic experience tailored to meet specific service and QoS requirements. In the case of MeeTime services and device-cloud services, collaboration is performed between AP and communication processing (CP) to enhance user experience through mutual awareness. In the context of 5G-A XR services defined in 3GPP Release 18, E2E service awareness is performed. For example, network elements (NEs) detect I-frames and P-frames and implement scheduling based on the frame types. Service awareness and service-based QoS guarantee are crucial in delivering a deterministic experience.

In real-world networks, different types of data services have distinct data characteristics, as shown in Figure 4. The use of communication pipes by apps is relatively primitive as they lack service awareness of each other, making it extremely challenging to guarantee QoS. Therefore, it is essential to restrict the usage of communication pipes by apps, and it becomes necessary to define data characteristics according to service types, service data, and networks. Based on
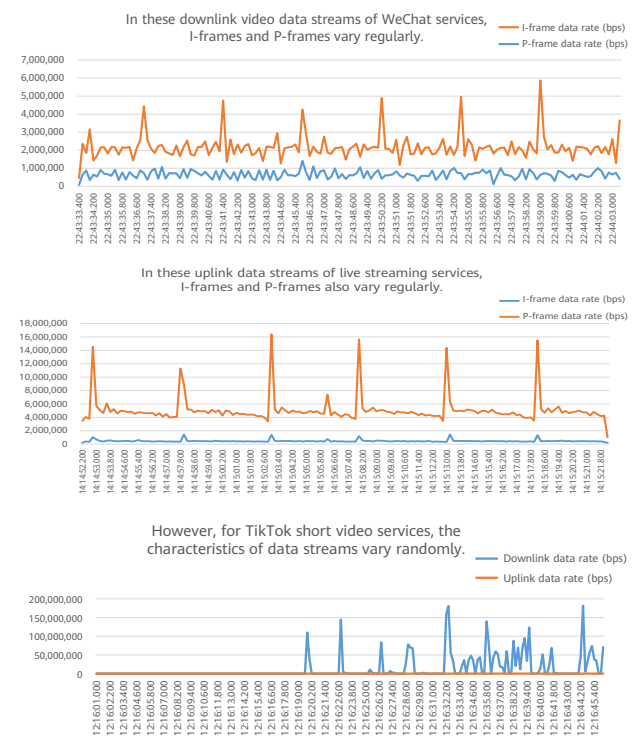

Figure 3 Actual uplink rate of a 5G site

Figure 4 Example of data characteristics of various data services

these characteristics, apps transmit data, and the network performs scheduling to ensure QoS. Moreover, power consumption can be reduced on both the network and UE sides.

In the device-cloud synergy system shown in Figure 5, apps and servers must handle user input in real time while rendering and encoding images. In certain situations, the processing time may exceed the E2E transmission time. The latency experienced by users encompasses both communication latency and the processing latency of apps and servers. In conventional implementation, data transmission in communications networks and data processing in apps are performed sequentially. As a result, it is difficult to effectively reduce the latency perceived by users only by conducting latency optimization in communications systems.

On the user side, there is a latency when apps process user input. Once the information is processed, the apps send it to the modem. The modem then requests air interface resources from the network. On the cloud side, the server must render and send images to the core network, which transmits them to the base station. The base station allocates air interface resources and then transmits the received data. Considering available timeslot resources (e.g., in TDD systems) and the overall resource scheduling status of the base station, an extra waiting time is introduced. Therefore, if data processing in apps and resource allocation in the network can be performed concurrently, the E2E user-perceived latency can be significantly reduced.

With this approach, the UE promptly notifies the modem of any user input it receives based on service characteristics. This notification triggers a procedure of modems to apply for resources and obtain wireless bandwidths. Once the apps on the UE side handle the data packet, a proper startup window is established. Consequently, when the data packet is transmitted to the modem, uplink resources are ready, and the data package can be directly transmitted, thereby minimizing the waiting time for resource allocation. When the apps and modem on the UE side are integrated,
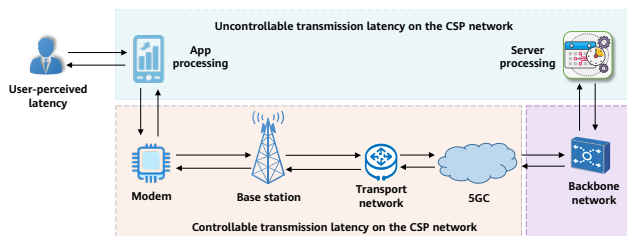
more in-depth interaction can be performed between them. For the downlink, the waiting time can also be reduced if the base station can acquire the information about the data arrival in advance. Therefore, in-depth interaction and awareness between the apps or server and the communications network can significantly reduce the E2E transmission latency, improving user experience.

2.5.2 QoS Scheduling

Slicing is a crucial method for achieving a deterministic experience and has already been implemented for certain services on live networks. However, it comes with high costs. To address this challenge, the existing QoS scheduling mechanism of the 5G network can be leveraged. QoS management and scheduling are core functions of all communications networks, including 5G networks. A mature and comprehensive QoS mechanism on 5G networks has been established, offering a wide range of QoS parameters for refined resource management.

In 5G, QoS management is implemented at the level of QoS flows, as depicted in Figure 6. Compared to 4G evolved packet system (EPS) bearers, 5G QoS management is more precise and yields superior results. The 5G QoS management enhancement focuses on mapping and management from IP flows to QoS flows and from QoS flows to data radio bearers (DRBs) over the air interface. Only carrier-class voice over NR (VoNR) and video over NR (ViNR) services on the current 5G network utilize high-priority QoS. All other data services rely on the default bearer. Over the top (OTT) data services use the network only as a pipe, limiting their ability to fully leverage the QoS capabilities of 5G networks.

In network implementation, QoS guarantee is usually accomplished through static splitting over the involved network domains. This approach decouples the association between subsystems, facilitating system design. However, to guarantee a deterministic experience, higher QoS requirements for subsystems are often posed, compromising
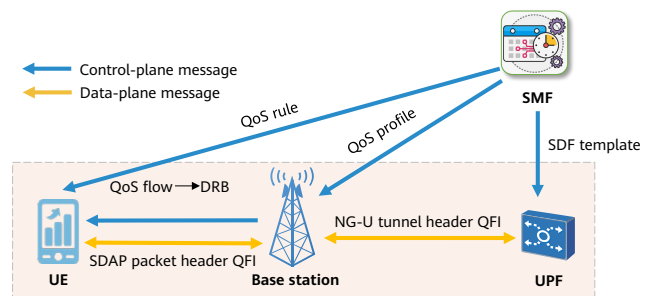

Figure 5 User-perceived latency and network transmission latency

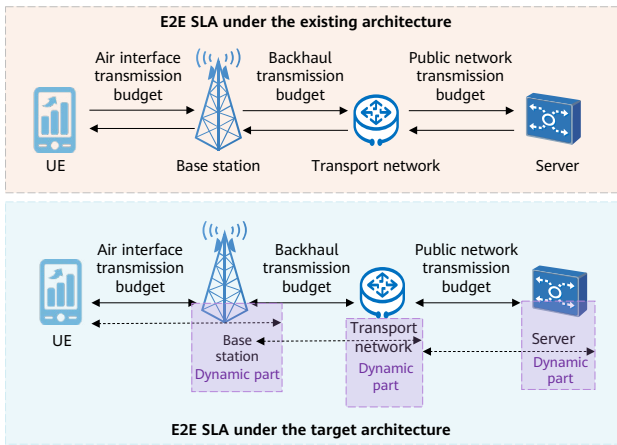**Figure 6** 5G QoS management mechanism

Figure 7 Dynamic QoS breakdown for E2E SLA assurance

efficiency. When a subsystem's QoS guarantee fails, it results in an overall failure of the E2E QoS guarantee.

To improve SLA assurance capability, the dynamic QoS decomposition approach can be utilized to break down the E2E SLAs into subsystems. This approach ensures that each subsystem's budget fluctuates within a specific range. Different subsystems can cooperate to achieve E2E SLA assurance. The data transmission of other systems can be sped up to compensate for the transmission latency of individual subsystems, thereby reducing the overall SLA risk to the system. This places greater demands on the existing QoS framework. The QoS of data packets is adjusted according to the transmission status of each subsystem. Consequently, it is necessary to reevaluate the dynamic QoS framework, as shown in Figure 7.

2.5.3 Open Network Interfaces

To fully utilize the network's QoS capability, network interfaces must be selectively opened up. There have been several use cases of opening up network interfaces. Based on the actual results, interfaces related to voice and short message service (SMS) services are most frequently called in either cloud vendor mode or CSP self-operated mode. Voice and SMS services are simple, and it is easy to encapsulate and invoke interfaces for such services. As a result, their quality is guaranteed on the CSP network. On the other hand, data services require integration into the communication technology (CT) network, which adds complexity. To reduce the complexity, we must address the challenges of standardizing and differentiating interfaces while ensuring their diversity.

At the Mobile World Congress 2023 (MWC 2023), the Open Gateway has emerged as a hot topic. It is a framework consisting of universal network Application Programmable Interfaces (APIs). Leading European CSPs and equipment vendors have expressed their desire to open up more network capabilities, unlock more potential, and further facilitate value monetization. Therefore, opening up network capabilities is an important direction in the future.

As shown in Figure 8, the Open Gateway network architecture opens up interfaces across multiple layers. Currently, the Global System for Mobile Communications Association (GSMA) has opened up eight interfaces, with plans to open up more interfaces in 2023. To drive the project forward, the industry has launched the CAMARA
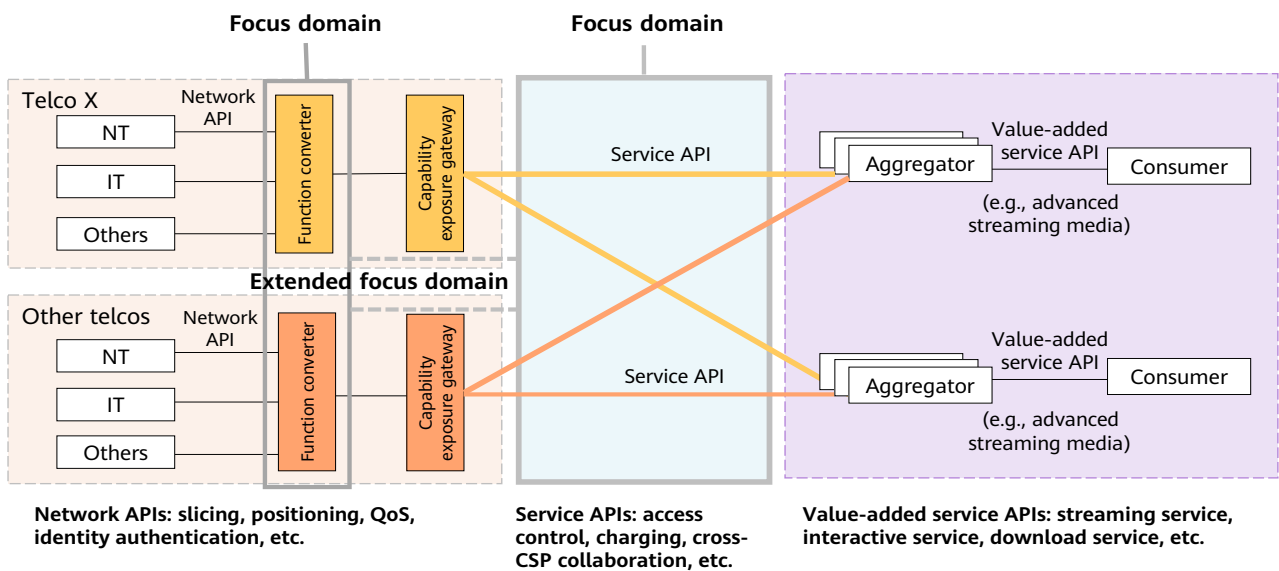

Figure 8 Open Gateway network architecture

initiative. Huawei is among the participants, along with 63 other companies. GSMA and the Linux Foundation have collaborated on an open-source project to define, develop, and release interfaces.

2.5.4 UE-centric Experience Measurement System

Measurement systems are essential for developing network construction standards, acceptance criteria, and evaluation systems. Therefore, it is necessary to set up measurement systems for deterministic experience to drive the evolution towards deterministic networks. A UE-centric experience measurement system consists of three levels, which must be implemented phase by phase.

1. Generate coverage-level grid maps based on call history records (CHRs), crowdsourcing data, and crowdtesting data. This level is also known as KPI-level measurement.

2. Generate grid maps based on KQIs on which service experience depends.

3. Generate grid maps based on QoE data.

2.5.5 Deterministic Network Architecture

The deterministic network architecture shown in Figure 9 is a visionary architecture focused on enhancing the user experience. When starting a service, the app sends QoS requirements to the scheduler and obtains spatiotemporal SLA prior data based on the SLA experience map. In this way, UEs determine whether the QoS can be guaranteed. If the QoS cannot be guaranteed, UEs determine whether to accept the QoS of a lower level. The procedure is as follows:

- The scheduler calls the API to send QoS requests to CSPs' OP domain.

- The OP domain parses the received requests into 3GPP parameters and forwards the parameters to the network exposure function (NEF), charging function (CHF), and multi-access edge computing (MEC) on the core network.

- The core network configures UE parameters over the user-network interface (UNI) and establishes 5G QoS identifier (5QI) flows as required.

- The OTT vendor pushes the service data to CSPs' MEC (optional), and the app starts the data service.

- The UE modem senses and measures the SLA service quality during the entire process.

3 Modem Power Consumption

3.1 Importance of Modem with Low Power Consumption

In order to achieve the goals of carbon peaking and carbon neutrality, it is necessary for all industries to reduce their carbon emissions, improve energy efficiency, and strive for sustainable environmental development. The pursuit of these goals is expected to add impetus to energy saving efforts within the mobile communications industry. Currently, the industry has identified energy saving as a
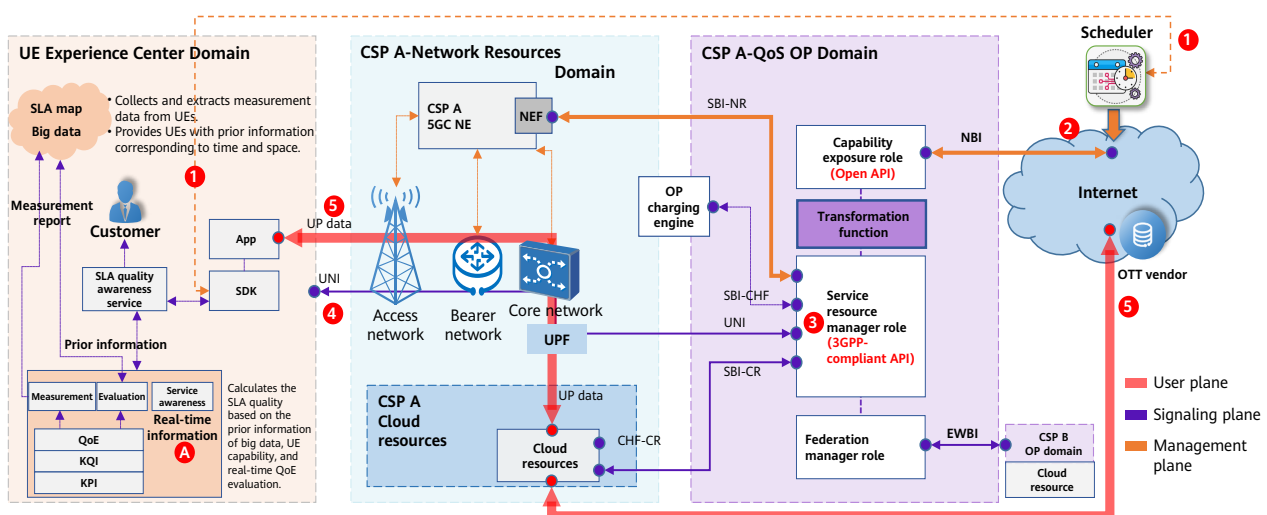

Figure 9 Deterministic network architecture

fundamental objective and distinguishing characteristic of 6G [6, 7, 9, 10, 12, 15]. In the mobile communications industry, CSPs place great emphasis on network energy saving due to its direct impact on their operating expense (OPEX). This is why CSPs have such strong requirements for reducing energy consumption of 5G networks [7, 9, 10]. However, it is just as important to address energy consumption of UEs, as their vast numbers mean their cumulative impact is substantial [7, 18]. In addition, the energy saving capabilities or standby time of UEs is also a key factor that significantly affects user experience.

As a core and fundamental capability of UEs, communication modems play a crucial role in running most services on these UEs. The communication energy consumption of these UEs, known as modem power consumption, directly affects user experience and serves as a key metric for evaluating the communication capabilities of UEs [2]. Firstly, the power consumption of a modem depends on the services it is running. For instance, a service that requires a higher communication data rate and/or lower latency will invariably increase power consumption due to its stricter processing time budget and extensive operational tasks. As the network evolves from 5G to 5G-A and on to 6G, we see the emergence of an increasing number of application scenarios and RTBC services, such as immersive XR. This poses stringent requirements on communication processing capabilities and presents challenges in modem power consumption. Secondly, the power consumption of a modem for specific services depends on its working mode, including factors such as operating bandwidth, number of receive antennas, and transmit antenna. For example, in the downlink transmission, a single NR carrier is usually configured with 100 MHz bandwidth and four receive antennas, whereas a single LTE carrier is typically configured with 20 MHz bandwidth and two receive antennas. Given the same service within the transmission capabilities of both NR and LTE and the same days-of-use (DOU) model, the power consumption of a single NR carrier is approximately 40% higher than that of a single LTE carrier. This is one reason why 5G UEs struggle with communication power consumption. Over the past few decades, each generation of cellular mobile communications has seen at least ten times better communication capabilities for UEs (such as peak data rate and user-perceived data rate) over the previous generation. If cellular mobile communications technology continues to evolve at such a rate, the continuous increase in operating bandwidth and number of antennas supported

by a UE will inevitably present significant challenges in modem power saving. Finally, modem power consumption is also subject to standards and protocol design. If low modem power consumption is not thoroughly considered and this goal lacks native protocol support, it will be difficult to give the full play to low modem power consumption features in the actual network, or the networks that have this feature enabled will be scarce. In this case, the features will have a limited impact. For example, the power consumption resulting from the existing NR physical downlink control channel (PDCCH) mechanism's invalid PDCCH blind detection accounts for 60% to 70% of total power consumption of a modem in connected mode. The PDCCH search space switching and PDCCH skipping mechanism, as defined in 3GPP Releases 16 and 17 for NR, can be utilized to reduce the power consumption associated with PDCCH-only detection. However, it is important to note that these features are optionally supported according to 3GPP specifications. Given this, it may take much time and efforts for the industry to establish a consensus on the commercial implementation of these non-native power-saving features in networks and UEs.

Moore's Law is gradually failing, with the power saving benefits provided by advanced processes constantly diminishing. As such, it will be difficult to fulfill future requirements for low modem power consumption through semiconductor process evolution. According to data provided in [17], in theory, each new generation of processes can reduce the modem power consumption by 20% to 30%. However, in practice, this reduction is significantly lower. In addition, various techniques may be utilized in different phases of the entire modem production process. This further undermines the actual power consumption benefit achieved by each generation of technologies. Consequently, communication modem across the industry will encounter common challenges in the realm of power saving as we move forward.

Furthermore, UEs encounter certain challenges in terms of battery capacity expansion and heat dissipation control, primarily due to their limited size and manufacturing processes.

To sum up, the low power consumption of modem should be listed as a key objective of the forthcoming 6G system design.

3.2 Balance Between Low Modem Power Consumption and Network KPIs

In the initial stages of designing the 5G basic version in 3GPP Release 15, low modem power consumption was not deemed a core objective. Consequently, the power consumption of 5G UEs compliant with 3GPP Release 15 was considerably higher than that of LTE UEs. The Balong modem team was dedicated to reducing power consumption through various methods, including optimizing architecture, enhancing software performance, and coordinating software and hardware components. To address the power consumption problem, a series of discussions was held during the development of 3GPP Releases 16 and 17. Eventually, various low UE power consumption features were introduced for 5G, including UE assistance information (UAI), wake-up signaling (WUS), paging early indication (PEI), PDCCH search space switching, and PDCCH skipping. That said, since the features introduced for NR after 3GPP Release 15 are essentially enhancements, upgrading both networks and UEs to 3GPP Releases 16 and 17 takes a long time. It is impossible to quickly upgrade all networks operated by a single CSP, let alone those of different CSPs. As a result, after the power saving features are specified in later 3GPP releases, reducing UE power consumption within a short period still is not feasible. Considering the significance of low modem power consumption, it is crucial to fully consider low UE power consumption at the very beginning of standards development in order to ensure its native support.

In addition, the UE power saving features introduced after 3GPP Release 15 pose a challenge in striking a balance between network KPIs and low UE power consumption in real-world applications. These features require coordination between the network and UEs to save UE power. However, some of them may impact network KPIs or fail to provide direct power-saving benefits for the network. As a result, these features are either not chosen for commercial use or not enabled, in turn resulting in high UE power consumption. Network energy saving features, such as those introduced in 3GPP Release 18, achieve energy savings by dynamically shutting down certain channels or signals on the network side. However, this approach adversely affects UE implementation by increasing processing complexity or power consumption. Consequently, these network energy saving features will face challenges and uncertainties in commercial use.

In the future, network and UE energy saving will be equally important. When developing standards, it is necessary to design unified KPIs for both network and UE energy saving to make them commercially viable. Furthermore, it is crucial to design network KPIs that truly reflect user experience.

3.3 Power-Efficient Modem Design

In the current communication framework, a modem primarily consumes power during invalid PDCCH detection in connected mode and paging detection in idle mode. This portion of power consumption does not contribute to services and so should be minimized whenever possible. In connected mode, the network typically does not notify UEs of the PDCCH sending occasion due to uncertainty in service arrival and complexity in network scheduling. As a result, UEs are required to perform blind PDCCH detection continuously, even if there is no service arrival or scheduling. Based on the DOU test data, the power consumption of PDCCH-only detection accounts for 60% to 70% of the total power consumption of a modem in connected mode. In idle mode, the modem needs to periodically wake up to detect paging information. Before detecting paging, the modem also needs to perform time and frequency synchronization. In LTE, each subframe carries common reference signals, whereas in NR, UEs in idle mode obtain one or multiple synchronization signal block (SSB) bursts based on periodically transmitted SSBs prior to a paging occasion for time and frequency synchronization. If no paging is performed for a UE, the invalid detection on paging increases power consumption.

Therefore, the modem should strive to achieve "zero bits and zero watts."

PDCCH Design for Low Power Consumption

As mentioned above, blind PDCCH detection when there is no service consumes a significant amount of modem power. This is one of the obstacles in achieving low modem power consumption. Therefore, it becomes essential to redesign the PDCCH mechanism. In both LTE and NR, PDCCH design prioritizes simplicity and flexibility of network scheduling. Normally, only the scheduling information for the current timeslot is sent to UEs in both systems. As a result, UEs continuously perform PDCCH blind detection and are unable to enter dormancy mode, leading to increased UE power consumption. If the network can notify UEs in advance about scheduling information, UEs can perform blind PDCCH detection only in the potential scheduling timeslots, thus avoiding numerous invalid PDCCH detections. However,

this approach requires the network to know UEs' service characteristics. By leveraging the service awareness function described in Section 2.5.1, service characteristics can be established for different types of services. With these service characteristics, apps transmit data and the network performs scheduling. In this way, the power consumption of both the network and UEs can be reduced. In conclusion, future designs should prioritize the development of a simplified PDCCH detection mechanism to greatly reduce unnecessary PDCCH detections.

Balance Between UE and Network Energy Saving

The network learns the service status of UEs based on UE assistance information, and either temporarily disables radio frequency (RF) circuits or reduces the number of enabled channels accordingly. This helps reduce network energy consumption while also putting all UEs in low energy consumption mode. Additionally, the network aggregates and schedules data packets based on UE assistance information such as service types and characteristics. It then notifies UEs about opportunities to detect the PDCCH, resulting in energy savings for both the network and UEs.

New Energy Acquisition Mode

In idle mode, a UE is paged only on a few paging occasions. If the UE is not paged, it means that the UE has no service, and therefore the UE should not consume energy during this period. In this scenario, we can implement a new energy harvesting technology that allows the network to transmit energy alongside regular signals. The UE can then harvest this energy to perform paging detection and confirmation, thereby achieving the goal of "zero bits and zero watts." This involves the new design of paging signals, including energy conversion efficiency and low energy consumption paging detection.

AI-assisted Power Saving

As a fundamental enabling technology, AI is expected to be natively supported by 6G systems. AI can be utilized for channel state information (CSI) measurement, beam management, positioning, and other purposes. These application scenarios are currently the subject of study in 3GPP Release 18. Furthermore, AI can aid in UE power saving, assisting UEs in predicting the arrival time of PDCCH by leveraging information such as service models, channel quality, and base stations' scheduling habits. This allows UEs to detect PDCCH at the predicted occasion, saving the power consumption that would be consumed for blind PDCCH detection.

4 Modem Complexity

4.1 Modem Complexity Analysis

As cellular mobile communications technologies evolve, the demand for stronger processing capabilities on modems continues to grow. This, in turn, leads to modems becoming more complex and therefore the cost increases. Although the benefits from evolving semiconductor technology can compensate for the increasing complexity and costs of modems, such benefits will diminish as the room for improvements becomes smaller [17]. Furthermore, new RTBC services require higher UE power and processing capabilities. Simplified modem designs should therefore be considered a crucial factor for formulating standards in the future [8]. It is necessary to avoid designs that significantly increase implementation complexity without improving user experience and instead focus on simplified modems that can deliver an optimal user experience and meet user requirements with a reasonable level of implementation complexity.

Peak data rate, one of the most important indicators in every generation of cellular mobile communication systems, refers to the maximum data rate (in bit/s) [19] that can be achieved under ideal channel conditions. The operating bandwidth, number of MIMO layers, and modulation scheme all affect the peak data rate, which ultimately determines the communications specifications and cost of a modem. In LTE, UEs do not have to reach the same peak rate as the network. There is a clear roadmap for LTE UE evolution, as more than 20 LTE UE categories are defined to adapt to a wide range of services and reduce UE costs, facilitating the development of different services. Unlike LTE UEs, NR UEs defined in 3GPP Release 15 for NR need to have the same peak rate as the network. For example, FR1 UEs need to support the carrier bandwidth of 100 MHz and four downlink layers. This means that certain NR UEs will have stronger capabilities than needed. Consequently, those UEs are more expensive and consume more power, making it harder to adapt to different services and hampering service development. These problems are what prompted discussions on reducing UE complexity in later 5G NR standards by introducing new technologies, such as 20 MHz & 1 Rx/2 Rx (RedCap), 50 MHz & 2 Rx, and 2 Rx XR UEs.

A number of companies have published 6G white papers detailing their objective of reaching a peak rate of 100 Gbit/s or even 1 Tbit/s [6, 12–14]. Achieving the same

level of peak data rate on both UEs and the network will increase both UE complexity and cost to an unacceptable level, making large-scale commercial use of UEs impossible. Regular users and even users in commercial scenarios barely require such a high peak data rate, meaning that a significant amount of resources would be wasted if modems with such a high peak data rate were developed. Peak data rate should reflect the upper limit of network capabilities that can be shared with different UEs rather than UEs being expected to have the same peak data rate as the network.

Standard design — another crucial factor for implementing a simplified modem — is used to support interoperations by clearly defining the interfaces between the network and UEs. A simplified interface can make interoperations easier and avoid UE compatibility problems. Conversely, while a complex interface can provide many options, it does not provide a unified solution for every basic function. These options make standards more flexible in meeting the requirements of various companies, however, they also introduce many potential combinations at the system level. The network can select some of these combinations based on CSPs' requirements or deployment conditions — an action impossible for modems. To function normally on networks operated by different CSPs or deployed using devices provided by different vendors, a modem must support all these potential combinations. However, many of these combinations may never be used on the network, existing only in standards. For example, among the 56 timeslot structures defined in NR standards, only two or three of them are widely used in commercial 5G systems. The need to support all the combinations significantly increases the implementation complexity and cost of UEs, leading to many potential issues in terms of UE compatibility (discussed in Section 5). UE Interoperability and Development Testing (IODT) and conformance tests alone cannot cover all compatibility issues, and these tests increase the UE test costs. In addition, as the large number of potential combinations increases, the commercial use of features will slow down because greater effort is required to promote an industry consensus. In future standards, the designs of the UE-network interfaces should be simplified as far as possible from the very beginning to reduce unnecessary options, because flexible configuration is not always a good choice.

The utilization of spectrum also needs to be considered during modem evolution. UEs need to support roaming on the spectrum in different regions and provided by different CSPs. However, given the diverse range of fragmented spectrum, the many use modes of spectrum, and the complex frequency band combinations that are possible, there are many daunting challenges involved in the implementation of the RF front end. Multiple filters are required to support the numerous frequency band combinations, significantly increasing the cost and size of a modem's RF front end while also incurring to performance deterioration due to insertion losses.

4.2 Flexible Utilization of Spectrum

4.2.1 Challenges Faced by Spectrum Usage Technologies

CA

Carrier aggregation (CA) is an important spectrum utilization method for both LTE and NR. To support CA, UEs need to simultaneously transmit or receive data on multiple carriers. However, the number and bandwidth of CA carriers affect the UE cost. Despite its usage on the live network, CA has created many challenges to the network and UEs, especially to the uplink capabilities of UEs. Emerging services, such as live broadcasting, industrial video surveillance, and device-cloud synergy, pose higher requirements on the uplink capabilities of the network.

First, CA is not flexible enough in the uplink because it cannot fully utilize uplink frequencies. Due to cost, power consumption, heat dissipation, and size factors, a UE can simultaneously work on only a few uplink RF channels — usually a UE can aggregate only two carriers in the uplink. When a CSP offers more than two frequency bands, CA cannot fully utilize these frequency bands in the uplink. Furthermore, CA requires uplink carriers to be associated with downlink carriers and cannot sufficiently adapt to services that have different requirements on uplink and downlink CA capabilities.

Second, CA has low power efficiency in the downlink — power is wasted when no data is transmitted. In CA, a UE needs to perform time and frequency synchronization, control channel monitoring, mobility measurement, and beam management (FR2) on each carrier independently. Consequently, enabling more carriers in the downlink will lead to lower power efficiency.

SUL

Supplementary uplink (SUL) is a new spectrum usage

feature introduced in 3GPP Release 15 to solve the problem of limited uplink coverage on typical 5G TDD frequency bands.

SUL has been continuously evolving to improve uplink capacity. Advanced versions of SUL include uplink carrier switching between two frequency bands in 3GPP Releases 16 and 17, and uplink multi-carrier switching between three or four frequency bands in 3GPP Release 18. To improve uplink performance and user experience in multi-band scenarios, uplink carrier switching technology dynamically switches the limited uplink power and channels of a UE to the optimal frequency bands or carriers for transmission. SUL and its advanced versions have become one of the key technologies of 5G and 5G-A networks, supporting a diverse range of new services that require large uplink bandwidth.

Although SUL supports flexible switching between three or four frequency bands, it still uses CA in the downlink. Like CA, it also requires uplink carriers to be associated with downlink carriers. Consequently, a UE that supports uplink multi-carrier switching must support two or three downlink CA frequency bands (including one SUL frequency band, or three or four downlink CA frequency bands). This limitation poses a high requirement on the downlink capability of a modem and significantly increases the modem cost.

4.2.2 Evolution of Multi-Band Access Technology

Each generation of cellular mobile communications technology introduces new frequency bands while also refarming some existing ones, and most CSPs have multiple frequency bands. The use of these frequency bands needs to be designed from the perspectives of network capacity, UE experience, power consumption, and complexity. Based on the analysis in Section 4.2.1, multi-band access technology should accommodate the following characteristics:

- First, uplink and downlink multi-carrier capabilities are decoupled. UEs that support multi-band switching in the uplink do not have to support downlink CA on the corresponding frequency bands. Improving the uplink capabilities of such UEs does not increase the cost of the downlink.

  In addition, uplink and downlink carrier configurations should not be limited by the capability of uplink and downlink carriers supported by UEs. Generally, the network supports more frequency bands and carriers than UEs do. As discussed in Section 4.1, due to modem

complexity and usability, UEs do not have to support the same carriers as the network does. This means that although all carriers supported by the network can be configured on UEs, those UEs do not need to support simultaneous reception or transmission on all these carriers. In summary, more carriers are configured on the network than needed, and UEs work in a carrier selection mode. Similar to uplink carrier selection, downlink carrier selection can be adopted in the future to reduce downlink complexity of UEs and reduce network energy consumption.

- Second, the definition of cells is decoupled from carriers to reduce network signaling overheads and processing power consumption of UEs. Currently, each cell is associated with one carrier. When a UE works in multi-carrier mode, it needs to monitor and measure multiple carriers, resulting in high power consumption. If a cell is decoupled from carriers, a single cell can offer multiple carriers (on the same or different frequency bands). One carrier (the anchor carrier) is used for access, time-frequency synchronization, idle mode listening, control channel listening, and mobility measurement, while other carriers (non-anchor carriers) are enabled only when data transmission is required. To decouple cells from carriers, the network needs to ensure that the carriers associated with a cell can be used to perform precise time-frequency synchronization. This allows UEs to determine the key parameters of non-anchor carriers by using information about the anchor carrier. From the perspective of L2 and L3 protocol stacks, decoupling cells from carriers can simplify the protocol stacks, thereby reducing the power consumption of UE measurement and control channel monitoring. Another benefit of carrier selection is mobility. UEs that support carrier switching are naturally capable of measuring neighboring cells or frequencies in advance. Because L2 and L3 can be decoupled from the physical layer, UE handovers between cells can evolve toward an information update and mapping to the physical layer of the new cell, without changing the L2 and L3 protocol stacks. This facilitates seamless handovers.

- Third, low and medium frequency bands (sub-3 GHz) and high-frequency bands (sub-6 GHz, mmWave) are mainly used for uplink and downlink transmission respectively. Unlike base stations, UEs are always restricted by size, battery capacity, power consumption, and heat dissipation. Increasing the antenna aperture and transmission power on UEs is not a viable option for reducing the path loss caused by high-band signals

of UEs and improving uplink capacity. Because low frequency bands have low path losses, wide coverage, and high power efficiency, they are more suitable for uplink transmission. In addition, the large spacing between sub-3 GHz and sub-6 GHz can facilitate duplex isolation, reduce transmit and receive interference within UEs, and increase uplink transmit power more easily.

However, some issues need to be solved before low and medium frequency bands are used for uplink transmission. Specifically, most low and medium frequency bands are scattered across a large span, making it difficult for a single RF component — such as a power amplifier (PA) or an antenna — to support these bands. Furthermore, the use of such bands is regulated for downlink transmission in mobile communications systems. Introducing these bands in the uplink requires new regulations and enabling technologies, such as traffic sensing-based downlink and uplink conversion.

In conclusion, the evolving multi-band access technology will significantly improve uplink performance and communications experience, and reduce the power consumption and cost of UEs.

4.2.3 High- and Low-Band Coordination

5G has defined new frequency bands FR1 and FR2, and the former has already been put into large-scale commercial use. China has deployed numerous FR1 base stations and is preparing for the commercial use of FR2 mmWave. The biggest advantage of mmWave over FR1 is large-bandwidth spectrum resources. However, due to the higher frequency and large path losses of mmWave, the transmission of mmWave signals is subjected to various environment factors (e.g., leaf blocking and penetration losses), making it difficult to deploy contiguous coverage by using mmWave. To address this challenge, large array antennas need to be deployed at base stations in order to improve antenna gains and implement transmit beam sweeping. This requires UEs to support array antennas, multiple panels, and beam sweeping. mmWave also leads to an increase in the cost and power consumption of modems, an issue that will become crucial for modem implementation.

Array antennas on a UE are often integrated with the RF circuit in the form of an antenna in package (AiP) (also called antenna panel) to reduce the line losses between the antennas and RF channels. A common AiP includes four or five pairs of cross polarization antenna array elements.

Each element is connected to an RF channel, which includes various components (such as PAs and phase shifters), meaning that one AiP requires more than eight RF channels. This increases the number of RF channels required by mmWave by more than four times when compared with typical sub-6 GHz bands. Furthermore, the spatial beams of an AiP can cover only parts of a sphere, meaning that multiple AiPs are required to provide full coverage. The AiP cost increases linearly with the number of AiPs.

The power consumption of a modem includes RF and baseband power consumption. For an AiP with 8 to 10 RF channels, the RF transmit and receive power consumption is several times higher than that of a sub-6 GHz UE, which usually supports only 2 Rx, 4 Rx, and 2 Tx. Using more AiPs on a UE will lead to a further increase in the transmit and receive power consumption.

Due to mmWave beam management, the baseband power consumption of mmWave is higher than that of sub-6 GHz. A UE needs to perform beam sweeping periodically in order to select an optimal transmit/receive beam pair, regardless of whether data needs to be transmitted. Take synchronization signal block (SSB) sweeping as an example. Assume that a UE has two panels, each of which has seven receive beams, and the base station transmits 16 SSBs. The UE needs to measure beam pairs 224 times in order to select an optimal transmit/receive beam pair, significantly increasing the power consumption of beam management.

Because FR1 has been commercially deployed at scale, we should fully leverage this advantage to overcome the coverage, UE cost, and power consumption challenges faced by mmWave through coordination between FR1 and FR2. A typical deployment mode in the future will be co-deploying mmWave and FR1 at the same base station, where FR1 and FR2 can be coordinated in CA. With this mode, CA can use FR1 to improve uplink coverage without increasing the transmit power of UEs, overcoming the uplink coverage bottleneck of mmWave. Furthermore, using FR1 in the uplink can significantly reduce the power consumption of UEs when compared with FR2. This mode also enables FR1 and FR2 to share the same radio environment. Despite different scattering and refraction characteristics of electromagnetic waves with different frequencies, FR1 and FR2 channels are correlated to a certain degree in terms of LOS or NLOS and the multipath angle, delay, and Doppler effect. Consequently, in CA scenarios with both high and low frequency bands, the channel state information measured by a UE on FR1 can be used to assist communication on FR2 (e.g., beam management and measurement). This

significantly reduces the number of transmit/receive beam pairs to be measured by the UE and the power consumption of mmWave beam management.

4.2.4 Frequency Band Combinations

Different combinations of frequency bands and specifications of transmit and receive antennas increase the implementation complexity of RF modules and the modem cost. First, UEs need to support global roaming, meaning that they should be able to work on a wide range of frequency bands (usually about 20) and the combinations of these bands. Second, the size of a UE is limited, making it difficult to integrate many high-performance RF components to a UE. 5G offers a total of 70 frequency bands, including the typical bands (2.6 GHz, 3.5 GHz, 4.9 GHz, 700 MHz, and mmWave bands), and typical 4G bands that have been refarmed (e.g., 1.8 GHz and 2.1 GHz). In the future, more frequency bands will be introduced to 5G, such as 6 GHz and those between 7 GHz and 24 GHz. Although some of these bands overlap with each other, they have different scopes and specifications due to the laws and regulations imposed by different countries and regions. An example is the 18 frequency bands between 700 MHz and 900 MHz in 5G — these bands require different filters. 5G supports a wide range of frequency bands between 410 MHz and 7.125 GHz, and between 24.25 GHz and 52.6 GHz. No RF component can support such a large frequency range, as doing so would require multiple antennas, PAs, filters, and other components to be integrated into a UE. A next-generation communications system needs to adapt to all the frequency bands supported by earlier generations, making it difficult to reduce the complexity and cost of UEs.

5G standards define numerous frequency band combinations and different band utilization methods, including LTE-NR dual connectivity (DC), NR-CA, and NR-DC. The corresponding band combinations for each method are also defined. The characteristics of each band combination reflect the differences in terms of the numbers of bands, the numbers of carriers and bandwidths on each band, and the uplink concurrency mode. This will result in a large number of band combinations. Different CSPs use different frequency band combinations based on their spectrum acquisition and network deployment requirements in their regions. 3GPP has defined over 6000 frequency band combinations.

5G UEs have advanced RF specifications and need to support 2 Tx and 4 Rx on the typical frequency bands as well as various frequency band combinations, posing high requirements on the specifications of RF modules, PAs, power supply modules, and antennas. Specifications are an important factor that affects complexity.

In addition to implementation, RF tests also affect complexity. UEs have to pass many mandatory verification tests before they can access the network. These tests are performed separately for different frequency bands and band combinations. Testing a UE that supports more frequency bands and band combinations is more complex and expensive. Additionally, because the antennas and RF units of 5G mmWave UEs are integrated, the conduction test (using conducting wires to send signals to UEs in the lab) for traditional sub-6 GHz UEs is not applicable to 5G UEs. An over-the-air test needs to be conducted in a microwave anechoic chamber to verify the antennas and RF links of 5G UEs at the same time, significantly increasing the UE cost and complexity.

Fully considering the implementation and test complexity of modems, we propose the following proposals:

First, refarm spectrum and use low frequency bands (sub-3 GHz) in the uplink and use medium and high frequency bands (sub-6 GHz, mmWave) in the downlink, forming a combination of low-frequency + medium- and high-frequency bands with a large frequency spacing. The combination can easily meet the requirement on duplex isolation and simplify filter designs. Low frequency spectrum can be integrated into a few band groups in order to guide industry players to provide cost-effective and high-performance components. However, UEs that were developed before 5G may not be able to support the band groups.

Second, the characteristic tests of frequency band combinations need to be decoupled from RF tests. In addition to integrating frequency bands and reducing frequency bands, avoid repetitively testing the same group of band combinations for LTE-NR DC, NR CA, and NR DC.

In conclusion, the key to modem evolution lies in optimizing the method for introducing new frequency bands, integrating frequency bands, reducing the complexity, cost, and power consumption of RF modules on UEs, and delivering optimal performance.

5 Modem Flexibility and Compatibility

5.1 Flexible Adjustment of UE Capabilities

Traditionally, the design of communication between UEs and the network adopts a network-centric mechanism to facilitate centralized management of numerous UEs. Features and parameters are configured based on UE capabilities. UEs usually report their maximum capabilities to the network so that it can configure features and parameters accordingly in order to deliver an optimal modem performance. During 5G evolution, UEs need to support an increasing number of function and feature combinations. To reduce the cost, modem features are not designed according to the maximum UE capabilities. Instead, the features share UE capabilities flexibly and dynamically. Without UE capability sharing, a modem can deliver optimal performance only when the features do not work in parallel. If the features need to work in parallel, flexible capability sharing is required. However, the existing standards do not define an effective mechanism for updating UE capabilities dynamically. UEs can update their capabilities only by re-initiating mobility registration, which causes service interruptions and a poor user experience.

5G has introduced a UE assistance information reporting mechanism based on power consumption and capability constraints, allowing UEs to request preference configurations from the network based on their capabilities. However, commercial networks rarely provide the preference configurations requested by UEs, and the mechanism makes little difference. As such, a fast UE capability update or requirement negotiation and compliance mechanism needs to be added to the standards in the future. The mechanism allows modems to allocate capabilities on demand to different features and systems under various constraints, thereby delivering optimal performance and user experience.

The UE capability reporting mechanism in 5G inherits the characteristics of high flexibility and scalability. Its good forward and backward compatibility can facilitate smooth evolution of the entire system. Because 5G supports numerous features and configurations, the capability design and reporting mechanism creates many challenges for UEs to encode and store capabilities. Whereas a UE supports only a few capabilities, the entire system supports many capabilities. For example, over 600 UE capabilities are

supported in 3GPP Release 15, and over 1700 UE capabilities in 3GPP Release 17. A UE needs a large space for capability encoding and storage. Even a UE with advanced capabilities must report its maximum capabilities to the network due to being unable to obtain the network's capability information. Reporting such a large amount of information consumes a large amount of processing, storage, and air interface transmission resources on the UE. In future standards, UE capability encoding and storage overheads should be considered at the beginning of system design. Furthermore, a mechanism needs to be designed for UEs to quickly sense network capabilities based on the flexible UE capability adjustment mechanism mentioned earlier. This will enable UEs and the network to obtain and match capability information from each other in real time, reducing UE costs and facilitating high-quality evolution of the entire system.

In the future, the system design needs to introduce a mechanism of equal negotiation between UEs and the network to achieve optimization of the entire system.

5.2 Lean Signaling

Radio resource control (RRC) is one of the most complex and crucial modules on a modem. It is responsible for managing the control-plane processes, physical-layer configurations, MAC, and other L2 entities, and coordinating with the non-access stratum (NAS). As described in Section 5.1, as the number of band combinations increases, so too does the implementation complexity and cost of modems. This leads to a huge waste of modem resources because most of these combinations will never be used by the network. Over 800 information elements (IEs) and 600 UE capability items were introduced by 3GPP Release 16 and Release 17, respectively. To address the challenges created by these combinations, 6G should redefine the mechanism for designing modem control signaling from the very beginning.

The prevailing ASN.1-based 3GPP control signaling design mechanism ensures flexible and free communication of UEs manufactured by different vendors and served by different CSPs in different countries. However, the current ASN.1 signaling mechanism adopts the release-based mode, and modems implement the features in a release based on the ASN.1 profile of the release. This mode has many limitations. For example, new features in a version cannot be added to modems until the ASN.1 corresponding to the release is frozen (which usually takes 1.5 to 2 years), regardless of whether the market urgently requires the

features. A long ASN.1 freezing time increases the time to market of some new features and cannot meet the market demand quickly. Even if only one simple feature is required, the modem needs to upgrade the encoding and decoding mechanism for control signaling to the ASN.1 signaling profile of the latest release. Otherwise, the modem cannot process signaling bit streams correctly. This significantly increases modem complexity and storage overhead.

To address this challenge, a modular control signaling design mechanism can be developed. Such a mechanism would eliminate the dependence of standardized basic signaling design, enabling the signaling design mechanism to evolve from release-based toward release-free or release-independent. In a modular ASN.1 signaling design, each module is associated with a feature or combination of features, and ASN.1 encoding and decoding are performed independently in each module. The most basic features can be selected to form a basic signaling module architecture, known as the base module, at the beginning of system design. With the base module, both the network and UEs can load new modules to support new service features or remove unnecessary modules on demand. Furthermore, the base module enables the network and UEs to dynamically adapt to the service requirements of various networks or customers and add signaling function modules accordingly to support new services, as shown in Figure 10.

In a modular ASN.1 signaling design, the fundamental signaling module is the only module that needs to be maintained on a modem. The design significantly reduces modem implementation complexity and storage overhead, and allows new features and related signaling modules to be designed for the network and UEs based on new requirements, accelerating the implementation of new features. Additionally, the design can spread compatibility risks across different modules to improve overall modem stability.

As such, we propose the scheme of designing a release-free modular signaling mechanism based on a simplified modem design and using the modular ASN.1 design as the basis for developing an ASN.1 signaling mechanism based on feature modules.
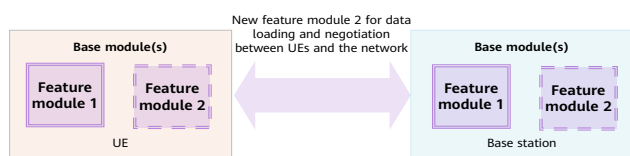
5.3 Native Robust Compatibility Mode

5G supports flexible extension of protocols to ensure both forward and backward compatibility and smooth evolution of standards. It also introduces many new features or combinations of features in new protocol versions, making it challenging to ensure protocol stability and UE compatibility. Compatibility issues will lead to painful repercussions for many UEs due to large-scale UE delivery. Especially, compatibility issues during the initial access state will cause UEs to lose basic connections. To avoid these potential risks, SDOs strive to maintain protocol stability and minimize compatibility risks during standards formulation. When the network or UEs are upgraded to support new features, the UEs, network, and CSPs will aim to ensure that legacy UEs are not affected by network upgrades and minimize the impact on user experience. Before features are put into commercial use, many full-coverage tests need to be conducted on UEs to avoid compatibility issues, significantly increasing the UE test cost. However, expertise accumulated on live networks shows that serious compatibility issues still may occur even if such tests have been conducted. For instance, because UEs and the network understand features differently or UEs have not been upgraded for a long time, the UEs cannot correctly adapt to protocol changes introduced by new features.

It is recommended that future standards introduce a built-in escape mechanism (e.g., built-in robust compatibility mode) to ensure that UEs can fall back quickly when severe compatibility issues occur. This mode can be used to implement automatic intervention and recovery, ensuring availability, as shown in Figure 11.

A built-in escape mechanism guarantees that UEs can set up a basic connection with the network or maintain basic communication even when serious compatibility issues occur, providing fundamental assurance for UEs to automatically recover communication and improving the robustness of the entire system.
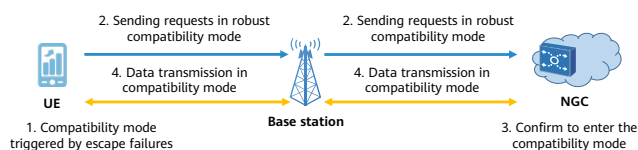

Figure 10 Modular ASN.1 signaling design

Figure 11 Robust compatibility mode

6 Modem Trustworthiness and Security

Over the past few years, Huawei has made tremendous efforts in improving the trustworthiness of every network component, including process and result trustworthiness. Trustworthiness, as discussed in many articles about 6G [6, 7, 9, 10, 12, 15], is regarded as a typical feature and requirement of future communications systems and the cornerstone for building an information society. To achieve E2E trustworthiness, the trustworthiness of modems should be considered in addition to the trustworthiness of core networks and base stations.
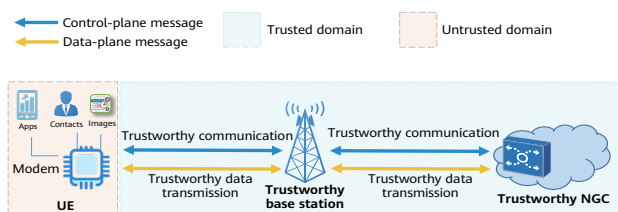
Figure 12 Modem trustworthiness is key to E2E trustworthiness

Figure 13 Potential threats to modems

The Balong team has introduced multiple mechanisms to improve the trustworthiness of modems. For example, it has provided various trustworthy mechanisms for UEs, including mutual authentication and user-plane integrity protection. However, modems still face many potential threats, as shown in Figure 13.

- Spoofing

  UEs camp on a fake base station and receive false system information (e.g., earthquake warnings), which may trigger the UEs to implement disaster handling mechanisms, such as the one for an emergency power failure.

- Tampering

  Signalling messages are tampered with before the air interface is activated in secure mode, or the user-plane rate exceeds the upper limit of the integrity protection rate, which causes data messages to be tampered with.

- Unforeseen protocol data

  Abnormal signaling formats, IE lengths, and field values lead to modem memory leakage or logic vulnerabilities, which causes the baseband processor to be hijacked into executing some hacked code (e.g., rewriting IMEIs).

- Information disclosure

  Null security algorithms (such as NEA0 or NIA0) that are designed for emergency services in restricted service mode are incorrectly used in other scenarios, causing the privacy information to be transmitted in plaintext.

- Denial of service (DoS)

  A fake base station attracts UEs to camp on it by transmitting high power signals, but it does not provide any services for them, and the UE is suspended on the fake base station.

- Elevation of privilege

  UEs still perform the "shall not" actions specified in the standards. For example, if a UE does not have the CAG permission, it should not attempt to access a private network cell and occupy air interface resources of other CAG-permitted UEs.

All information transmitted or received by UEs pass through the modem. If a modem — as a main communications interface for UEs — cannot be trusted, communication information on UEs and upper-layer services (including video call and remote industrial control) on UEs also cannot be trusted. A worse case is that an attacker may attack a UE through a modem to obtain the control permission of the UE or steal private information.

To ensure modem trustworthiness, it is recommended that a modem architecture that isolates secure environments from insecure ones be used in the design phase. This will prevent direct access from insecure environments to secure environments. Modem trustworthiness can be further improved from the perspectives of software (such as function modularization and software component boundary decoupling) and hardware (such as memory fault tolerance and ultra-high/low-temperature reliability).

Additionally, various modem tests need to be conducted during production in order to improve modem trustworthiness. For example, a configuration and operation check can eliminate possible human risks, and fuzz tests can thoroughly analyze the logic of all service state machines and identify potential logic vulnerabilities. At the later stage of production, third-party certification organizations are required to evaluate modem trustworthiness. Note that
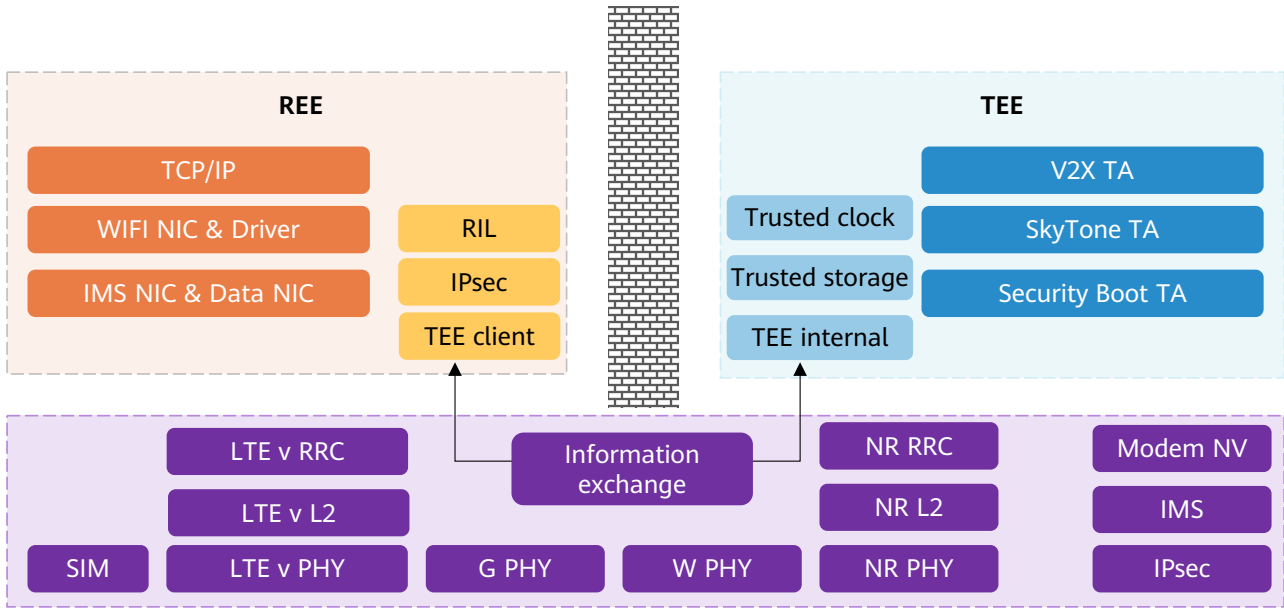
Figure 14 Trustworthy modem design

the prevailing certification mechanisms, including the CC certification system, are designed for the IT domain and can be used to evaluate only the trustworthiness of modem components (such as RAM and I/O interfaces shown in Figure 15) or execution environments. 3GPP RAN5 and GCF tests have been conducted for the CT function of modems, but these are positive tests designed to check whether a modem performs the corresponding actions according to the specifications. Reverse tests need to be conducted for the CT function of modems to check whether a modem can work properly in the case of the preceding threats.
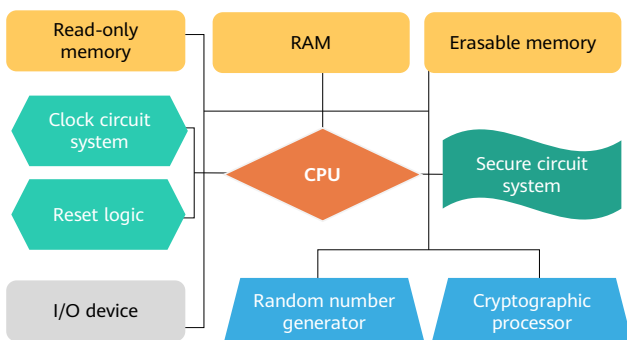
Figure 15 Modem from the perspective of IT

Assurance mechanisms, including fast fault detection, fault locating, and fault recovery, need to be implemented when modems are used to eliminate service interruptions. Besides, disaster recovery (DR) backup can be implemented to further improve modem trustworthiness. Because the DR backup cost of UEs continues to increase, a new backup

policy can be considered: perform module-level backup only for core modules. This would allow UEs to quickly switch to the backup module when the active module fails. Note that the module failure detection function on UEs may fail to identify some faults quickly (e.g., downlink sending failures). In this case, the network can help UEs detect module failures to guarantee that DR backup can be properly performed.

7 Conclusion

This paper describes our thoughts on the technical evolution of the Balong modem based on the study on new 5G-A services and technologies, potential requirements of 6G services and technologies, and the experience accumulated during the implementation and commercial use of the Balong modem. Technical insights are provided from the perspectives of deterministic network and experience, modem power consumption, complexity, compatibility, and trustworthiness. This lays a solid foundation for us to continuously improve the competitiveness of Balong modems as a world-leading modem provider.

References

[1] Wang Zhiqin, *IMT-2020 (5g) tuijinzu gongzuo jinzhan jieshao* [Work progress of the IMT-2020 (5G) Promotion Group], IMT-2020 (5G) Summit, 2019.7.17–18.

[2] China Mobile Terminal Lab, 5G *tongxin zhishu baogao* [5G communications index report, issue 1], 2020.5.

[3] Report ITU-R M.2516-0 (11/2022), Future technology trends of terrestrial IMT systems towards 2030 and beyond, 13th December 2022. https://www.itu.int/pub/R-REP-M.2516-2022

[4] ITU-R WP 5D, "Overview timeline for IMT towards the year 2030 and beyond," Workshop on "IMT for 2030 and beyond" during the ITU-R WP 5D meeting #41, 14th June 2022. https://www.itu.int/dms_ties/itu-r/oth/0a/06/R0A060000C80001PDFE.pdf

[5] IMT-2030 (6G) Promotion Group, "6G Vision and Candidate Technologies," 2021.6.

[6] IMT-2030 (6G) Promotion Group, 6G *dianxing changjing he guanjian nengli baipishu* [6G typical scenarios and key capabilities white paper], 5G Application and 6G Vision sub-forum of Digital China Summit, 2022.7.24.

[7] Wu Hequan, academician of the Chinese Academy of Engineering, "Ten thoughts on 6G R&D," 2nd Global 6G Technology Conference, 2022.3.23.

[8] Wu Hequan, academician of the Chinese Academy of Engineering, *Dui* 6G *yanjiu* de *zai sikao* [Rethinking 6G research]" Global 6G Conference, 2022.11.15.

[9] China Mobile, 6G *wangluo jiagou jishu baipishu* [6G network architecture technical white paper 2022]" 2022.6.22.

[10] China Unicom, 6G *baipishu* [6G white paper (V1.0)], 2021.3.

[11] NTT DoCoMo, White Paper: 5G Evolution and 6G, Jan. 2022 (Version 4.0). https://www.docomo.ne.jp/english/binary/pdf/corporate/technology/whitepaper_6g/DOCOMO_6G_White_PaperEN_v4.0.pdf

[12] Huawei, "6G: The next horizon white paper," 2022.1, https://www.huawei.com/en/technology-insights/future-technologies/6g-white-paper

[13] Ericsson, "6G-connecting a cyber-physical world," 11th Feb. 2022, https://www.ericsson.com/en/reports-and-papers/white-papers/a-research-outlook-towards-6g

[14] Samsung, "The next hyper-connected experience for all," https://research.samsung.com/next-generation-communications

[15] MediaTek, "6G vision whitepaper," 7th Jan. 2022, https://www.mediatek.com/blog/6g-whitepaper

[16] ITU-R WP 5D, "Framework and overall objectives of the future development of IMT for 2030 and beyond," June 2023.

[17] "TSMC unveils N2 process node: Nanosheet-based GAAFETs bring significant benefits In 2025," https://www.anandtech.com/show/17453/tsmc-unveils-n2-nanosheets-bring-significant-benefits

[18] Anders S. G. Andrae, Tomas Edler, "On global electricity usage of communication technology: Trends to 2030," *Challenges* 2015, *6*(1), 117–157.

[19] Report ITU-R M.2410-0, Minimum requirements related to technical performance for IMT-2020 radio interface(s).

[20] Erlang, Agner K. (1948), "Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges".

6G
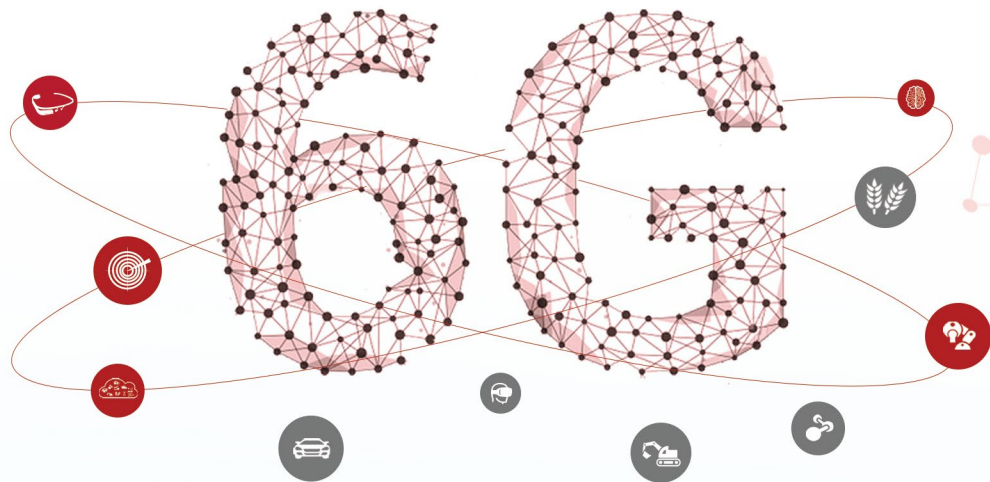
From Connected People and Things to Connected Intelligence

6G-Related Issues of *Communications of HUAWEI RESEARCH*

Issue 1

Issue 2

Issue 5 (Latest)

Videos from the 1st to 3rd Academic Symposiums on 6G ISAC

1st Symposium

2nd Symposium

3rd Symposium

HUAWEI TECHNOLOGIES CO., LTD.
Huawei Industrial Base
Bantian, Longgang
Shenzhen 518129, PRC
Tel: +86-755-28780808