An overview on integrated localization and communication towards 6G

Zhiqiang XIAO[1] & Yong ZENG[1,2*]

[1]*National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China;*
[2]*Purple Mountain Laboratories, Nanjing 211111, China*

Abstract While the fifth generation (5G) cellular system is being deployed worldwide, researchers have started the investigation of the sixth generation (6G) mobile communication networks. Although the essential requirements and key usage scenarios of 6G are yet to be defined, it is believed that 6G should be able to provide intelligent and ubiquitous wireless connectivity with Terabits per second (Tbps) data rate and sub-millisecond (sub-ms) latency over three-dimensional (3D) network coverage. To achieve such goals, acquiring accurate location information of the mobile terminals is becoming extremely useful, not only for location-based services but also for improving wireless communication performance in various ways such as channel estimation, beam alignment, medium access control, routing, and network optimization. On the other hand, the advancement of communication technologies also brings new opportunities to greatly improve the localization performance, as exemplified by the anticipated centimeter-level localization accuracy in 6G by extremely large-scale multiple-input multiple-output (MIMO) and millimeter wave (mmWave) technologies. In this regard, a unified study on integrated localization and communication (ILAC) is necessary to unlock the full potential of wireless networks for dual purposes. While there are extensive studies on wireless localization or communications separately, the research on ILAC is still in its infancy. Therefore, this article aims to give a tutorial overview on ILAC towards 6G wireless networks. After a holistic survey on wireless localization basics, we present the state-of-the-art results on how wireless localization and communication inter-play with each other in various network layers, together with the main architectures and techniques for localization and communication co-design in current two-dimensional (2D) and future 3D networks with aerial-ground integration. Finally, we outline some promising future research directions for ILAC.

Keywords wireless localization, integrated localization and communication, cellular networks, B5G, 6G

1 Introduction

Starting from the second generation (2G), wireless localization has been included as a compulsory feature in cellular networks, with continuous enhancement on the localization accuracy over each generation, e.g., from hundreds of meters accuracy in 2G to tens of meters in the fourth generation (4G). For the forthcoming fifth generation (5G) mobile networks, localization is regarded as one of the key components, due to its fundamental support for various location-based services, and the requirement on localization accuracy is up to sub-meter level [1]. While the deployment of 5G networks is ongoing, researchers around the world have already started the investigation on the sixth generation (6G) mobile communication targeting for network 2030, with various visions proposed [2–8]. For example, it was envisioned that 6G should achieve "ubiquitous wireless intelligence" [5], for providing users smart context-aware services through wireless connectivity anywhere in the world. This renders that acquiring the accurate location information of users becomes more critical than ever before, with potentially centimeter (cm)-level localization accuracy for 6G.

However, most current localization services provided by global navigation satellite systems (GNSS), wireless local area networks (WLAN) or existing cellular networks can at best achieve meter-level localization accuracy in clutter environments. Such coarse localization services are difficult to meet the

---

* Corresponding author (email: yong_zeng@seu.edu.cn)

cm-level localization accuracy requirements of many emerging applications. For example, the following three promising usage scenarios of 5G-and-beyond networks, namely, intelligent interactive networks, smart city, and automatic factory, all highlight the critical role of highly accurate localization in future network design.

(1) Intelligent interactive networks. It is believed that the ultimate goal of communication networks is to promote the intelligent interactions across the world, in terms of people-to-people, people-to-machine, and machine-to-machine. An unprecedented proliferation of new internet-of-things (IoT) services, like multisensory extended reality (XR) encompassing augmented/mixed/virtual reality (AR/MR /VR) [9], brain-computer interfaces (BCI) [3], as well as tele-presence and tele-control services [5], brings excellent opportunities to realize the goal of interaction with everything. To implement such new applications, it is necessary to achieve high localization performance, as elaborated in the following.

• **Multisensory XR.** XR services will enable users to experience and interact with virtual and immersive environments through first-person view [9]. To enable truly immersive XR applications, XR systems must be deployed through wireless networks, and the accurate tracking of XR devices is of paramount importance. For wireless XR applications, a control center collects the tracking information of the XR devices, and sends data to those devices through wireless links. Therefore, the accuracy of device tracking and the delay of signal measurements will significantly affect the XR information transmission and hence impact the user experience. For instance, an inaccurate head-tracking may cause cybersickness, like nausea, disorientation, headaches, and eye strain [10]. In general, for XR services, depending on the usage scenarios, the requirement for localization accuracy ranges from 1 to 10 cm, and the time delay should be typically less than 20 milliseconds (ms) [9].

• **Wireless BCI (WBCI).** The forthcoming 5G and future 6G networks bring new opportunities to tailor communication networks into versatile networks integrated with human-centric communication, wireless sensing, and remote control [3], where people will be enabled to interact with their surrounding environment by using various IoT devices connected through the WBCI technology. It opens the door for people to control their neighboring IoT devices through their brain implants, gestures, empathic as well as haptic messages [9]. Such a breathtaking technology requires communication support of extremely high data rate, ultra-low latency, and high reliability, together with high-performance localization services, e.g., cm-level accuracy. In addition, the cooperative localization among IoT devices is also quite important for WBCI.

• **Tele-presentation and Tele-control.** With the advancement of various supporting technologies including high-resolution imaging and sensing, wearable displays, mobile robots and drones, it is expected that the technologies of tele-presentation and tele-control will become reality in the near future [5]. For tele-presentation, a remote environment can be represented through real-time environment capturing, information transmission, and three-dimensional (3D) holographic rendering, which makes the accurate location information critical for 3D mapping. Furthermore, people may operate the remote IoT devices through wireless networks, just like manipulating them face-to-face, which is referred to as tele-control or tele-operation. An exemplary application of tele-presentation and tele-control is tele-surgery (or remote surgery), which will enable doctors to perform emergent surgery remotely. Note that in such use cases, highly-accurate, ultra-reliable, and low-latency localization is vital. For tele-control, the remote and neighboring localization systems usually have two separate coordinates, perform different localization algorithms, and use different reference nodes, which may cause mismatch errors. Therefore, for such applications, the real-time infrastructure calibration between the two localization systems is of paramount importance for reducing the mismatch errors.

(2) Smart city. A smart city has the ability to efficiently analyze different requirements from the society, and reasonably manage and optimize public resources, such as electricity, water, transportation, and healthcare, to provide better public services [11]. A truly smart city entails many different aspects. Here, we elaborate two major application scenarios to outline the importance of the accurate localization for smart city.

• **Smart indoor services.** Over the last decade, IoT technology has been advanced rapidly, which will flourish the smart indoor services, like smart homes, indoor navigation in shopping malls, and crowd monitoring. Different from outdoor scenarios, one critical issue for indoor localization is the severe non-line-of-sight (NLoS) signal propagation that may significantly degrade the localization accuracy [12,13]. Meanwhile, the privacy protection of location information is another important issue for public indoor localization services [5]. One of the key problems is to identify what kind of location information needs to be protected. For example, for some public devices, their location information should be accessible to

Table 1 The localization requirements of different applications for 5G/6G networks

Usage scenario	Application	Requirement
Intelligent interactive networks	Multisensory XR	Centimeter-level accuracy (i.e., 1–10 cm);
		low latency (less than 20 ms)
	WBCI	Centimeter-level accuracy;
		low latency (millisecond-level);
		high requirements on cooperative localization among IoT devices
	Tele-presentation and Tele-control	Centimeter-level accuracy;
		high reliability;
		calibration between two different localization systems
Smart city	Smart indoor services	NLoS-based localization;
		privacy information classification and protection
	Smart transportation	Submeter-level accuracy;
		wide coverage;
		high mobility tracking;
		cooperative localization in V2X communication cases
Automatic factory	CRAS	Cooperative localization among massive IoT devices;
		at least submeter-level accuracy;
		high reliability;
		low latency

all user devices, while that for user personal devices or some kernel public devices needs to be protected.

• **Smart transportation.** The research on smart transportation is still ongoing, with several standards proposed, like dedicated short-range communications (DSRC) [14] and vehicle-to-everything (V2X) [15]. Autonomous driving [16] and vehicle-to-vehicle (V2V) communications [17] are envisioned as two attractive developing trends of smart transportation, both of which call for advanced localization technologies. For autonomous driving, the 3D mapping for real-time scenarios is critical, which requires accurate distance information between the vehicle and obstacles to construct the environment model. Compared with other use cases, for smart transportation, the localization systems should be designed not only for high accuracy, but also for wide coverage and robustness in highly mobile scenarios.

(3) Automatic factory. The development of connected robotics and autonomous systems (CRAS) like autonomous robotics and drone-delivery systems, promotes the progress of automatic factory [3], such as smart storage, autonomous production, and autonomous delivery. To enable effective cooperation among various IoT devices, the accurate localization is the prerequisite. Different from other applications, for automatic factory, cooperative localization among massive IoT devices is of critical importance, which will require the highly-accurate, low-latency, and highly-reliable location information of the massive devices to build the end-to-end (E2E) communication links.

Table 1 summarizes the main localization requirements of different applications for 5G/6G networks. It is observed that achieving highly-accurate, low-latency, and highly-reliable localization will play an important role in future wireless networks.

To achieve the above goals, the underlying technologies of 5G-and-beyond mobile communication, like millimeter wave (mmWave), massive multiple-input multiple-output (MIMO), and ultra dense networks (UDNs), can be utilized for improving the localization performance. The mmWave communication operates at frequencies around 30 GHz and above, with signal bandwidth up to 2 GHz, which is able to provide high temporal resolution for time-based localization [18]. Furthermore, extremely large-scale MIMO (XL-MIMO) consisting of thousands of antenna elements can achieve angular resolution less than 1° for angle-based localization [19]. Besides, UDNs are expected to increase the likelihood of line-of-sight (LoS) links, which can also be exploited for improving localization performance [18].

Furthermore, one promising trend of future wireless networks is to integrate communication, computing, control, localization, and sensing (3CLS) [3], which can maintain the key performance indicators (KPIs) by appropriately managing the radio resources according to the real-time locations of mobile terminals. Therefore, as a part of the 3CLS, the integrated localization and communication (ILAC) is critical to explore the full potential of wireless networks in the following aspects.

• Radio signals can simultaneously carry data and location-related information of the transmitters, which renders ILAC a natural choice for the most efficient utilization of radio resources and cost-effective sharing of wireless infrastructure.

• Channel estimation for communication and location estimation for localization share similar signal

processing procedures, which makes it possible to share certain hardware components, for cost-effective and compact wireless devices.

• The location information of mobile terminals can be utilized for improving communication performance, which is referred to as the location-aware communication [20–22], such as location-aided channel estimation, beam alignment, and network optimization.

• Advanced communication technologies can be utilized for localization performance enhancement, and a unified design on signal waveforms, coding, modulation, and radio resource allocation can be pursued.

While there are extensive studies focusing on wireless communications or localizations [23–30], to our best knowledge, a tutorial overview on ILAC to fully utilize the network infrastructure and radio resources for dual purposes is still missing. In [23–25], the authors provided surveys on indoor localization. Specifically, Ref. [23] focused on the principles of different localization approaches and provided an overview on various localization infrastructures. In [24], the advanced techniques, such as data fusion, cooperative localization, and game theory, were highlighted to improve the localization performance. In [25], the authors provided an up-to-date overview on indoor localization with emphasis on IoT scenarios. Both Refs. [26, 27] gave overviews on the enabling technologies for network localization, tracking, and navigation, where Ref. [26] mainly focused on the mathematical theories of different indoor tracking and mapping methods. and Ref. [27] gave a comprehensive review on localization techniques in cellular network, WLAN, and wireless sensor networks (WSNs). In [28, 29], the authors discussed the localization techniques from the signal processing perspective. In [30], a survey of cellular-based localization was given, which discussed the evolution of the conventional cellular localization methods, together with the envision on the 5G new radio (NR) localization.

In this article, we aim to provide a tutorial overview on ILAC towards 6G. To this end, we first give a holistic introduction on wireless localization basics, in terms of the main definitions and classifications of localization systems, different localization approaches, fundamental performance analysis and metrics, the major localization infrastructures, and some advanced localization related techniques. Then we focus on the ILAC targeting for the future wireless network design. To this end, we first provide an overview on the recent third generation partnership project (3GPP) standardization for 5G localization, and then discuss the enabling technologies of 5G networks towards cm-level localization accuracy. After that, we present the state-of-the-art location-aided communication techniques to expose how wireless localization and communication inter-play with each other in different network layers, and discuss the techniques of localization and communication co-design to fully utilize the network infrastructures and radio resources of wireless networks. Furthermore, a discussion on ILAC for aerial-ground integrated networks will be given, which aims to facilitate the ubiquitous wireless coverage, moving from the conventional two-dimensional (2D) plane to the 3D space. Finally, we give an architecture of future wireless networks, and point out some promising future research directions for ILAC.

2 Wireless localization basics

2.1 Definition and classification of wireless localization

As illustrated in Figure 1, a wireless localization system aims to estimate the location of the targeting object based on the reference signals propagated between the reference nodes and the targeting object, Following the standard localization literature [26], we term the targeting object as the agent node and the reference nodes as the anchor nodes (ANs) throughout this article. Compared to a closely related terminology, wireless positioning, which estimates the position of the agent node relative to the ANs, wireless localization further locates the estimated position on a coordinate of a map based on the locations of the ANs. Nonetheless, the terminologies positioning and localization are often used interchangeably [30].

Typically, a wireless localization system consists of two major components: a set of ANs and a location estimation unit that can be deployed either on the agent node or a remote site. The procedures of localization usually include two stages. In the first stage, specific reference signals are transmitted either by the ANs or the agent node, which are measured by the receiver to obtain some location-related information, such as received signal strength (RSS), time of arrival (TOA), time difference of arrival (TDOA), or angle of arrival (AOA). In the second stage, such measurements are collected by the location estimation unit to estimate the location of the agent node. Localization systems can be categorized from various perspectives, like based on location estimation algorithms [27] or localization infrastructures [24].
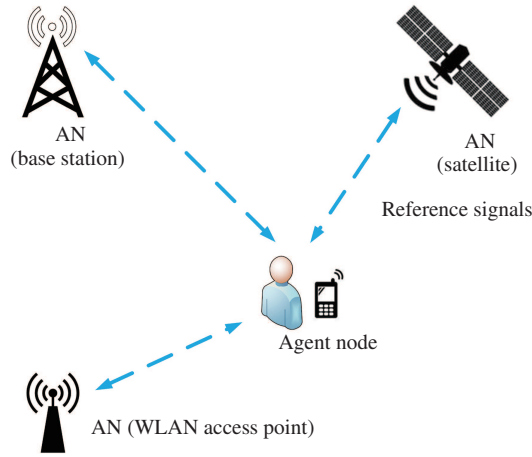
Figure 1 (Color online) A general architecture of wireless localization, with heterogeneous wireless infrastructures including cellular networks, navigation satellites, and WLAN.
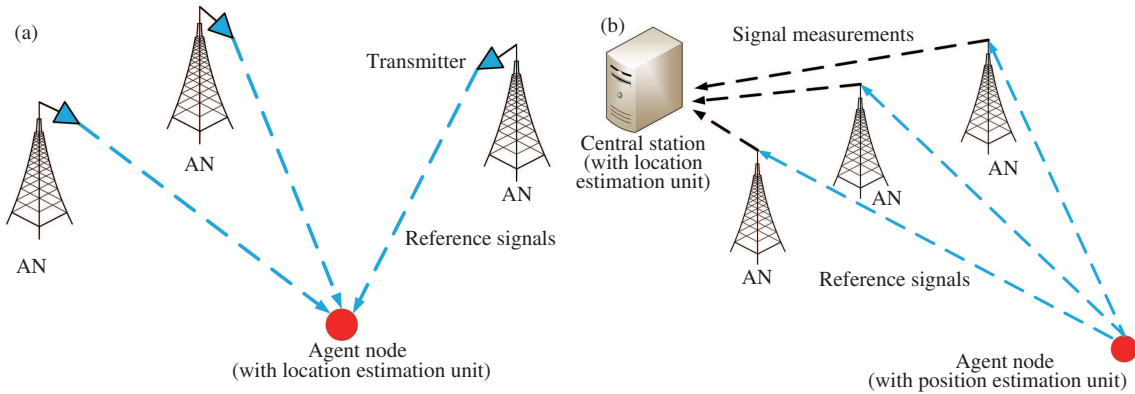

Figure 2 (Color online) An illustration of self-localization and remote localization systems. (a) For self-localization, the location estimation is performed by the agent node itself. (b) For remote localization, the location estimation is performed by a remote central station.

One popular classification is to consider where the location estimation is performed [31], based on which we have self-localization or remote localization systems.

2.1.1 *Self-localization*

As illustrated in Figure 2(a), for self-localization systems, a location estimation unit is deployed on the agent node, which receives reference signals transmitted from several ANs. The agent node has the capability to perform appropriate signal measurements, based on which its location is estimated. Self-localization systems have several advantages. First, since almost all localization-relevant operations are performed locally at the agent node, the localization speed is mainly dependent on the computational capability of the agent node. Therefore, such systems are easier for performance enhancement via updating the computational or measurement units of the agent node, without having to modify the network infrastructure. Second, self-localization systems have the inherent mechanism for user privacy protection, since the agent node only passively receives signals transmitted from ANs, with little risk of location information leakage from the user side. Finally, for some dynamic localization scenarios like tracking and navigation, various onboard sensors like inertial measurement units (IMUs) can be handily equipped on the agent node to provide some motion information for improving the localization accuracy [28]. However, self-localization systems have high hardware requirements on the agent node, such as high caching and computational capability for signal measurement and location estimation. Therefore, such systems, like global positioning system (GPS) or inertial navigation system (INS), can only be deployed on the devices with powerful caching and computational components.

2.1.2 *Remote localization*

For remote localization systems, as illustrated in Figure 2(b), the reference signals are transmitted from the agent node to ANs. Upon receiving the reference signals, the ANs would send their respective signal measurements to a remote central station, where the location estimation is performed. The main advantage of remote localization systems over the self-localization counterparts is the less demanding on the agent node, since almost all time-consuming and complex computing operations are performed at the remote central station, like the cellular base stations (BSs) or the computing center. Therefore, remote localization is especially appealing for resource-limited devices, such as IoT devices and wireless sensor nodes. In addition, different from self-localization systems where the location information is only acquired by the agent node, remote localization systems can preserve locations for all agent nodes in the area of interest, which can be utilized for various purposes, like location-aware communication [20]. However, since all location information of agent nodes is stored in a remote server, the information privacy and security is a critical issue for remote localization.

2.2 Basic localization techniques

In general, localization techniques can be classified into two main categories: direct localization and two-step localization. For direct localization [32–34], the received signals are directly used to estimate the location of the agent node, whereas for two-step localization, the location-related information, such as RSS, TOA, TDOA, and AOA, is firstly extracted from the received signals, and then used for target localization. Note that in principle, direct localization can achieve better performance than two-step localization. However, by considering the complexity and implementation constraints, two-step localization approaches are usually used in practical systems. Typically, the two-step localization approaches can be further classified into geometric-based, scene analysis (also known as fingerprinting), and proximity approaches [23, 35], as discussed in details in the following.

Consider a basic wireless localization system consisting of N ANs and one agent node. The locations of the ANs are known, denoted by $\boldsymbol{p}_n, n = 1, \ldots, N$, while that of the agent node needs to be estimated, denoted by $\boldsymbol{w}$. Regardless of self- or remote localization, a two-step localization approach can be interpreted as a parameter estimation problem. For the first step for signal measurement, the location-related information is obtained from the received signals, which are in general affected by multi-path effects and NLoS propagation. The general measurement model can be expressed as

$$r_n = h(\boldsymbol{p}_n, \boldsymbol{w}) + e_n, \ n = 1, 2, \ldots, N, \tag{1}$$

where r_n denotes the generic signal measurement associated with the nth AN, $h(\cdot)$ is a nonlinear function which contains all necessary information to compute the location of the agent node, and e_n represents the measurement error. Note that for geometric-based methods, the exact expressions of $h(\cdot)$ can be established (as given in Subsection 2.2.1) based on algebraic relationships between ANs and the agent node. However, for fingerprinting-based methods, the pre-built fingerprints are usually treated as $h(\cdot)$. For the second step for location estimation, the main task is to solve the systems of nonlinear equations in (1) to estimate $\boldsymbol{w}$ based on the obtained signal measurements $\{r_n\}_{n=1}^{N}$. Such nonlinear equations are difficult to solve directly and deserve detailed discussion. In this subsection, we mainly focus on the principles and characteristics of different measurement models, while a comprehensive analysis about the location estimators to solve those nonlinear equations is deferred to Subsection 2.3.

2.2.1 *Geometric-based localization*

As illustrated in Figure 3, geometric-based localization technologies exploit the geometric properties of triangles to locate the agent node. Typically, geometric-based methods have two variations: trilateration and triangulation. Trilateration uses the distance-related signal measurements, like RSSs [36–38], TOAs [39–44] or TDOAs [45–47] of the received signals to locate the agent node. On the other hand, triangulation usually measures the AOAs of the received signals, and locates the agent node at the intersection of angle direction lines [48, 49], where the AOAs can be measured with the aid of directional antennas or antenna arrays.

RSS. RSS-based localization uses the average signal power attenuation between the ANs and the agent node to estimate their distances, based on which a geometric model is formulated to estimate the location of the agent node [36–38]. For example, as illustrated in Figure 3(a), the location of an agent node can
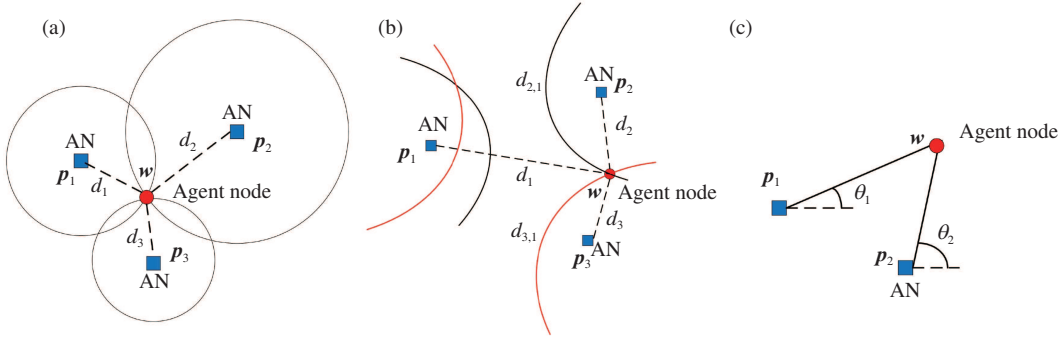
Figure 3 (Color online) Geometric-based localization in 2D space with perfect signal measurements, where $\boldsymbol{p}_n \in \mathbb{R}^{2 \times 1}$, $n = 1, 2, 3$, denote the ANs, $\boldsymbol{w} \in \mathbb{R}^{2 \times 1}$ is the agent node to be located. (a) RSS- or TOA-based localization, where d_n, $n = 1, 2, 3$ denote the actual distances between the agent node to the three ANs. (b) TDOA-based localization, where $d_{n,1}$, $n = 2, 3$, represent the relative distance between a pair of ANs to the agent node. The two pairs of red and black lines correspond to the hyperbola curves. (c) AOA-based localization, where θ_i, $i = 1, 2$ are the angles between the ANs and the agent node.

be ideally determined in 2D space with three ANs. In general, the average received power $P_{r,n}$ in dB associated with the nth AN can be modelled as [50]

$$P_{r,n} = P_0 - 10\alpha_n \log_{10} d_n + e_{\text{RSS},n}, \quad n = 1, 2, \ldots, N, \tag{2}$$

where P_0 is the reference received power at a reference distance of 1 meter (m), $d_n = \|\boldsymbol{p}_n - \boldsymbol{w}\|$ is the actual distance between the nth AN and the agent node, α_n denotes the path loss exponent, and $e_{\text{RSS},n}$ represents the RSS measurement error. Assuming that P_0 and α_n, $n = 1, 2, \ldots, N$ are known, the distance between the agent node and each of the ANs can be estimated. The generic model in (1) for RSS-based localization can be written as

$$r_{\text{RSS},n} = h_{\text{RSS}}(\boldsymbol{p}_n, \boldsymbol{w}) + e_{\text{RSS},n}, \quad n = 1, 2, \ldots, N, \tag{3}$$

where $r_{\text{RSS},n} = P_{r,n} - P_0$, and $h_{\text{RSS}}(\boldsymbol{p}_n, \boldsymbol{w}) = -10\alpha_n \log_{10}(\|\boldsymbol{p}_n - \boldsymbol{w}\|)$. The remaining task is to estimate $\boldsymbol{w}$ based on the obtained $\{r_{\text{RSS},n}\}_{n=1}^{N}$ in (3), which is discussed in Subsection 2.3.

The main advantage of RSS-based localization lies in that it does not require the time synchronization among different nodes, and RSS measurements are readily available for almost all practical wireless systems. In addition, different from TOA, TDOA, or AOA-based localization, RSS measurements do not rely on line-of-sight (LoS) signal propagation. However, the main drawback of RSS-based approaches is the poor localization accuracy, especially in clutter environments, since the signal attenuation in these environments is only weakly correlated with distance, leading to poor accuracy for distance estimation [36]. Besides, an accurate signal propagation model is necessary for the reliable RSS-based distance estimation, which is challenging due to the unpredictable variations of the channel behavior. Therefore, in practice, RSS-based localization is mostly adopted for applications with coarse localization accuracy requirements.

TOA. As illustrated in Figure 3(a), TOA-based approaches first estimate the distances between the agent node and each of the ANs by using the signal propagation delay or time of flight (TOF), denoted by τ_{f}, and then build the trilateration model to locate the agent node [39–44]. Typically, depending on how τ_{f} is defined, TOA-based methods can be further divided into one-way TOA (OW-TOA) [42] and two-way TOA (TW-TOA) [43].

For OW-TOA localization, node A (either the AN or the agent node) transmits to node B a packet that contains a timestamp τ_s recording the time when the packet was sent, and node B then measures the TOA of the received signal, denoted by τ_r. If the time clock between the ANs and the agent node is perfectly synchronized, it is clear that τ_{f} can be determined by node B as $\tau_{\text{f}} = \tau_r - \tau_s$, and the distance between nodes A and B can be calculated as $d = \tau_{\text{f}} \cdot c$, where c is the speed of light. However, OW-TOA methods have two main drawbacks. First, a small time synchronization error between the agent node and ANs may significantly degrade the distance estimation. Second, the transmitted signal must be labeled with a timestamp to allow the receiving node to calculate τ_{f}, which increases the complexity of the signal structures and may cause additional estimation error.

On the other hand, for TW-TOA localization, node A transmits a packet to node B, which responds by sending an acknowledgement packet to node A after a response delay τ_d. Provided that τ_d is known, node A can calculate its distance to node B based on the signal round-trip time of flight (RTOF), i.e.,

$\tau_{\text{RT}} = 2\tau_{\text{f}} + \tau_d$. TW-TOA addresses the first drawback of OW-TOA by avoiding the requirement of time synchronization between the two nodes. However, in practice, it is still difficult for the measurement node, i.e., node A, to know the exact response delay τ_d. Although τ_d could be ignored if it is relatively small compared with τ_{f} in long-range signal propagation, it critically affects the performance for short-range scenarios. Furthermore, while TW-TOA method eliminates clock synchronization error between the two nodes, relative clock drift could compromise the distance estimation accuracy. In addition, timestamp is still needed for TW-TOA to compute the RTOF of the transmitted signal.

Mathematically, a general TOA-based measurement model is formulated as [39,41]

$$\text{c} \cdot \tau_{\text{f},n} = d_n + e_{\text{TOA},n}, \quad n = 1, 2, \ldots, N, \tag{4}$$

where $\tau_{\text{f},n}$ is the measured TOF of signal propagation between the nth AN and the agent node, which is typically affected by positive bias due to the NLoS propagation and system errors, as captured by the additional measurement error $e_{\text{TOA},n}$. Following similar notations in (3), let $r_{\text{TOA},n} = \text{c} \cdot \tau_{\text{f},n}$ and $h_{\text{TOA}}(\boldsymbol{p}_n, \boldsymbol{w}) = d_n = \|\boldsymbol{p}_n - \boldsymbol{w}\|$, the generic model in (1) for TOA-based localization can be written as

$$r_{\text{TOA},n} = h_{\text{TOA}}(\boldsymbol{p}_n, \boldsymbol{w}) + e_{\text{TOA},n}, \quad n = 1, 2, \ldots, N. \tag{5}$$

By solving systems of nonlinear equations in (5), the location of the agent node can be estimated.

TDOA. For each TDOA measurement, the agent node will lie on a hyperboloid with a constant distance difference between a pair of ANs to the agent node. Specifically, the equation of the hyperboloid is given by [23]

$$d_{i,j} = \|\boldsymbol{p}_i - \boldsymbol{w}\| - \|\boldsymbol{p}_j - \boldsymbol{w}\|, \quad i \neq j, \tag{6}$$

where $d_{i,j}$ denotes the relative distance between the ith and jth ANs to the agent node, which can be estimated by the TDOA measurements. For instance, in 2D localization cases, as illustrated in Figure 3(b), the location of the agent node $\boldsymbol{w}$ could be theoretically estimated from the two intersections of the two hyperbolas with two pairs of foci, i.e., $\boldsymbol{p}_1$ versus $\boldsymbol{p}_2$ and $\boldsymbol{p}_1$ versus $\boldsymbol{p}_3$. Assuming that a reference signal was sent at an unknown time τ_0, which is then received by the ith and jth measurement units at time τ_i and τ_j, respectively. The TDOA between the ith and jth measurement units is [35]

$$\tau_{i,j} = (\tau_i - \tau_0) - (\tau_j - \tau_0) = \tau_i - \tau_j. \tag{7}$$

For a localization system consisting of N ANs, there are $N(N-1)/2$ TDOAs from all possible pairs of ANs, but only $N-1$ of them are non-redundant. Without loss of generality, we consider the first measurement unit as the reference and the non-redundant TDOAs are $\tau_{n,1}, n = 2, 3, \ldots, N$. In practice, TDOA measurements suffer from positive bias errors introduced by clock synchronization error among ANs, the multi-path effects, and so on. Therefore, similar to (4), the TDOA measurement model is

$$c \cdot \tau_{n,1} = d_{n,1} + e_{\text{TDOA},n}, \quad n = 2, 3, \ldots, N, \tag{8}$$

where $d_{n,1}$ is the relative distance between the nth AN and the 1st AN to the agent node, which is defined in (6), and $e_{\text{TDOA},n}$ represents the measurement error. Following the definition in (3) and (5), the generic model in (1) for TDOA-based localization can be expressed as

$$r_{\text{TDOA},n} = h_{\text{TDOA}}(\boldsymbol{p}_n, \boldsymbol{w}) + e_{\text{TDOA},n}, \quad n = 2, 3, \ldots, N, \tag{9}$$

where $r_{\text{TDOA},n} = \text{c} \cdot \tau_{n,1}$ and $h_{\text{TDOA}}(\boldsymbol{p}_n, \boldsymbol{w}) = \|\boldsymbol{p}_n - \boldsymbol{w}\| - \|\boldsymbol{p}_1 - \boldsymbol{w}\|$.

TDOA-based methods overcome the drawbacks of the TOA-based methods. First, it only requires time synchronization among ANs, which can be achieved by using wire backbone networks [35]. Another advantage of TDOA-based localization is that the timestamp is no longer needed, as evident from the TDOA computation in (7). This simplifies the structure of the transmitted signals and avoids the potential sources of error. However, all time-based positioning methods, like TOA and TDOA, rely heavily on the LoS path to compute TOA or TDOA information, which renders them venerable to NLoS environment.

AOA. AOA-based localization uses the angles between ANs and the agent node to determine the location of the agent node [48,49]. As shown in Figure 3(c), for 2D localization, AOA-based methods require only two known ANs with two measured angles to determine the location of the agent node. To measure the AOAs, the ANs should be equipped with antenna arrays or directional antennas with spatial resolution capabilities.

For 2D localization, the actual AOA between the agent node $\boldsymbol{w} = [w_x, w_y]^{\mathrm{T}}$ and the nth AN $\boldsymbol{p}_n = [p_{x,n}, p_{y,n}]^{\mathrm{T}}$ can be expressed as [35]

$$\theta_n = \tan^{-1}\left(\frac{w_y - p_{y,n}}{w_x - p_{x,n}}\right), \quad n = 1, 2, \ldots, N, \tag{10}$$

where $\theta_n \in (-\pi, \pi)$ represents the azimuth angle in a 2D plane. In the presence of angle estimation errors, the AOA measurements can be modeled as

$$r_{\mathrm{AOA},n} = h_{\mathrm{AOA}}(\boldsymbol{p}_n, \boldsymbol{w}) + e_{\mathrm{AOA},n}, \quad n = 1, 2, \ldots, N, \tag{11}$$

where $r_{\mathrm{AOA},n}$ is the measured AOA with the error $e_{\mathrm{AOA},n}$, and $h_{\mathrm{AOA}}(\boldsymbol{p}_n, \boldsymbol{w}) = \tan^{-1}(\frac{w_y - p_{y,n}}{w_x - p_{x,n}})$. For 3D localization, considering the agent node at $\boldsymbol{w} = [w_x, w_y, w_z]^{\mathrm{T}}$ and the ANs at $\boldsymbol{p}_n = [p_{x,n}, p_{y,n}, p_{z,n}]^{\mathrm{T}}, n = 1, 2, \ldots, N$, their azimuth and elevation angles can be expressed as [49]

$$\begin{bmatrix} \theta_n \\ \phi_n \end{bmatrix} = \begin{bmatrix} \tan^{-1}\left(\dfrac{w_y - p_{y,n}}{w_x - p_{x,n}}\right) \\ \tan^{-1}\left(\dfrac{w_z - p_{z,n}}{(w_x - p_{x,n})\cos\theta_n + (w_y - p_{y,n})\sin\theta_n}\right) \end{bmatrix}, \tag{12}$$

where $\phi_n \in (-\pi/2, \pi/2)$ represents the elevation angle. Therefore, similar to (11), the AOA measurement model in 3D space is

$$\boldsymbol{r}_{\mathrm{AOA},n} = \boldsymbol{h}_{\mathrm{AOA}}(\boldsymbol{p}_n, \boldsymbol{w}) + \boldsymbol{e}_{\mathrm{AOA},n}, \quad n = 1, 2, \ldots, N, \tag{13}$$

where $\boldsymbol{r}_{\mathrm{AOA},n} = [\hat{\theta}_n, \hat{\phi}_n]^{\mathrm{T}}$ is a vector consisting of the measured azimuth $\hat{\theta}_n$ and elevation $\hat{\phi}_n$ with measurement errors $\boldsymbol{e}_{\mathrm{AOA},n} = [e_{\theta_n}, e_{\phi_n}]^{\mathrm{T}}$, and $\boldsymbol{h}_{\mathrm{AOA}}(\boldsymbol{p}_n, \boldsymbol{w}) = [\theta_n, \phi_n]^{\mathrm{T}}$ is defined in (12).

The advantage of AOA-based localization lies in that it does not require time synchronization between the measuring units, and 2D localization only requires two ANs. However, the ANs for AOA-based localization need to equip with highly directional antennas or large antenna arrays for highly accurate angular resolution. Besides, in the complex multi-path environment, like urban areas or indoor scenarios, the AOA estimation is subject to significant errors due to the NLoS propagation.

2.2.2 *Scene analysis/fingerprinting-based localization*

Since the performance of geometric-based localization approaches degrades significantly in complex environments, alternative approaches based on scene analysis or fingerprinting have been proposed [51–56]. Such methods first exploit the data collected by the sensors, like cameras, accelerometer, or specific WiFi access points (APs), to extract unique geotagged signatures, i.e., fingerprints, and then pinpoint the location of the agent node by matching the online signal measurements against the pre-recorded geotagged fingerprints. Depending on the types of fingerprints, fingerprint-based localization can be classified into visual-, motion-, and signal-based [54]. In this article, we mainly focus on radio frequency (RF) signal fingerprint-based methods. In many applications, signal fingerprints usually correspond to the RSS indicators (RSSIs) instead of TOA, TDOA or AOA [55]. The main reason is that the time- or angle-based signal measurements rely heavily on LoS geometric assumption, which is hardly satisfied in complex environments. Therefore, a variety of RSSI fingerprint-based localization systems have been developed, like RADAR system [57] that employs WiFi signals for indoor localization.

In general, fingerprinting-based localization includes offline training and online localization phases. As illustrated in Figure 4, taking the 2D localization as an example, the area of interest is firstly divided into L cells, where the location of the lth cell is known, denoted by $\boldsymbol{z}_l = [z_{x,l}, z_{y,l}]^{\mathrm{T}}, l = 1, 2, \ldots, L$. For offline training, a mobile test node travels the L cells and communicates with the N ANs to measure the RSS fingerprints, which are denoted by $\boldsymbol{s}_l = [s_{l,1}, s_{l,2}, \ldots, s_{l,N}]^{\mathrm{T}}$ for measurements at $\boldsymbol{z}_l, l = 1, 2, \ldots, L$. Then, the entire radio map of the area of interest is obtained as $\boldsymbol{F} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_L]^{\mathrm{T}}$, which would be stored in a database for online localization. For online localization, a real-time RSS measurements of the agent node at the location of $\boldsymbol{w} = [w_x, w_y]^{\mathrm{T}}$ are measured as $\boldsymbol{s}_{\boldsymbol{w}} = [s_{\boldsymbol{w},1}, s_{\boldsymbol{w},2}, \ldots, s_{\boldsymbol{w},N}]^{\mathrm{T}}$, and the location of the agent node can be estimated as $\hat{\boldsymbol{w}} = [\hat{w}_x, \hat{w}_y]^{\mathrm{T}}$ based on a rule $g(\cdot)$ that compares the received online measurements $\boldsymbol{s}_{\boldsymbol{w}}$ against the radio map $\boldsymbol{F}$ [52, 56], i.e.,

$$\hat{\boldsymbol{w}} = g\left(\boldsymbol{F}, \boldsymbol{s}_{\boldsymbol{w}}\right). \tag{14}$$
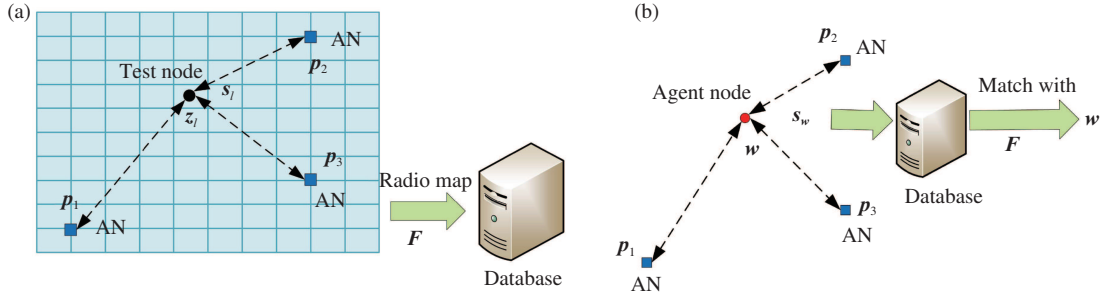
Figure 4 (Color online) The procedures of fingerprint-based localization, where $\boldsymbol{p}_n, n = 1, 2, 3$ denote the ANs. (a) Offline training phase, where $\boldsymbol{s}_l$ is the RF fingerprint at the cell $\boldsymbol{z}_l$. After the test node travels the whole area of interest, the radio map $\boldsymbol{F}$ consisting of all fingerprints in the area will be stored in a database. (b) Online localization phase, where the online RSS measurements $\boldsymbol{s}_w$ of the agent node is sent to the database. After pattern matching $\boldsymbol{s}_w$ against with the radio map $\boldsymbol{F}$, the location of the agent node can be estimated as $\hat{\boldsymbol{w}}$.

Depending on the mathematical model of RSS fingerprints, there are two main conventional localization approaches: deterministic and probabilistic.

Deterministic. For deterministic approaches, the measured RSSs are assumed to be static. In general, the location of the agent node is estimated to be the location of the cell, whose fingerprint is the closest to the online RSS measurements [56], i.e.,

$$\hat{\boldsymbol{w}} = \arg \min_{l=1,\ldots,L} d\left(\boldsymbol{s}_l, \boldsymbol{s}_w\right), \tag{15}$$

where $d(\cdot)$ denotes a certain distance metric. For example, applying the Euclidean distance metric in (15), we can obtain

$$\hat{\boldsymbol{w}} = \arg \min_{l=1,\ldots,L} \left\| \boldsymbol{s}_l - \boldsymbol{s}_w \right\|. \tag{16}$$

Solving (16) to estimate the location of the agent node is known as the nearest neighbor (NN) method. Another well-known deterministic method is the K-nearest neighborhood (KNN) [23, 53], where the location estimation is obtained by averaging the locations of the K cells in the radio map that have the nearest distances. The weighted KNN (WKNN) is a variant of the KNN method [52, 56], where the selected locations of the closest cells are combined by assigning weights to estimate the location of the agent node, where the weights are usually proportional to the inverse of their corresponding $d\left(\boldsymbol{s}_l, \boldsymbol{s}_w\right)$. In general, KNN and WKNN methods can achieve better performance than NN method. However, as the density of the radio map increases, NN method may achieve comparable performance as KNN or WKNN methods. Technically, RF fingerprinting methods originate from machine learning classification, so other machine learning methods such as support vector machine (SVM) and linear discriminant analysis (LDA) can be also used for location fingerprinting [23]. Such methods can achieve better localization accuracy compared with KNN, WKNN, or NN, but with higher computational complexity. The main advantage of the deterministic approaches is their simplicity. However, a single static RSS fingerprint of a cell may not be sufficient to uniquely represent the feature of the cell due to the time-varying nature of wireless signal propagation, so probabilistic approaches are developed.

Probabilistic. Probabilistic methods use statistical inference between online signal measurement $\boldsymbol{s}_w$ and the stored radio map $\boldsymbol{F}$ to estimate the location of the agent node, where the RSS fingerprint at a cell is treated as a random vector and the knowledge of RSS distribution in the area of interests is acquired through offline training. The underlying principle of probabilistic localization is the maximum a posteriori (MAP) estimation, which estimates the location of the agent node by maximizing the conditional probability of the location given the online RSS measurements [58], i.e.,

$$\hat{\boldsymbol{w}} = \arg \max_{\boldsymbol{z}_l, l=1,\ldots,L} P\left(\boldsymbol{z}_l | \boldsymbol{s}_w\right), \tag{17}$$

where $P\left(\boldsymbol{z}_l | \boldsymbol{s}_w\right)$ is the conditional probability of the agent node at location $\boldsymbol{z}_l$ given the online RSS measurements $\boldsymbol{s}_w$. In the absence of a priori knowledge about the location of the agent node, the probabilities of the agent node at the each cell of the radio map are equal. Then by using the Bayes' formula, Eq. (17) can be further transformed into

$$\hat{\boldsymbol{w}} = \arg \max_{\boldsymbol{z}_l, l=1,\ldots,L} P\left(\boldsymbol{s}_w | \boldsymbol{z}_l\right), \tag{18}$$
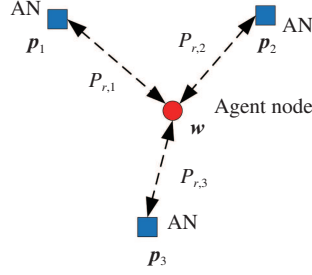
Figure 5 (Color online) Proximity-based localization using RSS measurements, where $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ and $\boldsymbol{p}_3$ denote the ANs, $\boldsymbol{w}$ denotes the location of the targeting agent node, and $P_{r,n}, n = 1, 2, 3$ represent the RSS measurements. The location of the agent node would be determined by comparing the RSSs under certain proximity constraints.

which is known as maximum likelihood (ML) estimation, where $P\left(\boldsymbol{s_w}|\boldsymbol{z}_l\right)$ is the probability of RSS distribution at the given location $\boldsymbol{z}_l$. Therefore, the probabilistic localization methods rely on the estimation of the conditional probability $P\left(\boldsymbol{s_w}|\boldsymbol{z}_l\right)$. There are two main approaches to approximate $P\left(\boldsymbol{s_w}|\boldsymbol{z}_l\right)$, namely, parametric and non-parametric estimation [23, 56]. For parametric estimation methods, the known analytical distribution functions, such as Gaussian, lognormal, and kernel functions, are used to approximate temporal RSS characteristics. However, these parametric estimation methods usually require some probabilistic assumptions (like the probabilistic independence), which makes them challenging to apply in some practical situations. For non-parametric estimation, the fingerprint distributions are proportional to the current centralized histogram, which is known as histogram matching [56]. However, for such methods, a large number of time samples are needed for each cell to generate a histogram.

Compared with geometric-based methods, the main advantage of fingerprinting methods is their robustness to signal measurement errors introduced by multi-path effects. This is because that fingerprinting methods transform localization problems into the problems of pattern matching by dividing the process of localization into offline training and online matching phases. In fingerprinting methods, a location is characterized by its detected signal patterns. Therefore, without having to know the exact locations of ANs, fingerprinting requires neither distance nor angle measurement, rendering it especially feasible in cluttered environments, like indoor or urban scenarios. However, such methods also have some drawbacks. First, for offline training, it is labor intensive and time-consuming to extract the RF fingerprints and construct the radio map. The selected RF fingerprints must uniquely correspond to a given location and should have low variability during a certain time interval. The process of training RF fingerprints is time-consuming since the distribution of RSS fingerprints is usually non-Gaussian, skewed, and multimodal. In addition, since the signal propagation environment is inherently time-varying, the radio map needs to be updated regularly. Furthermore, for online localization, it is necessary to limit the search size and exploit efficient pattern matching algorithms for reducing the cache consumption and computational complexity. Therefore, compared with geometric-based localization, fingerprinting is more suitable for small-size environments.

2.2.3 *Proximity-based localization*

The principle of proximity-based localization is that the location of the agent node is determined according to proximity constraints [59–61]. The mathematical model behind such proximity methods is similar to that of the deterministic approaches in fingerprinting-based localization. As illustrated in Figure 5, such methods usually use RSS measurements to detect the agent node in a dense grid of ANs, and the location of the AN which has the strongest RSS is treated as the location of the agent node. However, different from fingerprinting methods, proximity location estimation depends on the locations of the actual ANs, thus the localization accuracy relies on the density of ANs, which can only be improved by increasing the number of ANs. In general, proximity method is simpler to implement than other localization techniques. However, this method can only provide a very coarse localization service, so it is usually used in systems with low requirements on the location accuracy. The representative application for this method include Cell-ID (CID) [30], RFID [24], and bluetooth-based localization systems [27], which are often used in cellular and IoT networks. Another use case of proximity-based localization is to reduce the search size of fingerprint-based localization before fingerprints pattern matching is performed.

A summary of the aforementioned localization approaches is provided in Table 2, which compares their

Table 2 Summary of different localization techniques

Model	Location-related information	Advantages	Disadvantages
Geometric-based	RSS	Time synchronization is not required; LoS path is not necessary; available for almost all wireless systems	Low accuracy; vulnerable to complex environments
	TOA	High accuracy for LoS scenarios	Time synchronization between ANs and the agent node is necessary; timestamp is required; LoS path is assumed
	TDOA	High accuracy for LoS scenarios; time synchronization between ANs and agent node is not needed; timestamp is not needed	LoS path is assumed
	AOA	Time synchronization is not required; two ANs are sufficient for 2D localization	Directional antennas or arrays are needed; LoS path is assumed in general (some advanced methods can work on NLoS condition, which will be discussed in Subsection 3.2)
Scene analysis (or fingerprinting)	Fingerprints	Accuracy can be very high; robust to clutter environments; LoS path is not required	Training phase is labour intensive and time-consuming; accurate mapping from locations to fingerprints is needed; radio map needs to be updated regularly
Proximity	RSS	Simple and inexpensive; easy to implement	Very low accuracy

differences on measurement models, advantages, and disadvantages.

2.3 Location estimators

In general, there are two categories of location estimators, namely, nonlinear and linear, to solve the localization problems defined in (1). The nonlinear estimators directly solve the problems by minimizing a cost function constructed from (1). Such nonlinear estimators usually result in high localization accuracy. However, sometimes the global solution of such schemes may not be guaranteed since their cost functions are usually multi-modal, and nonlinear estimators usually have high time complexity if grid or random search is involved. By contrast, linear estimators which convert the nonlinear equations into a set of linear equations can find efficient solutions quickly, with degraded localization accuracy compared to nonlinear estimators.

2.3.1 *Nonlinear estimators*

Typical nonlinear estimators include nonlinear least squares (NLS), weighted NLS (WNLS) and ML estimators [35]. Base on the generic model (1), the general cost function of the NLS estimator is defined as [41]

$$V_{\text{NLS}}\left(\boldsymbol{w}\right) = \sum_{n=1}^{N}\left(r_n - h(\boldsymbol{p}_n, \boldsymbol{w})\right)^2$$
$$= \left(\boldsymbol{r} - \boldsymbol{h}\left(\boldsymbol{w}\right)\right)^{\text{T}}\left(\boldsymbol{r} - \boldsymbol{h}\left(\boldsymbol{w}\right)\right), \tag{19}$$

where $\boldsymbol{r} = [r_1, \ldots, r_N]^{\text{T}}$ and $\boldsymbol{h}(\boldsymbol{w}) = [h(\boldsymbol{p}_1, \boldsymbol{w}), \ldots, h(\boldsymbol{p}_N, \boldsymbol{w})]^{\text{T}}$ are N dimensional vectors. The solution of NLS estimator corresponds to the estimated location $\hat{\boldsymbol{w}}$ that minimizes the cost function (19), i.e.,

$$\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} V_{\text{NLS}}(\boldsymbol{w}). \tag{20}$$

The NLS estimator does not rely on any assumption about the error statistics. However, when the covariance of the error vector $\boldsymbol{e} = [e_1, \ldots, e_N]^{\text{T}}$ is available, we can obtain the WNLS estimator, which is defined as [58]

$$\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} V_{\text{WNLS}}(\boldsymbol{w})$$

$$= \arg\min_{\boldsymbol{w}} \left(\boldsymbol{r} - \boldsymbol{h}\left(\boldsymbol{w}\right)\right)^{\mathrm{T}} \boldsymbol{C}^{-1}(\boldsymbol{e}) \left(\boldsymbol{r} - \boldsymbol{h}\left(\boldsymbol{w}\right)\right), \tag{21}$$

where $\boldsymbol{C}(\boldsymbol{e}) = \mathbb{E}\left[\boldsymbol{e}\boldsymbol{e}^{\mathrm{T}}\right]$ denotes the covariance of $\boldsymbol{e}$, and $\mathbb{E}\left[\cdot\right]$ represents the expectation operation. Furthermore, when the error probability distribution $P_e(\boldsymbol{e})$ is known, the ML estimator can be used for location estimation [62,63], i.e.,

$$\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} V_{\mathrm{ML}}(\boldsymbol{w})$$
$$= \arg\min_{\boldsymbol{w}} \log P_e(\boldsymbol{r} - \boldsymbol{h}(\boldsymbol{w})). \tag{22}$$

Note that when the errors satisfy the zero-mean Gaussian distribution, the ML and WNLS estimators have the same performance. In summary, the NLS estimator is simpler than WNLS or ML estimators and can be a practical choice if the noise information is unavailable, while when the error covariance matrix is available, WNLS can perform better than NLS, and the ML is optimal, since it can attain the the Cramer-Rao lower bound (CRLB) [64,65], which is discussed in Subsection 2.5.

In general, there are two ways to solve the optimization problem in (20)–(22). The first one is to perform a global exploration by using grid or random search techniques, such as genetic algorithm [66] or particle swarm optimization [67]. However, although such methods can achieve high localization accuracy, they are time-consuming and the global convergence may not be always guaranteed. The other way is the iterative search algorithm, which requires a good initialization to avoid trapping at the undesired local minima. There are commonly three iterative search schemes, namely, Newton-Raphson [68], Gauss-Newton [69], and steep descent methods [70].

2.3.2 *Linear estimators*

For geometric-based localization, since the function $h(\cdot)$ has an explicit expression given by the algebraic relationships between ANs and the agent node, as discussed in Subsection 2.2, the corresponding localization problems may be also solved in closed-form through linear estimators. The linear estimators mainly include the linear least squares (LLS) and weighted LLS (WLLS). The aim of linearization is to covert the nonlinear equations in (1) into linear forms, based on which the location of the agent node can be estimated through ordinary least squares (LS) techniques. Taking the 2D TOA-based measurement model in (5) as an example, by taking squares on both sides of (5), we have

$$r_{\mathrm{TOA},n}^2 = \|\boldsymbol{p}_n - \boldsymbol{w}\|^2 + e_{\mathrm{TOA},n}^2 + 2e_{\mathrm{TOA},n} \|\boldsymbol{p}_n - \boldsymbol{w}\|, \tag{23}$$

where $n = 1, 2, \ldots, N$. Let $m_{\mathrm{TOA},n} = e_{\mathrm{TOA},n}^2 + 2e_{\mathrm{TOA},n} \|\boldsymbol{p}_n - \boldsymbol{w}\|$ be the error term, and Eq. (23) can be further simplified into

$$\begin{aligned} r_{\mathrm{TOA},n}^2 &= \|\boldsymbol{p}_n - \boldsymbol{w}\|^2 + m_{\mathrm{TOA},n} \\ &= p_{x,n}^2 + p_{y,n}^2 + w_x^2 + w_y^2 - 2w_x p_{x,n} - 2w_y p_{y,n} + m_{\mathrm{TOA},n} \\ &= p_{x,n}^2 + p_{y,n}^2 + \xi - 2w_x p_{x,n} - 2w_y p_{y,n} + m_{\mathrm{TOA},n}, \end{aligned} \tag{24}$$

where $\xi = w_x^2 + w_y^2$ is a dummy variable in the third simplification step. According to [58], with the assumption that the errors are relatively small, we can eliminate $m_{\mathrm{TOA},n}$ and linearize (24) into the following compact form:

$$\boldsymbol{A}\boldsymbol{\theta} = \boldsymbol{b}, \tag{25}$$

where

$$\boldsymbol{A} = \begin{bmatrix} -2p_{x,1} & -2p_{y,1} & 1 \\ \vdots & \vdots & \vdots \\ -2p_{x,N} & -2p_{y,N} & 1 \end{bmatrix}, \quad \boldsymbol{\theta} = [w_x, w_y, \xi]^{\mathrm{T}}, \quad \boldsymbol{b} = \begin{bmatrix} r_{\mathrm{TOA},1}^2 - p_{x,1}^2 - p_{y,1}^2 \\ \vdots \\ r_{\mathrm{TOA},N}^2 - p_{x,N}^2 - p_{y,N}^2 \end{bmatrix}.$$

Then the LS solution of (25) is found by

$$\hat{\boldsymbol{\theta}} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1} \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}, \tag{26}$$

where the first and second entries of $\hat{\boldsymbol{\theta}}$ are the estimated location of the agent node. An alternative way for LLS TOA-based localization is proposed in [71], which subtracts the first equation of (25) from the

Table 3 Comparison of different location estimators

Measurement model	Estimator	Advantage	Disadvantage
Geometric-based; fingerprinting; proximity	NLS	High accuracy; error statistics are not required	Global optimal solution cannot be guaranteed; high complexity
	WNLS	Higher accuracy than NLS	Error covariance is needed; global optimal solution cannot be guaranteed; high complexity
	ML	Highest accuracy compared with other estimators; can achieve the theoretical CRLB	Requires error probability distribution information; global optimal solution cannot be guaranteed; high complexity
Geometric-based	LLS	Closed-form solution is guaranteed; computationally efficient; error statistics are not required	Low accuracy especially for clutter environments
	WLLS	Higher accuracy than LLS; computationally efficient	Error statistics are needed; may require iterative computation

remaining equations. Assuming that the noise is sufficiently small, the linearization form of (5) can be obtained by (25) with

$$
\boldsymbol{A} = \begin{bmatrix} -2(p_{x,2} - p_{x,1}) & -2(p_{y,2} - p_{y,1}) & 1 \\ \vdots & \vdots & \vdots \\ -2(p_{x,N} - p_{x,1}) & -2(p_{y,N} - p_{y,1}) & 1 \end{bmatrix},
$$

and

$$
\boldsymbol{b} = \begin{bmatrix} r_{\text{TOA},2}^2 - r_{\text{TOA},1}^2 + \|\boldsymbol{p}_1\|^2 - \|\boldsymbol{p}_2\|^2 \\ \vdots \\ r_{\text{TOA},N}^2 - r_{\text{TOA},1}^2 + \|\boldsymbol{p}_1\|^2 - \|\boldsymbol{p}_N\|^2 \end{bmatrix}.
$$

Similar to TOA, the equivalent closed-form solutions for TDOA [72,73], AOA [74,75], and RSS [76] based on localization can also be obtained. Although the LLS estimators provide a closed-form solution, the solutions are sub-optimal in general, due to the discarding of information in the linearization process, and it only performs well when the noise is relatively small. The WLLS estimator [77] is more generic than LLS scheme by making the use of the mean and covariance information of the measurement errors, which can provide higher localization accuracy.

A comparison of different location estimators, in terms of their advantages, drawbacks, and their corresponding measurement models, is provided in Table 3.

2.4 Main sources of error and mitigation techniques

Localization performance is fundamentally limited by various estimation biases and measurement errors. Here, we discuss three main sources of error, together with their corresponding mitigation techniques.

2.4.1 *Multi-path fading*

Multi-path fading commonly exists in wireless channels, which can considerably degrade the localization performance, especially for indoor localization scenarios. In particular, for narrowband localization systems in clutter environments, the signals that arrive at the receiver via different paths are superimposed with each other, resulting them unresolvable at the receiver. Moreover, the multi-path effect varies with signal propagation environments, making the signal detection more difficult. To mitigate this effect, some diversity combining techniques are proposed, and for the ultrawide bandwidth (UWB) systems, the multi-path components are usually resolvable temporally without resorting to complex algorithms [78,79]. However, in harsh environments, a large number of multi-path components still degrade the localization performance, especially for the geometric-based algorithms which need to distinguish the LoS path from a large number of NLoS paths to obtain the location information of the agent node. On the other hand, multi-path effect may also be exploited to achieve cm-level localization accuracy in indoor environment, via the technique called time-reversal (TR) [80,81]. TR is a transmission technique that leverages the multi-path channel as a matched filter and focuses the energy of transmitted signals onto the intended location. The authors proposed a TR-based fingerprinting localization method in [80], which can achieve

10 cm localization accuracy in indoor environment, where the channel impulse responses in different locations were stored as fingerprints, and a TR-based method was used to capture the features of the fingerprints. In [81], a TR-based indoor tracking system with decimeter level accuracy was proposed, where the TR-base method was used for distance estimation, combined with the IMUs for angle estimation.

2.4.2 *NLoS propagation*

The adversary impact of NLoS propagation lies in that the received NLoS signals weaken the correlation between signal measurements and link distance, since it will introduce a positive bias to the range estimate. In general, there are three methods to cope with the NLoS condition [35]. The first method is based on the statistical information of the NLoS error. By assuming a scattering model of the environment, the statistics of signal measurements can be obtained, and then the well-known techniques, like MAP or ML, can be used to mitigate the effects of NLoS errors. However, the difficulty of such methods is to obtain an accurate model, which may change with terrain and/or the construction/demolition of buildings. The second method uses all NLoS and LoS measurements with appropriate weights to minimize the effects of the NLoS contributions, where the weights are generated from the localization geometry and the ANs layout. Although this method is effective even in the cases without LoS measurements, its solution is unreliable because the NLoS errors are always present. The third method is to identify and discard those NLoS measurements, and perform localization only based on the LoS measurements. In essence, the problem of NLoS identification in this method is converted into a statistical detection problem, where the NLoS and LoS conditions are considered as two hypotheses, and the goal of the problem is to figure out a metric to differentiate the NLoS and LoS hypotheses. For instance, we can identify the NLoS path based on the statistics of range measurements. Usually, the NLoS range measurements which are positively biased with non-Gaussian distribution tend to have a larger variance compared with the LoS counterpart with Gaussian distribution. However, in some harsh environments, almost all measurements come from NLoS paths, so there are insufficient LoS measurements for localization. For such cases, the localization methods using NLoS measurements and geometrical information are proposed [35]. In general, these NLoS localization techniques can be divided into two categories. One is NLoS localization using signal measurements combining with the priori knowledge of the environment map. The other is the localization using the measurements from scatters [13, 82], which turns the multi-path "from foe to friend". Note that in the latter method, the NLoS measurements are first identified, and then the geometrical relationship among the ANs, the agent node, and the scatters is used to locate the agent node.

2.4.3 *Systematic error*

Systematic errors refer to those errors that originate from the localization system itself, such as the imperfect signal measurements and radio miscalibration. For instance, in time-based localization systems, the ANs are equipped with the oscillators for time synchronization. However, the oscillators often experience independent frequency drifts, resulting in clock drift and offset, that may degrade the localization accuracy. Systematic errors often bias the location estimators, making the mean of the estimator larger than the true value. These errors are usually constant with respect to the targeting location and cannot be eliminated by averaging over multiple repeated measurements. Nevertheless, some techniques can effectively mitigate these errors. For example, real-time infrastructure calibration can mitigate the localization performance degradation, where wireless links among ANs are made periodically to calibrate the parameters of the localization system. Alternatively, some techniques like clock offset correction and recursive Bayesian approach have been proposed to tackle systematic errors [83, 84].

2.5 Performance metrics

The performance of localization systems can be evaluated from various aspects. In this subsection, we outline the main performance evaluation metrics, including accuracy, precision, complexity, coverage and scalability, while a performance comparison across several localization infrastructures is given in Subsection 2.6.

2.5.1 *Accuracy*

Accuracy (or location error) is usually measured as the Euclidean distance between the estimated location $\hat{\boldsymbol{w}}$ and the actual location $\boldsymbol{w}$ of the agent node, which is typically the most important performance metric

to analyze the overall system performance. In practice, various statistics can be adopted for this evaluation criterion, such as the mean square error (MSE) of the location estimates, which is defined as

$$
\begin{aligned}
e_{\mathrm{MSE}}(\hat{\boldsymbol{w}}) &= \mathbb{E}\left[\|\hat{\boldsymbol{w}} - \boldsymbol{w}\|^2\right] \\
&= \mathrm{tr}\left(\boldsymbol{C}\left(\hat{\boldsymbol{w}}, \boldsymbol{w}\right)\right) + \|\mathbb{E}\left(\hat{\boldsymbol{w}}\right) - \boldsymbol{w}\|^2,
\end{aligned}
\tag{27}
$$

where $\mathrm{tr}(\cdot)$ indicates the matrix trace, and $\boldsymbol{C}\left(\hat{\boldsymbol{w}}, \boldsymbol{w}\right)$ denotes the covariance matrix of $\hat{\boldsymbol{w}}$ and $\boldsymbol{w}$, which is defined as

$$
\boldsymbol{C}\left(\hat{\boldsymbol{w}}, \boldsymbol{w}\right) = \mathbb{E}\left[(\hat{\boldsymbol{w}} - \boldsymbol{w})(\hat{\boldsymbol{w}} - \boldsymbol{w})^{\mathrm{T}}\right].
\tag{28}
$$

The first and second terms of (27) represent the variance and bias of the estimated location, respectively. Note that the bias is usually a constant unknown error introduced by the signal measurement, which can be mitigated through appropriate methods. For unbiased cases, the CRLB gives the lower bound on the variance of $\hat{\boldsymbol{w}}$ [64, 65], i.e.,

$$
\boldsymbol{C}\left(\hat{\boldsymbol{w}}, \boldsymbol{w}\right) \succeq \boldsymbol{I}^{-1}\left(\boldsymbol{w}\right),
\tag{29}
$$

where $\succeq$ denotes that the matrix $\boldsymbol{C}\left(\hat{\boldsymbol{w}}, \boldsymbol{w}\right) - \boldsymbol{I}^{-1}\left(\boldsymbol{w}\right)$ is positive semidefinite, and $\boldsymbol{I}(\boldsymbol{w})$ is the Fisher information matrix (FIM) of $\boldsymbol{w}$, given by [41]

$$
\boldsymbol{I}\left(\boldsymbol{w}\right) = -\mathbb{E}\left[\frac{\partial^2 \ln P\left(\boldsymbol{r} \mid \boldsymbol{w}\right)}{\partial \boldsymbol{w} \partial \boldsymbol{w}^{\mathrm{T}}}\right],
\tag{30}
$$

where $P\left(\boldsymbol{r} \mid \boldsymbol{w}\right)$ denotes the conditional probability density function (PDF) of the measurement vector $\boldsymbol{r}$. Taking the 2D TOA-based localization as an example, the measurement vector can be generated by (5), which is a vector of the measured distances between the agent node and each of ANs, denoted by

$$
\boldsymbol{r} = \hat{\boldsymbol{d}} = \left[\hat{d}_1, \hat{d}_2, \ldots, \hat{d}_N\right]^{\mathrm{T}}.
\tag{31}
$$

For convenience, here we consider the measurement errors $e_{\mathrm{TOA},n}, n = 1, \ldots, N$ that are zero-mean Gaussian distributed, and the conditional PDF of measured distance $\hat{\boldsymbol{d}}$ is

$$
P\left(\hat{\boldsymbol{d}} \mid \boldsymbol{w}\right) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{(\hat{d}_n - d_n)^2}{2\sigma_n^2}\right),
\tag{32}
$$

where $d_n, n = 1, \ldots, N$ denote the actual distances between each AN and the agent node. By substituting (32) into (30), the FIM can be calculated as [63]

$$
\boldsymbol{I}\left(\boldsymbol{w}\right) = \begin{bmatrix} \displaystyle\sum_{i=1}^{N} \frac{(w_x - p_{x,i})^2}{\sigma_i^2 d_i^2} & \displaystyle\sum_{i=1}^{N} \frac{(w_x - p_{x,i})\left(w_y - p_{y,i}\right)}{\sigma_i^2 d_i^2} \\ \displaystyle\sum_{i=1}^{N} \frac{(w_x - p_{x,i})\left(w_y - p_{y,i}\right)}{\sigma_i^2 d_i^2} & \displaystyle\sum_{i=1}^{N} \frac{\left(w_y - p_{y,i}\right)^2}{\sigma_i^2 d_i^2} \end{bmatrix}.
\tag{33}
$$

Then substituting (33) into (29), the CRLB for the TOA-based localization method can be obtained. In a similar manner, the FIMs for TDOA, RSS, and AOA measurements can also be obtained. The CRLB implies that the MSE of location estimates satisfies the following bound:

$$
\begin{aligned}
e_{\mathrm{MSE}}(\hat{\boldsymbol{w}}) &= \mathbb{E}\left[\|\hat{\boldsymbol{w}} - \boldsymbol{w}\|^2\right] \\
&\geqslant \mathrm{tr}\left(\boldsymbol{C}\left(\hat{\boldsymbol{w}}, \boldsymbol{w}\right)\right) \geqslant \mathrm{tr}\left(\boldsymbol{I}^{-1}\left(\boldsymbol{w}\right)\right).
\end{aligned}
\tag{34}
$$

Another useful evaluation criterion is the root mean square error (RMSE), which is the root of the MSE with the following bound [28]:

$$
\begin{aligned}
e_{\mathrm{RMSE}}(\hat{\boldsymbol{w}}) &= \sqrt{\mathbb{E}\left[\|\hat{\boldsymbol{w}} - \boldsymbol{w}\|^2\right]} \\
&\geqslant \sqrt{\mathrm{tr}\left(\mathrm{Cov}\left(\hat{\boldsymbol{w}}\right)\right)} \geqslant \sqrt{\mathrm{tr}\left(\boldsymbol{I}^{-1}\left(\boldsymbol{w}\right)\right)}.
\end{aligned}
\tag{35}
$$

The CRLB determines the attainable location accuracy of the unbiased system. However, many practical estimators are biased because of signal NLoS propagation and other factors, so in practice the system performance may not achieve the CRLB. Other bounds like the Bayesian Cramer-Rao bound [85], Weiss-Weinstein bound [86], and extended Zik-Zakai bound [87] are tighter but require more complicated evaluations compared with CRLB.

2.5.2 *Precision*

Precision reveals the variation of location estimation with respect to the localization accuracy [23]. Specifically, precision measures the statistical characterization of the accuracy which varies over many localization trials. In some studies, the geometrical dilution of precision (GDOP) is also used to measure the variation of localization errors [23]. Taking the TOA-based localization as an example, the GDOP is defined as [41]

$$\text{GDOP} = \frac{e_{\text{RMSE}}(\hat{\boldsymbol{w}})}{e_{\text{RMSE}}(\hat{\boldsymbol{d}})}, \tag{36}$$

where the numerator and denominator are the RMSE of the location estimate and the range estimate, respectively. The smaller GDOP value means the better performance on localization precision. In addition, the GDOP also reveals the relation between the achievable localization accuracy and the geometry distribution of the ANs, which can be adopted as a criterion for ANs placement and selection to minimize the GDOP value.

Another evaluation metric is the localization error outage (LEO), which is defined as the probability when the localization error exceeds a certain threshold e_{th} [26], i.e.,

$$\text{LEO}(e_{\text{th}}) = \Pr\left\{\|\hat{\boldsymbol{w}} - \boldsymbol{w}\| \geqslant e_{\text{th}}\right\}. \tag{37}$$

An equivalent expression of (37) is the cumulative distribution function (CDF) of the localization error defined by

$$\text{CDF}(e_{\text{th}}) = 1 - \text{LEO}(e_{\text{th}}), \tag{38}$$

which denotes the success probability of location estimations with respect to a predefined accuracy. In practice, the LEO or CDF reveals the probability of confidence in the location estimate. When the accuracies of two localization algorithms are the same, the algorithm that gives lower LEO or higher CDF values has better precision [26]. For example, a localization system with $\text{CDF}(1.5) = 0.9$ (a precision of 90% within 1.5 m) performs better than that with $\text{CDF}(1.5) = 0.5$ (a precision of 50% within 1.5 m).

2.5.3 *Complexity*

The complexity of a localization system depends on the hardware, process of signal measurement, and computational complexity of the localization algorithm [23]. In general, it is difficult to analytically derive the complex formula of different localization techniques. Therefore, the computational complexity of the location estimators is usually treated as the complexity of the localization system. There is always a trade-off between accuracy and complexity in the sense that more accurate localization usually requires higher computational complexity. On the other hand, the location update rate or latency can be also used as a criterion to evaluate the system complexity, which reflects the time delay between two consecutive location updates for the same agent node and is very important for navigation.

2.5.4 *Coverage and scalability*

In general, the localization performance degrades as the distance between each AN and the agent node increases. The coverage refers to the maximum area where the localization system can provide effective localization services with guaranteed performance in terms of accuracy, precision, latency, and so on. In general, the coverage can be roughly classified into global, local, and indoor coverage depending on different localization infrastructures. On the other hand, as the localization coverage increases, the wireless channels may become congested and the localization system needs to perform more signal measurement and calculation operations, so the scalability reflects the adaptive capability of the localization system when the localization scope gets large [23].

Table 4 Comparison of different localization infrastructures

Localization infrastructures	Example	Techniques	Type	Performance
GNSS	GPS	OW-TOA/TDOA	Trilateration	10–20 m accuracy; global coverage; about 30 s latency
Cellular networks	2G	CID + TA	Proximity	About 550 m accuracy; local area coverage; low latency
		E-OTD	Trilateration	50–300 m accuracy; local area coverage; medium latency
	3G	OTDOA/UTDOA	Trilateration	50–200 m accuracy; local area coverage; medium latency
		RFPM	Scene analysis	Over 50 m accuracy; local area coverage
		A-GPS	Trilateration	10–50 m accuracy; global coverage; high latency
	4G	E-CID	Angulation + Proximity	About 150 m accuracy; local area coverage; low latency
		OTDOA/UTDOA	Trilateration	25–200 m accuracy; local area coverage; medium latency
		A-GNSS	Trilateration	Less than 10 m accuracy; global coverage; high latency
WiFi	IEEE 802.11	RFPM	Scene analysis	1–5 m accuracy; 50–100 m coverage in general; medium latency
UWB	IEEE 802.15.4a	TOA/TDOA	Trilateration	0.1–1 m accuracy; indoor area coverage; medium latency
VLP	IEEE 802.15.7	RSS/TOA/TDOA/AOA	Trilateration/triangulation/ fingerprinting/promixty	cm-level accuracy; medium latency; indoor area coverage

2.6 Localization infrastructures

There are two basic approaches to deploy the localization systems. The first one is to build a dedicated localization infrastructure, like the GNSS. The second approach is to reuse the existing wireless network infrastructures with the signals of opportunity (SoOP), like cellular networks, WLAN, to provide wireless localization services, in addition to communication services. For the first approach, the main advantage is that it can achieve high localization performance by using specific reference signal and professional hardware, while the drawback is the cost of the hardware and the limitation on the system scalability. For the second approach, it avoids the expensive and time-consuming deployment of infrastructure, but such systems usually rely on sophisticated algorithms to improve the performance. In this subsection, we mainly discuss the most popular localization infrastructures, including GNSS, cellular networks, WiFi, and UWB based on localization systems, and a detailed comparison on their performance is given in Table 4.

2.6.1 *GNSS*

Several countries in the world have already developed and launched various GNNSs, including the GPS by the U.S.A, the Galileo by Europe, the Beidou by China, the GLONASS by Russia, as well as other regional systems like the Japanese Quasi-Zenith satellite system (QZSS) and the Indian regional navigational satellite system (IRNNS). Although all of these GNNSs introduce different innovations at the system and signal levels, they share common theoretical and functional principles. Taking the most popular GPS as an example, it is a self-localization system based on a man-made constellation of 27 earth-orbiting

satellites. The localization techniques behind GPS are the OW-TOA or TDOA methods, where at least four clearly visible satellites are required for the agent node to localize itself in terms of latitude, longitude, and altitude. In general, the localization accuracy of GPS ranges from 10 to 20 m [88], and the precision performance is highly dependent on the geometric distribution and selection of satellites, which can be measured by the GDOP value. The stand-alone GPS suffers from the problem of time to first fix (TTFF) or cold start; that is, when a GPS receiver is first turned on, it needs a long time (about 30 s or even a few minutes) to acquire the satellite signals. To this end, an assisted-GPS (A-GPS) technique has been developed [88,89], which uses a location server equipped with the GPS receiver that can simultaneously detect the satellites to help the user equipment (UE) to acquire GPS signals more quickly. Although GPS can provide localization services globally, its performance is degraded in severe scattering environments, like indoor or urban areas.

2.6.2 *Cellular networks*

Initially, cellular networks were designed for communications, and all UEs in the networks that need location services resorted to the GNSS. However, due to the poor performance of the GNSS in urban and indoor environments, the cellular-based localization was proposed as a good complementary of GNSS to enhance its performance and robustness. The first cellular-based localization system is the E-911 [90] introduced by the Federal Communication Commission (FCC) of the U.S.A to provide emergency services. Moreover, since the location information within the cellular networks can be exploited for commercial services and network optimization, in the subsequent cellular networks standards from 2G to 4G, more advanced cellular-based localization techniques were proposed. For cellular networks, both the uplink and downlink between the BSs and UEs can be exploited for localization. In 2G cellular standard, the localization schemes in global system for mobile communications (GSM) and code-division multiple access (CDMA) network include CID, timing advance (TA), and enhanced observed time difference (E-OTD) [31,91–93], which are mainly based on the uplink time-based measurements that required strict time synchronization, leading to coarse localization accuracy, ranging from 50 to 550 m. In 3G networks, the specific location measurement units (LMUs) were introduced in wideband CDMA (WCDMA) systems, which can be integrated with the BSs to improve the signal measurement performance. The uplink localization methods in 3G networks are similar to those in 2G networks, where the enhanced CID (E-CID) and uplink TDOA (UTDOA) are the improved versions of CID and E-OTD, respectively [30]. Different from 2G localization, the localization method based on the downlink signals was specified in 3G networks, i.e., observed TDOA (OTDOA). The OTDOA information is measured as the reference signal time difference (RSTD) by the UE, which then reports the RSTD to cellular networks to calculate its location. Moreover, an RF pattern matching (RFPM) method was proposed in the 3GPP Release 10 [94]. In addition, the A-GPS method is adopted by 3G networks, which can provide location services in GPS-denied areas. In general, the localization accuracy of 3G networks ranges from 50 to 200 m [95–97]. The location services in 4G long term evolution (LTE) were firstly defined in Release 9 [98]. Different from 2G or 3G networks, the 4G networks specify the LTE positioning protocol (LPP) [99] to exchange information between UE and the remote location server. Furthermore, a dedicated positioning reference signal (PRS) was introduced in LTE standards, which has high configurability in terms of power, time, and frequency allocation, and can improve localization performance [100]. In the forthcoming 5G networks, the location information is becoming more critical for various applications [18], including content prefetching, radio environment mapping, proactive radio resource management (RRM). All of the prospective applications require the localization services with higher accuracy. Moreover, the availability, scalability, security, as well as privacy are also new challenges for localization systems [24]. Fortunately, the 5G communication systems enabled by higher carrier frequencies, wider signal bandwidths, denser networks, and massive MIMO technologies can bring new opportunities for localization, which will be discussed in Section 3.

2.6.3 *WiFi*

WiFi technology can be used as a promising indoor localization scheme, thanks to its ubiquitous availability and handy for short-range RSS measurements. It usually operates in two unlicensed bands, i.e., 2.4 GHz (IEEE 802.11b/g) and 5 GHz (IEEE 802.11a), with a range of 50–100 m in general [55], and has now increased to about 1 km in IEEE 802.11ah [101]. The main advantage of WiFi-based localization is its almost ubiquitous availability since most smart devices today are WiFi enabled. However, WiFi signals transmit on the unlicensed industrial, scientific and medical (ISM) band which are vulnerable

to interference [55]. The most commonly known WiFi-based localization approach is the RSS-based fingerprinting, while TOA-, TDOA- and AOA-based methods are relatively less used since angular and time delay measurements are complex. In general, the accuracy of typical WiFi localization systems is approximately 1–5 m with a few seconds update rate [102–104]. Recently, the accuracy of WiFi-based localization systems has achieved decimeter-level in certain scenarios [105, 106]. There are many survey papers providing reviews about WiFi-based localization [54, 55, 104]. In [55], the authors provided a comprehensive overview about WiFi-based indoor localization techniques, while in [54] the authors focused on the outdoor fingerprinting-based localization with WiFi signals. In [104], the localization methods with the use of available measurements performed on smartphone are reviewed.

2.6.4 *UWB*

The UWB signal refers to the signal whose spectrum is larger than 500 MHz in the frequency range from 3.1 to 10.6 GHz [107, 108]. In general, the UWB spectrum can be acquired either by generating a series of extremely short duration pulses less than 1 ns or by aggregating a number of narrowband subcarriers. The first UWB standard is IEEE 802.15.4a, which was designed for low-rate wireless personal area networks (WPAN) in short range [109]. Different from conventional narrowband signals, UWB signals have higher temporal resolutions, lower transmission power consumption, and the ability to resolve multipath and penetrate obstacles, which makes it quite promising to provide centimeter-level localization services [110, 111]. In general, the localization accuracy of UWB-based systems ranges from 0.1 to 1 m, but the main drawback is its short coverage, which renders it most suitable for indoor environments [110, 111]. The UWB technology can be implemented on self- or remote localization systems, and it can be incorporated with different localization techniques to improve the performance. For fingerprinting localization, UWB fingerprinting enables the small ambiguity region even with a single AN, so that high localization accuracy can be achieved for both LoS and NLoS scenarios. For geometric-based localization, due to the high temporal resolution of UWB signals, the time-based localization methods can achieve very high accuracy [109, 110]. However, UWB is not suitable for RSS- or AOA-based localization methods, since the RSS and AOA measurements cannot benefit from the huge bandwidth of UWB.

2.6.5 *Visible light positioning (VLP)*

Visible light communication (VLC) is an emerging technology for high speed transmission, which uses light devices and optical sensors to transmit data. The first well-known standard of VLC is IEEE 802.15.7, which defines the physical layer protocols for optical signal transmission [112]. On the other hand, visible light can also be used for localization, which is referred to as VLP [113]. Different from RF-based wireless localization, VLP can easily achieve centimeter-level localization accuracy. However, the main drawback of VLP systems is their limited coverage, which makes them mostly suitable for indoor scenarios. In [46], the authors proposed a TDOA-based localization method by using visible LEDs light, which can achieve localization accuracy of less than 1 cm in an indoor space. In [114], an LED-based fingerprinting localization method was proposed, which can achieve 0.81 m accuracy in a 30 m × 30 m 2D indoor space. In [115], the authors proposed an angle-based localization method, which measures the angle difference of the arrival optical signal transmitted to a receiver for localization, and it can achieve 3.2 cm accuracy with time cost of 0.36 s.

2.7 Advanced localization techniques

In the above subsections, we mainly focus on the static localization problems for one agent node. In this subsection, we discuss some advanced techniques for localization, including tracking, simultaneous localization and mapping (SLAM), cooperative localization, and data fusion.

2.7.1 *Tracking*

The problems of tracking can be viewed as a sequence of independent localization problems, but in a more general sense, besides location estimation, it further involves the estimation of velocity, acceleration, and all past states of the mobile agent node. Compared with static localization, tracking entails mobility modeling to describe the agent node's movement. Mathematically, the problems of tracking can be formulated as

$$\boldsymbol{x}_t = \zeta(\boldsymbol{x}_{t-1}) + \boldsymbol{z}_t, \tag{39a}$$

$$\boldsymbol{y}_t = h(\boldsymbol{x}_t) + \boldsymbol{v}_t, \tag{39b}$$

where $\boldsymbol{x}_t$ denotes the global state of the agent node at time step t, including its location, velocity, acceleration, while $\boldsymbol{y}_t$ is the measurement at step t; the function $\zeta(\cdot)$ models dynamics of the agent node, and the function $h(\cdot)$ is the measurement model; $\boldsymbol{z}_t$ and $\boldsymbol{v}_t$ represent the additive random noise of the state dynamics and measurement, respectively. The goal of tracking is to obtain a time succession of the agent node's states $\boldsymbol{x} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T\}$ from a set of measurements $\boldsymbol{y} = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T\}$. Therefore, the tracking problem can be viewed as an estimation problem. To operate in real-time, various filtering-based techniques have been proposed, such as the Bayesian filtering methodology, like Kalman filer (KF) [83], extended KF (EKF), and particle filter (PF) [84]. For a more comprehensive review on tracking, readers may refer to [26, 27].

2.7.2 SLAM

Different from the localization problems, for SLAM, the locations of the ANs may not always be prior known, and the goal is to locate a set of fixed ANs and construct a map of the surrounding environment when a mobile node navigates through a predetermined path. Typically, for the simplest SLAM problem, only one mobile node performs the environment surveying and locates the ANs, while for more sophisticated cases cooperative SLAM is involved. Different from tracking, for SLAM, the navigation scheme of the mobile node may also affect its state, thus the dynamics equation in (39a) can be revised as [26]

$$\boldsymbol{x}_t = \zeta(\boldsymbol{x}_{t-1}) + \boldsymbol{u}_t + \boldsymbol{z}_t, \tag{40}$$

where $\boldsymbol{u}_t$ is a control parameter used to guide the mobile node following the predetermined path. Note that since the location of ANs is unknown, the state $\boldsymbol{x}_t$ should contain the mobile's state for the classical tracking as well as the ANs' locations. In practice, the mathematical tools to solve the tracking problems are also suitable for the SLAM problem, especially the EKF which is widely used for SLAM. More comprehensive overviews on this topic are given in [116–118].

2.7.3 Cooperative localization

In many application scenarios, due to the presence of NLoS propagation, some agent nodes are difficult to directly communicate with a sufficient number of ANs for localization, which may degrade the localization accuracy. To this end, the cooperative localization techniques have been proposed [119–121], which can improve the localization performance, particularly in complex environments. For noncooperative localization, all agent nodes need to communicate with ANs, thus a high density of ANs or long-range ANs transmission coverage is required. Compared with noncooperative localization, cooperative localization allows the inter-communication among agent nodes, and the agent node can obtain information from both ANs and other agent nodes, so cooperative localization can not only improve accuracy but also extend the localization coverage. The cooperative localization is also a parameters estimation problem, which can be solved through two kinds of methods. One is the deterministic method, which includes the classical LS, multidimensional scaling, multilateration [122–124]. However, such methods rely on the assumption of a Gaussian model for all measurement uncertainties, which may not be effective in some practical scenarios. The other category is the probabilistic method, which is known as belief propagation (BP) [121, 125]. Such methods can not only obtain the location estimations but also measure the uncertainty of these estimates. A detailed fundamental analysis about cooperative localization can be found in [64], where the equivalent Fisher information (EFI) for cooperative networks was derived. In [121], the authors discussed the main cooperative localization approaches from the perspective of estimation theory and factor graphs.

2.7.4 Data fusion

Due to the multi-path effects, the performance of localization systems that only use a single type of measurement is severely limited in complex environments. To this end, the study on fusing different types of information to improve the localization performance has gained momentum [83,84,126,127]. So far, the wireless network has become heterogeneous with various wireless technologies, such as cellular networks, WLAN, RFID, and Bluetooth. Therefore, hybrid data fusion (HDF) has attracted research interest for unlocking the full potential of localization systems [27]. An example of data fusion localization is the SiRstarV [27], which combines real-time data from GNSS satellites, WLAN, cellular, as well as multiple IMU sensors to improve the localization performance. Numerical results show that its positioning error is

within 9 m for 68% cases and 13.1 m for 95% cases over several tests [128]. A generic framework of fusion technologies used for tracking is provided in [126], where fusion exists across all stages of localization systems. In the signal measurement stage, different signal measurements like TOA, TDOA, AOA, and RSS are combined. In the position estimation stage, different localization techniques such as trilateration, triangulation, and fingerprinting are performed. Finally, a temporal filter like EKF is applied to smooth the estimated user trajectory with the help of information gathered from IMU sensors.

3 Integrated localization and communication for 5G and beyond

The future mobile communication networks are expected to realize the vision of Internet of Everything (IoE) with versatile networks not only for ubiquitous communications, but also for seamless localization and intelligent automatic control with high accuracy [3,129]. To fulfill this magnificent goal, it is necessary to develop high performance localization techniques, so as to not only meet the requirements of various emerging commercial and industrial location-based services, but also to improve the communication performance in various aspects at different network layers. In this section, we first overview the recent 3GPP standardizations on localization technologies in 5G NR from Release 15 to Release 17, and discuss the developing trends of the localization systems. Then we discuss the enabling technologies of 5G networks towards centimeter-level localization. Since the accurate location information is beneficial for communications, we overview the location-aware communication techniques in different network layers. After that, due to the correlation between communication and localization, we discuss the co-design of localization and communication systems and attempt to give some insights for future network design. Finally, as a promising vision of 6G, we discuss the development trends of localization and communication techniques in future aerial-and-ground integrated networks.

3.1 3GPP standardizations for 5G localization

Compared with 4G-LTE, 5G networks enabled by higher carrier frequencies, wider bandwidth, and massive antenna arrays are expected to achieve enhanced localization performance in terms of accuracy, reliability, coverage and latency [130]. In general, the localization accuracy of LTE is between 25 and 200 m, while that of 5G networks is expected to be on the order of submeter-level or even cm-level, as reported by 3GPP [1]. In Release 15, a general description of location services (LCS) and the corresponding requirements are given in TS 22.071 [1]. Release 15 specifies the CID and radio access technology (RAT)—independent positioning methods by reusing LPP [99], but the RAT-dependent positioning methods are excluded. In Release 16, the recent new use cases of localization are identified in TR 22.872 [131], which mainly includes wearables, advertisement push, flow monitoring and control, as well as an emergency call. The accuracy requirements for such applications range from 1 to 3 m in indoor scenarios, and below 50 m horizontal and 3 m vertical in outdoor scenarios, with less than 10 s TTFF and 1 s latency. Such requirements cannot be satisfied with the current cellular networks, e.g., LTE, so a study item was concluded in March 2019 to investigate the NR positioning support, and the technical report is summarized in TR 38.855 of Release 16 [132]. In Release 16, the regulatory requirements on positioning are listed as follows.

- Horizontal positioning error is less than 50 m for 80% of UEs.
- Vertical positioning error is less than 5 m for 80% of UEs.
- E2E latency and TTFF are less than 30 s.

These are regraded as the minimum performance targets for NR positioning studies. Furthermore, for commercial use cases in indoor and outdoor scenarios, the localization technologies should meet the following requirements.

- For indoor scenarios, the horizontal and vertical positioning error are less than 3 m for 80% of UEs.
- For outdoor scenarios, the horizontal positioning error is less than 10 m, and vertical positioning error is less than 3 m for 80% of UEs.
- E2E latency is less than 1 s for both indoor and outdoor scenarios.

In order to fulfill the above requirements, Release 16 recommends the following RAT-dependent localization methods, including downlink TDOA (DL-TDOA), downlink AOD (DL-AOD), uplink TDOA (UL-TDOA), uplink AOA (UL-AOA), multi-cell RTT (Multi-RTT), and E-CID. In addition, the combination of RAT-dependent and RAT-independent techniques like GNSS, Bluetooth, WLAN, and sensors is also considered for NR positioning. The simulation results of these proposed NR positioning methods

are presented in [132], which considers three different scenarios, i.e., urban macro (UMa), urban micro (UMi), and indoor office (InH). The simulation results show that DL-TDOA can meet the regulatory requirements in all scenarios, while under some specific evaluation assumptions, some other techniques can meet the requirements of commercial performance [130]. To satisfy the increasing requirements on localization accuracy resulting from new applications and industrial IoT (IIoT) use cases, the studies on NR positioning enhancements are ongoing in Release 17 (TS 38.857). As the recommendation of TS 38.857, the NR positioning in Release 17 should meet the following exemplary performance targets.

- For general commercial use cases (e.g., TS 22.261 [133]), the submeter-level positioning accuracy should be guaranteed.
- For IIoT use cases (e.g., TR 22.804 [134]), positioning error should be less than 0.2 m.
- The latency requirement is less than 100 ms generally, while for some IIoT use cases, the 10 ms latency is desired.

Two main goals of TS 38.857 are: (i) study enhancements to support high accuracy, low latency, network efficiency (or scalability), and device efficiency; (ii) study solutions to support integrity and reliability of assistance data and position information. Furthermore, the enhancements on the combination of diverse positioning techniques and the flexibility of the networks are further emphasized. To meet the requirements of the emerging autonomous applications, the future localization technology should not only guarantee the high accuracy, but also be able to determine the reliability and uncertainty or confidence level of the location-related data.

3.2 Towards centimeter localization for 5G and beyond

Massive MIMO, mmWave, UDNs and device-to-device (D2D) communication are four underlying technologies of 5G networks, which can not only improve the communication performance, but also potentially benefit localization [18]. For massive MIMO, the BS equipped with a large antenna array can steer highly directional beams and provide high angular resolution, which can be utilized for angle-based localization [135]. The mmWave technology can provide large bandwidth on the order of GHz, which offers high temporal resolution, thereby ensuring more accurate time-based localization [136]. In addition, since mmWave signals suffer from high path loss than the sub-6 GHz counterparts, the cell size in 5G networks will shrink, which renders UDN a promising technology for 5G networks. The high density of BSs in UDN will increase the LoS probability, which may also improve the localization performance. On the other hand, the promising D2D communication, where two neighboring devices directly communicate with each other while bypassing the BS, may flourish the device-centric cooperative localization.

3.2.1 *mmWave massive MIMO localization*

The main advantage of massive MIMO is its unprecedented potential of high spectral efficiency [137–139]. In massive MIMO systems, the BS equipped with a large number of antennas can serve a large number of UEs simultaneously with high data rates in the same frequency through dense spatial multiplexing. This is achieved via beamforming/precoding by large antenna arrays with high angular resolution, where the channels among different UEs are asymptotically orthogonal [137, 140]. In general, the large antenna array can be deployed as a one-dimensional linear array or two-dimensional planar array, e.g., uniform linear arrays (ULAs) [141] versus uniform rectangular arrays (URAs) [142]. For ULA with N antenna elements, the beam can steer in 1D angles θ, and the unit-norm array response vector is [141]

$$\boldsymbol{a}(\theta) = \frac{1}{\sqrt{N}} \left[1, \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}d\sin(\theta)}, \dots, \mathrm{e}^{-\mathrm{j}(N-1)\frac{2\pi}{\lambda}d\sin(\theta)} \right]^{\mathrm{T}}, \tag{41}$$

where λ is the signal wavelength and d is the distance between the adjacent antenna elements. For URA with $N = N_x \times N_y$ antenna elements, the operation of beamforming lies in 2D angles referred to azimuth and elevation angles (θ, ϕ), and the unit-norm array response vector is [142]

$$\begin{aligned} &[\boldsymbol{a}(\theta, \phi)]_{n_x, n_y} = \frac{1}{\sqrt{N}} \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}d\sin(\phi)[(n_x-1)\cos(\theta)+(n_y-1)\sin(\theta)]}, \\ &n_x \in \{1, \dots, N_x\}, \quad n_y \in \{1, \dots, N_y\}. \end{aligned} \tag{42}$$

Another promising technology of 5G networks is mmWave communication, which can achieve high data rate with low latency, due to its availability of the large bandwidth on the order of GHz [143]. In
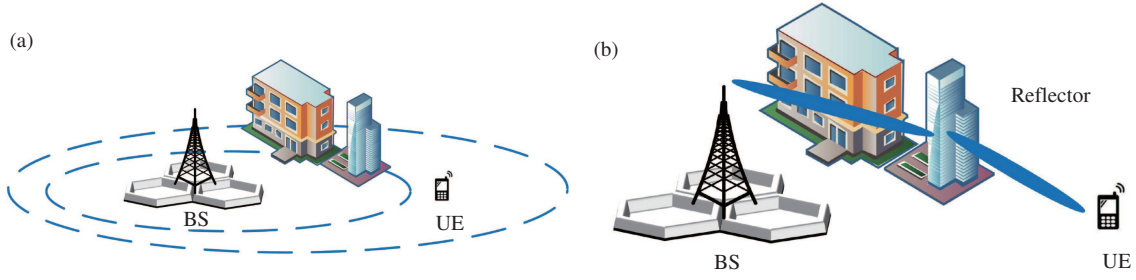
Figure 6 (Color online) Signal propagation in sub-6 GHz and mmWave systems. (a) For sub-6 GHz system with omnidirectional antennas, signals suffer from rich scattering and poor multi-path separation capability. (b) For mmWave system with directional antennas, signals will arrive at the receiver with very few paths due to the high carrier frequency and accurate antenna direction.

particular, mmWave communication operates at a carrier frequency range from around 30 to 300 GHz. The mmWave communication at 60 GHz with bandwidth up to 7 GHz has been standardized for WPANs, e.g., IEEE 802.11ad [144] and IEEE 802.15.3c [145]. In [136], the authors compared the raw resolution of time-based localization across different frequency bands, where the raw resolution is defined as the ratio of the speed of light and the available bandwidth. They revealed that mmWave signal specified in IEEE 802.11ad with bandwidth over 2 GHz can achieve raw resolution of roughly 15 cm, while that of the UWB signal with bandwidth over 500 MHz is only about 60 cm. Compared with conventional location reference signals, which are usually transmitted on the ISM frequency bands, and suffer from severe interference and multi-path effects, mmWave signals can reduce the probability of interference and have the capability to resolve the multi-path components thanks to the very large bandwidth [146]. Another important feature of mmWave transmission is the channel sparsity; that is, only a limited number of propagation paths can reach the receiver due to the short signal wavelength, which can be also exploited to enhance the localization performance [147, 148]. However, to overcome the high path loss associated with mmWave signals due to the short wavelength [149], mmWave transmission is usually combined with massive MIMO for directional beamforming, for which the accurate angular information can be extracted and utilized for localization. Therefore, mmWave massive MIMO-based localization may significantly outperform the conventional angle-based localization at lower frequencies, which suffer from rich scattering and poor multi-path separability [141]. As illustrated in Figure 6, compared with traditional sub-6 GHz system, the limited scattering and high directivity are unique characteristics of mmWave massive MIMO communications, which bring new opportunities for localization. In the following contents, we discuss the mmWave massive MIMO localization in terms of channel model, parameter estimation, and localization approaches.

Channel model. For convenience, we consider a MIMO system equipped with ULAs with N_t antennas at the BS and N_r antennas at the UE. The corresponding MIMO channel can be modeled as [150]

$$\boldsymbol{H}(t) = \sum_{l=1}^{L} \alpha_l \boldsymbol{a}(\theta_r, l) \boldsymbol{a}(\theta_t, l)^{\mathrm{H}} \delta(t - \tau_l), \tag{43}$$

where L denotes the total number of multi-paths, which is usually small due to the multi-path sparsity in mmWave regime [150]; α_l and τ_l represent the complex path gain and the delay of the lth path, respectively; $\boldsymbol{a}(\theta_{r,l})$ and $\boldsymbol{a}(\theta_{t,l})$ are the antenna array response vectors for angles $\theta_{r,l}$ and $\theta_{t,l}$ seen by the UE and BS, respectively. The sparsity of mmWave channel can be exploited to simplify the estimation of channel parameters [147]. This method uses a virtual representation for (43) by sampling in time, frequency, and/or space with the aid of the discrete Fourier transform (DFT). For instance, the beamspace MIMO channel can be represented as

$$\boldsymbol{H}_b(t) = \boldsymbol{U}_r^{\mathrm{H}} \boldsymbol{H}(t) \boldsymbol{U}_t, \tag{44}$$

where $\boldsymbol{H}_b(t)$ is the beamspace channel matrix, which is a linear equivalent representation of the antenna domain channel matrix $\boldsymbol{H}(t)$, $\boldsymbol{U}_r$ and $\boldsymbol{U}_t$ are the DFT matrices whose columns are response and steering vectors, respectively, which are orthogonal. Estimating parameters from $\boldsymbol{H}_b(t)$ is much easier than the parameter estimation based on $\boldsymbol{H}(t)$, since the latter determines the parameters in a nonlinear manner.

Parameter estimation. For MIMO systems, the source localization involves the AOAs and AODs estimation, which can be classified into localization for point sources and distributed sources [151, 152].

For point source localization, the signal of each source emitted from a single AOA or the AOAs of different sources is distinguishable. On the other hand, for the distributed source localization, the signal of each source is emitted from an angular region. In general, point sources usually correspond to the LoS propagation scenarios [151], while distributed sources are commonly used for the multi-path scenarios [152]. For distributed sources, they can be further classified into coherently distributed (CD) sources [153] and incoherently distributed (ID) sources [154], depending on whether they are slowly time-varying channels or rapidly time-varying channels.

In general, the parameters for estimation include the azimuth and elevation angles of signal departure and arrival, signal propagation delay, Doppler shift, and the corresponding uncertainty of these parameters. The channel parameter estimation approaches can be categorized based on the source type. For point sources, the authors in [155] categorized the channel parameter estimation methods into subspace-based and compressive sensing methods. The subspace-based methods treat the parameters into tensors, and use the tensor decomposition method to reduce the dimensionality. Such methods can achieve a good balance between estimation accuracy and computational complexity [156]. Compressive sensing is a promising method to recover the sparse signals, which is particularly suitable for mmWave massive MIMO system due to its sparsity in angular and delay domain [157]. For distributed sources, the parameter estimation can be classified into CD sources and ID sources estimation. The classical estimation approaches for point sources have been successfully applied into CD sources estimation [158–160], and the multiple signal classification (MUSIC) algorithms were employed to estimate the AOA of CD sources [153]. However, the parameter estimation for ID sources is more complicated. These methods can be generally divided into parametric methods and non-parametric methods. The parametric approaches include ML [161], covariance matching [162] and pseudo-subspace [163]. The ML estimation is optimal [161], but it suffers from high complexity. The LS estimator is proposed by using the covariance matrix matching techniques to reduce the computational complexity [162]. The non-parametric approaches like beamforming approaches [164] have lower computational complexity but with compromised performance compared with parametric approaches.

Localization approaches. The conventional localization approaches can be also used in mmWave MIMO systems for improving localization performance. In [165, 166], a 3D localization method was proposed by using hybrid signal measurements in mmWave systems, which can jointly estimate the position and velocity of the UEs and construct the environment maps. In [167], the method based on RSS measurements of mmWave signals was proposed, which can achieve localization accuracy of one meter. The mmWave-based object tracking was proposed in [168], where the RSS and signal phase were used as the features for tracking an object. In [136], a set of feasible localization approaches in mmWave bands were discussed, and the results demonstrated that mmWave localization can achieve decimeter level localization accuracy. In [141,169], the methods for estimating the object's position and orientation by using mmWave MIMO were proposed, where the CRLB on position and rotation angle estimation was derived. Moreover, due to the highly directional narrow beamforming of mmWave signals, the beam training protocols that have been standardized in IEEE 802.11ad [144] can be used for improving localization. With the beam training protocol, the strongest signals are selected for AOA, AOD, and TOA estimation. Furthermore, the mmWave localization can turn multi-path from a foe into a friend [12]. By considering the signal reflectors as virtual transmitters [12, 13], the high accurate localization is possible even without the LoS link. A hybrid localization approach for massive MIMO systems combining with TDOA, AOA, and AOD measurements was proposed in [170]. In [171], a fingerprinting-based localization method in massive MIMO systems was proposed, where the uplink RSS measurements were used as the fingerprints. In [172], the direct source localization (DiSouL) was proposed by jointly processing the observations at the distributed massive MIMO BSs. In [173,174], the combination of TDOA and AOA measurements by using the EKF was proposed. A comprehensive discussion on massive MIMO localization was given in [155]. An overview about the 5G mmWave localization for vehicular networks is given in [175].

3.2.2   *D2D communication and cooperative localization*

To meet the ever-increasing throughput demand of various applications on mobile devices, one promising method is to shift the current cell-centric architecture to device-centric architecture [176]. In traditional cellular networks, all communications must go through the BSs even if both communication entities are very close with each other, while D2D communication enables that two devices can communicate without
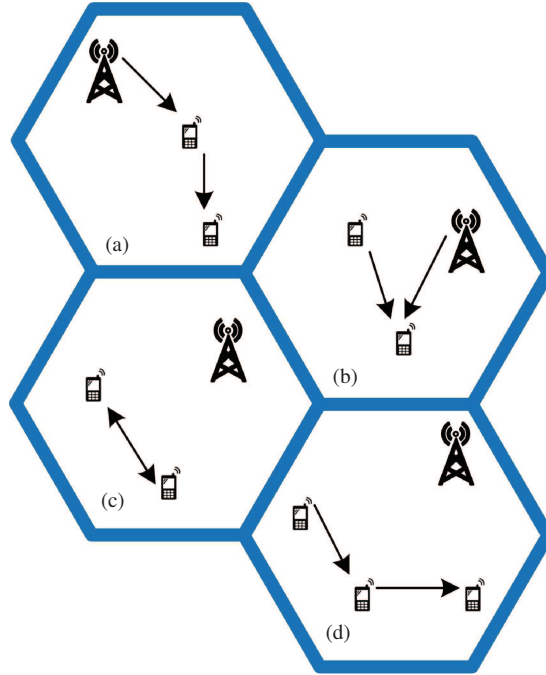
Figure 7 (Color online) An illustration of the device-centric networks and four types of D2D communication. (a) Device relaying from BS; (b) BS-aided D2D communication; (c) direct D2D communication; (d) device relaying from other devices.

traversing the BSs. The related use cases include multi-hop relaying [177], peer-to-peer (P2P) communication [178], machine-to-machine (M2M) communication [179], cellular offloading [180]. A detailed survey of D2D communication was given in [17]. In general, D2D communication combining with densely deployed small cells can achieve high data rate with high spectral efficiency and low latency, which paves the way for cooperative localization. In this subsection, we outline the localization schemes based on different types of D2D communications in cellular networks.

Depending on the role that the BSs played in different communication schemes, D2D communication can be classified into four categories [181], namely, device relaying from BS, BS-aided D2D communication, direct D2D communication, and device relaying from other devices, as illustrated in Figure 7. For device relaying from BS, a device located at the edge of a cell can only receive weak signals transmitted from the BS, which is assisted by other devices via relaying. In this case, device localization for the cell edge users can be achieved by utilizing the assisting devices as the pseudo BSs with known location, and the signal measurements among these devices can be used for localization. However, the location information of the assisting devices might be inaccurate, which will degrade the localization accuracy of the target device. Therefore, how to mitigate the effect of error propagation is a critical issue. For BS-aided D2D communication, two devices communicate with each other via their direct link, together with the assisted information provided by the BS. In this case, a device can measure the signals from both the BS and the other device, which can be utilized for localization. Therefore, the assisting device acts like an additional AN to provide extra information to improve the localization performance. By contrast, in the architecture of direct D2D communication or device relaying from other devices, two devices communicate with each other directly or indirectly via information relaying by other devices without traversing the BS. In such architectures, for the target device, the relative distance and/or angles to other devices can be obtained, and if the locations of other devices are known, cooperative localization can be achieved. For the multi-hop based cooperative localization, the key problem is to perform location estimation using multiple signal measurements from different devices. Based on the spectrum usage, D2D communication can be classified into inband D2D communication, which reuses the cellular spectrum for D2D data transmission, and outband D2D communication, which usually uses unlicensed spectrum. From a localization perspective, the main advantage of inband D2D lies in the better spectrum utilization, which avoids the consumption of extra hardware for other unlicensed spectrum, while the main drawback is the potential high interference, which may severely degrade the localization performance. On the other hand, the outband D2D communication transmits signals by using different wireless technologies, and the

location-related information is measured from heterogeneous networks. In this case, how to coordinate the signal transmission over different bands and data fusion for improving localization accuracy is a key problem. Furthermore, for all cooperative localization schemes, it is critical to protect the location privacy of the users and avoid the significant performance degradation simultaneously.

3.2.3 *Localization in UDNs*

Network densification is a promising technology to meet the ever-increasing demands on area coverage and capacity enhancement in 6G mobile networks. In the extreme case, we would have UDNs, which refer to such networks with more cells than users [182–184]. In [184], a quantitative definition of the UDN was given, where a network can be considered as ultra-dense if the density of cells is no smaller than 10^3 cells/km^2. In general, the densely deployed cells in UDNs can be classified into full-functioning BSs (picocells and femtocells) and macro-extension APs (relays and remote radio heads) [183]. The coverage area of all these cells is typically small, ranging from a few meters up to 100 m. The fundamental difference of UDNs from traditional networks lies in that it not only enables higher data rates with less energy consumption for communication, but also brings new opportunities for localization. First, since more small cells are in the close vicinity of users, the localization error of CID-based methods can be reduced. Furthermore, as the cell shrinks in UDNs, the probability of LoS link increases, which is beneficial for accurate location-related measurements [174]. In [185], the authors proposed a joint localization method combining TOA and DOA measurements, as well as a real-time UE tracking method by using an EKF. The simulation results showed that the localization accuracy is below 1 m for 95% of the case with the signal bandwidth below 10 MHz using one or two BSs in the 5G UDN. In [186], a joint positioning and synchronization method based on centimeter wave dense 5G network was proposed, which is able to estimate the clock offsets in addition to the UE's location.

3.3 Location-aware communication

Accurate location information can be used not only for location-based services, like location-based advertising, autonomous driving, but also for improving the communication performance, which is known as location-aware communication or location-aided communication. Location-aware communication has received fast growing attentions in recent years, where the location information can be utilized in a variety of ways to improve the communication performance in 5G-and-beyond networks, like reducing the communication overheads and delays, minimizing the energy consumption, and increasing the communication capacity. The utilization of location information in cognitive wireless networks was studied in [22], where the location-assisted network optimization was discussed, including location-assisted dynamic spectrum management, handover, as well as network planning and expansion. In [21, 187], the potential of location-aware communication in multi-user and multi-cell systems was discussed, with special emphasis on utilizing location information for resource allocation. In [20], a comprehensive survey about the location-aware communication was given based on the layers of the protocol stack, including the physical, MAC, and network layers.

3.3.1 *Physical layer*

In the physical layer, the location information is usually used for channel estimation, beamforming, and generating the radio environment maps (REMs) to reduce the interference and signaling overhead.

For location-aided beamforming, the accurate AOD and AOA of the LoS path can be exploited to design the beam vectors of the transmitter and receiver, respectively. Compared with the conventional beamforming methods based on full-band channel state information (CSI), the location-based beamforming schemes do not require the full-band reference signals, and the narrowband pilots are sufficing [174], which can reduce the energy and resource consumption especially for the UE. In [188], a beamforming method was proposed based on the AOAs or AODs of the LoS path between ANs and the UE by using the EKF to estimate and track the location of the UE. The results showed that if the LoS path is available, with the angular error below 2°, the location-based beamforming scheme outperforms conventional CSI-based schemes in terms of mean user-throughput and time-frequency resources efficiency. Another major problem of beamforming with large antenna arrays is the huge training overhead, which can be reduced by utilizing accurate location information of the transmitter and receiver [189–191]. Location-aided beamforming method for mmWave vehicular communication has been studied in [189], which can reduce

channel estimation overhead and speed up the initial access. In [190, 191], a compressive sensing based mmWave beamforming method was proposed, where the location information of the transmitter/receiver was used to design the sensing matrix. Currently, the accuracy of location aided for beamforming is within meter-level, while the more accurate location information of the UE can further reduce the overhead of beamforming.

The high-precision REM [192–194] can be constructed by using the accurate location information combined with channel quality metric (CQM) [20]. The main difference between the REM and geolocation database is that REM contains the radio elements like the knowledge of large-scale fading and location-based radio condition, which can be utilized for RRM without the CSI between the BSs and the UE [174]. The REM construction entails a training process, where a measurement center collects the CQM from different locations of the area and performs a learning algorithm to obtain a number of radio scene parameters, like the path loss exponent and shadowing variance, tagged with the actual locations [20]. In general, the number of sampling points in the area of interest, the location accuracies of these points, the dynamics of the propagation environment, and the accuracy of the propagation modeling can significantly affect the REM performance. To compare the estimated REM with the true REM, some quality metrics are proposed, including RMSE [195], correct detection zone ratio (CDZR) and false alarm zone ratio [196]. In general, a REM which can predict the large-scale channel fading in the area of interest needs at least meter-level localization accuracy. The more accurate location resolution leads to more accurate REM. However, since the REM construction must perform online, there is a trade-off between complexity and accuracy.

An accurate REM can be used for RRM at different network layers to reduce the overhead and delay of communications. In the physical layer, one of the best known applications is the REM-based spatial spectrum access scheme in cognitive radio (CR) [197], where before the CR devices initiate a communication, they first query the REM database for the available frequencies depending on their locations, and select the frequency bands from a set of received unoccupied frequencies. Furthermore, with the additional motion information of the UE, like velocity and orientation, we can predict the locations of the UE. Combining the REM and the predicted UE locations, the proactive RRM can be achieved, which can not only maintain the communication performance at a certain time but also adapt to the upcoming events. For example, the authors in [198] outlined the adaptive mobile communication for the predicted capacity. Moreover, if the predicted location of the UE and the accuracy of the REM are guaranteed, the proactive RRM can enable a high robustness and low latency E2E communications without the instantaneous CSI between the BS and UE [199]. Other applications of location-aided communication applications in the physical layer have also been studied, such as location-aided MIMO interference channels and coordinated multi-point (CoMP) transmission [200].

3.3.2 *MAC layer*

In the MAC layer, the accurate location information is beneficial for MAC layer RRM, like location-based multicasting and broadcasting, scheduling, and load balancing. Compared with the physical layer, the location accuracy requirements on the MAC layer is lower, usually range from several meters to tens of meters. The geographical location aided broadcasting is referred to as geocasting. For mobile ad hoc networks (MANET), broadcasting is an important method to quickly deliver message to the specific set of nodes, especially when the route to the destinations is still unknown. The traditional broadcasting methods based on flooding suffer from large bandwidth consumption and the broadcast storm problem (BSP) [201]. In [202], the authors proposed a broadcasting protocol based on location-related information in terms of distance and angle between transmitter and receiver, which can achieve high reachability and bandwidth efficiency. Recently, in [203], the authors proposed a location aided probabilistic broadcast (LAPB) algorithm for MANET routing, which selects the more effective nodes, according to an adaptive probability based on location information, to broadcast route request. The simulation results showed that the LAPB method is able to significantly reduce the overhead and alleviate the BSP in MANETs. Another essential transmission scheme of ad hoc networks is multicasting, which can also benefit from the accurate location information, and some location-based multicast routing protocols have been proposed in [204–206].

Another well-known application is the location-aided radio resources scheduling, where the same resources can be shared by two different communication links if the inter-interference level is below a certain threshold. In [187], the authors proposed a location-aided round robin scheduling algorithm to solve the

problem introduced by fractional frequency reuse that the fractional bandwidth can meet the heavy traffic demand during rush hours, where some of the cell-central and cell-edge users are selected to share the same frequency band with minimum intra-cell interference. Compared to alternative methods that select users based on the instantaneous channel knowledge requiring all users to feedback CSIs, the location-based scheme requires less feedback information and can achieve higher total throughput than the conventional methods. In [207], a passive location resource scheduling scheme based on an improved genetic algorithm was proposed. In [208], the authors proposed proactive scheduling schemes for delay-constrained traffic, where the current user location and the priori statistics were used to predict the request arrival time slots, which can significantly reduce the transmission energy compared to the reactive methods. For the WSNs, the energy efficiency is a critical issue due to the limited lifetime of the sensors' batteries, so the authors in [209] proposed the location-based sleep scheduling schemes, which can dynamically schedule the awake or asleep status of the sensors to reduce the total energy consumption.

The future wireless networks are expected to be highly heterogenous and dense, where different types of APs, like WiFi, BlueTooth, and light fidelity (LiFi), coexist in the small areas, leading to the problems of unbalanced traffic loads and inefficient resource utilization [210]. The load balancing technique is critical to solving such problems [211], which can offload some traffic load from busy APs to idle APs. For example, in [212–214], the authors considered the utilization of load balancing techniques to optimize the handover overhead and throughput of the LiFi and WiFi integrated hybrid networks. The location information can help the dynamic load balancing, which is an adaptive scheme depending on the UE distribution within a cell [215]. Since the distribution of UEs can be obtained based on the accurate locations of UEs, the location information can help the network APs to allocate resources more efficiently. For instance, in [214], the authors proposed a location-aided load balancing scheme for hybrid LiFi and WiFi networks, which aims to maximize the system throughput.

3.3.3 *Network layer*

In the network layer, location information can be used in various aspects, like geographic routing, location-aware content delivery, to improve scalability and reduce network overhead and latency. The geographic routing is usually referred to as the georouting, which utilizes the geographic location information of the nodes in the networks to transmit the data packets to their destination [20]. Different from the topology-dependent routing schemes, the georouting protocols depend on the physical location, eliminating the requirements on topology storage and reducing the associated costs, which is especially suitable for wireless ad hoc networks. The routing performance in ad hoc networks is affected by the following factors. First, due to the lack of centralization, the wireless nodes in ad hoc networks may not operate with unified standards. Furthermore, since ad hoc networks are usually dynamic and mobile (e.g., MANET), which are formed and changed depending on the particular goals, with nodes joining in and leaving the network from time to time, thus the topology of the ad hoc network is constantly changing [216]. All such features of ad hoc networks bring new challenges to the design of routing protocols. The conventional ad hoc routing protocols can be classified into proactive and reactive protocols. The main drawback is that both of them require the topology information for message routing, which entails high maintenance cost. Compared to the conventional routing schemes, the georouting protocols based on the geographic location of the nodes can eliminate the need for topology information. In the georouting scheme, when a node receives a packet, it will select the most appropriate node from its neighbours and deliver the packet in a hop-to-hop manner based on the location of the targeting nodes. There is no specific route for a particular destination, since the selection of the nodes depends on the different network states. One classical georouting scheme is greedy forwarding, where the packets are forwarded to the neighboring node which is closer to the destination at each hop [217]. In addition to improving routing performance, other georouting protocols have been designed with particular goals. For example, in [218], the authors proposed a georouting scheme which can feature quality of service (QoS) predictions based on the mobility of the UE. In [219], an energy-efficient georouting protocol was proposed, which can guarantee delivery. A more comprehensive overview on georouting was given in [216].

With the accurate location and tracking information of users, the location-aware adaptive content delivery schemes can be achieved, which include adaptive quality streaming, in-network caching, and content prefetching [220]. Such applications require the location and trajectory information of the mobile users. The well-known adaptive quality steaming scheme is the HTTP-based adaptive streaming (HAS) [221], where the mobility trajectories of users were exploited to optimize HAS quality adaptation,

thereby preventing data stream stalls. In [221], the authors developed an adaptive algorithm which can proactively switch the transmission rates based on the predicted user location and the stored REM. The in-network caching techniques enable caching media content closer to the mobile user to reduce the delay and the prevent congestion in busy hours [222], where the popular media content is stored at the edge of the network. The location of users and their mobility patterns are used to predict the hot network regions where the media content will most likely be requested, so in such cases, the coarse localization method with tens or even hundreds of meter accuracy suffices. On the other hand, the content prefetching refers to that with the help of the REMs and the predicted user's location, the network is able to deliver content proactively. For accurate content delivery, it may require meter-level location service. Specifically, when the networks are aware that the mobile UE will experience poor QoS based on the predicted UE's location, it can load a part or all of media content into the local storage of the UE in advance [220]. Therefore, content prefetching has the capability to provide seamless streaming services to users. In this case, the accurate location information of UE can improve the effectiveness of prefetching [223].

3.4　Localization and communication co-design

Since radio signals can simultaneously carry data and location-related information of the transmitters, a unified study on ILAC tends to be a natural choice. Currently, although localization systems can utilize the existing communication infrastructures like cellular networks cost-effectively, the localization and communication systems are mostly designed separately. This is mainly due to the fact that the two lines of work in general have different design goals: one is to maximize the localization accuracy, while the other is to maximize the reliable data transmission rate through the wireless channels subject to fading and noise. Nevertheless, there is a growing trend that the designs of the communication and localization systems can be aligned. For instance, increasing the signal bandwidth or signal-to-noise ratio (SNR) can increase the localization accuracy as well as the channel capacity. Hence, in this subsection, we attempt to discuss the study on ILAC for giving some insights on future network design.

Although there is extensive work on localization in cellular networks as discussed above, the studies on localization and communication co-design are relatively rare. In [224–228], some beamforming schemes for joint localization and data transmission in mmWave for 5G networks were studied, where the different trade-offs between the localization accuracy and the data rate were derived. In [224], the authors studied the trade-off between localization efficiency and downlink data rates by tuning the mmWave BS transmit power. Specifically, the total transmit power of the BSs is fixed and divided into two parts: one associated with localization and the other dedicated for communication. Hence, the optimal transmission scheme exists to select the proper power splitting factor for supporting different QoS requirements. In [225], the authors focused on optimizing the beam vectors at the BSs to minimize the overall power consumption under certain data rate and localization accuracy requirements, where the data rate and the localization accuracy were measured as the functions of beam vectors, and the CRLBs of localization accuracy for both TOA- and TDOA-based localization methods were derived. In [226], by assuming a finite coherence time, a trade-off between communication rate and localization accuracy for single-user LoS mmWave communication was studied. Specifically, the total communication duration is fixed and partitioned into beam alignment and data transmission stages, and both of which are quantified as functions of the codebook size for beam alignment. The following trade-off between localization accuracy and communication rate was revealed: more time spent for beam alignment leads to better SNR and improved localization accuracy, but it will result in less time for data transmission and hence lower data rate. Moreover, in [227], the impact of imperfect beam alignment on the rate-localization trade-off was considered. In [228], the authors extended their prior work in [226] into a multi-user scenario, and the trade-off between the sum-rate and localization accuracy in the uplink for multi-user mmWave communication system was researched.

In this subsection, we give the general architectures of the ILAC systems and consider their performance trade-offs. For simplicity, we focus on the basic downlink case in the additive white Gaussian noise (AWGN) channel. At the transmitter side, the signals for communication and localization can be designed and transmitted separately, or a common signal can be shared and reused for the both of purposes, which are referred to as separated signals and shared signals, respectively. At the receiver side, the localization and communication systems can have their respective receivers that operate as an information decoder (ID) and the LMU, respectively, or share a common receiver, where the channel estimation unit (CEU) for communication can be reused for measuring location information from the common received signal, which is referred to as separated receivers and shared receiver, respectively. Therefore, as illustrated in
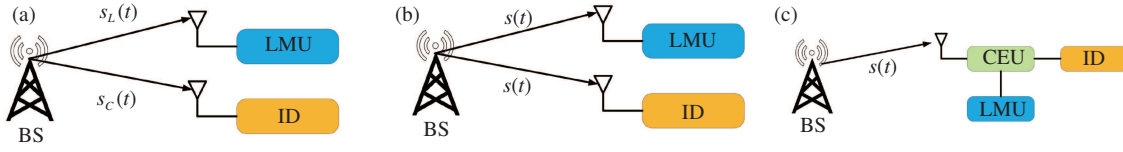
Figure 8 (Color online) An illustration of different architectures of ILAC systems. (a) Separated signals and receivers; (b) shared signals but separated receivers; (c) shared signal and receiver.

Figure 8, the architectures of the ILAC systems can be classified into three categories: separated signals and receivers, shared signal but separated receivers, as well as shared signal and receiver.

3.4.1 *Separated signals and receivers*

As illustrated in Figure 8(a), in this case, we consider a scenario that the total bandwidth is fixed, denoted by B_t, which is divided into two orthogonal frequency bands for localization and communication respectively. The signal for localization denoted by $s_L(t)$ has the bandwidth $B_L = kB_t$, where $0 \leqslant k \leqslant 1$. Therefore, that for communication, denoted by $s_C(t)$, has the bandwidth $B_C = (1 - k) B_t$, and there is no interference between them due to orthogonality in the frequency domain. It is assumed that the two signals have the same transmission power, i.e., $\mathbb{E}[|s_C(t)|^2] = \mathbb{E}[|s_L(t)|^2] = P$. The signals received at ID and LMU are

$$y_C(t) = h_C(t)s_C(t) + n(t), \tag{45a}$$

$$y_L(t) = h_L(t)s_L(t) + n(t), \tag{45b}$$

where $t \in [0, T]$ is the observation time; $h_C(t)$ and $h_L(t)$ are the communication and localization channel coefficients, respectively, where we assume $\mathbb{E}[|h_C(t)|^2] = \mathbb{E}[|h_L(t)|^2] = 1$ without loss of generality; $n(t)$ denotes the AWGN, with the noise power spectral density (PSD) of N_0 for all frequencies. According to Shannon's channel capacity theorem [229], the maximum data transmission rate is

$$R = B_C \log_2 \left(1 + \frac{P}{N_0 B_C} \right). \tag{46}$$

On the other hand, considering the time-based ranging for localization, the CRLB is formulated as [109, 230]

$$\mathbb{E}\left[(d - \hat{d})^2 \right] \geqslant \text{CRLB} = \frac{c^2}{8\pi^2 \beta^2 PT/N_0} = \frac{c^2}{8\pi^2 \beta^2 \frac{P}{N_0 B_L}}, \tag{47}$$

where it assumes a rectangular-shaped signal transmitted for localization, with $B_L T \approx 1$, and

$$\beta^2 := \frac{\int_{-B_L}^{B_L} f^2 |S_L(f)|^2 \, \mathrm{d}f}{\int_{-B_L}^{B_L} |S_L(f)|^2 \, \mathrm{d}f} \tag{48}$$

denotes the mean square bandwidth (MSB) of the signal $s_L(t)$, and $S_L(f)$ is the Fourier transform of $s_L(t)$, which will increase as the signal bandwidth increases. Eqs. (46) and (47) reveal that increasing the dedicated bandwidth can improve the data rate and decrease the CRLB. Therefore, as the total bandwidth is fixed, there is a trade-off to select the proper bandwidth splitting factor to support different QoS requirements for communication and localization, which is referred to as bandwidth splitting scheme. Alternatively, we can consider that the two types of signals use the same frequency band with bandwidth B, but with different transmission power, i.e., $\mathbb{E}[|s_C(t)|^2] = P_C$ and $\mathbb{E}[|s_L(t)|^2] = P_L$. The total transmission power for a single-user is fixed, i.e., $P = P_C + P_L$. For each receiver, the received signal is

$$y(t) = h_C(t)s_C(t) + h_L(t)s_L(t) + n(t). \tag{49}$$

Similarly, here we assume $\mathbb{E}[|h_C(t)|^2] = \mathbb{E}[|h_L(t)|^2] = 1$ for notational simplicity. In this case, the receiving performance is typically characterized by signal-to-interference and noise ratio (SINR), and the communication and localization performance can be obtained as

$$R = B \log_2 \left(1 + \frac{P_C}{N_0 B + P_L} \right), \tag{50a}$$

$$\text{CRLB} = \frac{c^2}{8\pi^2\beta^2 \frac{P_L}{N_0 B + P_C}}. \tag{50b}$$

This reveals that signals with high power are beneficial for either data transmission or localization accuracy, so the appropriately splitting the transmission power is necessary to balance the trade-off between communication and localization functionalities, which is referred to as the power splitting scheme.

3.4.2 *Shared signal but separated receivers*

As illustrated in Figure 8(b), in this case, a common signal is shared by localization and communication systems, so the study on joint waveform optimization for improving the performance of such a dual-purpose system is critical. Although the role of localization was already highlighted in the 3G era, the dedicated reference signals for localization were not used until the 4G-LTE, with the introduction of the PRS. Therefore, compared with the extensive research on waveform design for communications, that localization in cellular networks is relatively rare. In [231, 232], the authors studied the impact of different signal PSD on time-based ranging accuracy. Conventionally, the signal power is uniformly distributed over the available spectrum. However, the authors revealed that the uniform PSD is strictly sub-optimal for signal propagation delay measurement. Actually, although the signal waveforms for localization and communications are usually designed separately, they have some similar key requirements, like low latency, high reliability, and low device complexity, rendering that their waveforms can be co-designed in future networks. For instance, considering the single-carrier transmission scheme, the signal that has more power concentrated at the edges of the spectrum is generally beneficial for time-based localization, which in the time domain can result in an impulse-like autocorrelation for time-based ranging. Furthermore, in such cases, the MSB of the signal can be greater according to (48), which leads to the lower CRLB for better localization accuracy. However, for communication, the optimal signal PSD scheme is to concentrate the signal power at the center of the mainlobe to reduce the inter-symbol-interference (ISI). Therefore, there is a trade-off in terms of waveform design between localization and communication with different PSD requirements. Moreover, it is important to integrate the geometrical information into waveform design for future networks, and the waveforms should be adaptive with reconfigurable features, which can flexibly configure their bandwidth, signal power, etc., depending on the real-time environment [232].

3.4.3 *Shared signal and receiver*

As illustrated in Figure 8(c), localization and communication systems can also share a common receiver with the functionalities of LMU and ID, which can extract location-related information and decode data from a common signal. In general, when channel parameters are unknown, the location-related information estimation in multi-path environments is closely related to channel estimation, so the CEU for communication can be also reused for localization. For instance, the energy detector for non-coherent demodulation and the matched filter for coherent demodulation in communication systems can be cost-effectively exploited for TOA estimation [233]. The path amplitudes and TOAs can be jointly estimated using the ML approach, which can achieve the CRLB for large SNRs. A primary barrier of ML estimators is the computational complexity, which results in time-consuming computation for accurate TOA estimation. Therefore, a natural trade-off between localization and communication is that the more time spent on channel estimation can give the more accurate localization performance, but results in the higher communication latency.

3.5 Localization and communication in aerial-ground integrated networks

Integrating terrestrial networks (especially the cellular networks) and aerial networks to achieve ubiquitous wireless connectivity in the 3D space is one of the visions for 6G [5]. In particular, with the high mobility and on-demand deployment capability, unmanned aerial vehicles (UAVs) have been regarded as a powerful tool to expand the wireless networks from the ground to the air space [234]. In general, UAVs may be used as the low-altitude platforms (LAPs) to assist the terrestrial wireless communication from the sky, which are typically deployed at an altitude below several kilometers, while the high-altitude platforms (HAPs) consisting of floating BSs (e.g., balloons) are usually deployed in the stratosphere with tens of kilometers above the earth surface. Compared with HAPs, UAV-based LAPs are easier and faster for deployment and more flexible for reconfiguration for critical missions, and it is able to establish a strong LoS communication link with the ground UEs directly without relying on extra communication

infrastructure like dish antennas [235]. In general, UAV-aided communications have three main use cases, i.e., UAV-aided ubiquitous coverage, relaying, and data collection [235]. On the other hand, UAVs with their own missions may also be integrated into cellular networks as new aerial users, leading to the other paradigm known as cellular-connected UAVs [236]. The 3GPP started a study item to exploit the potential of LTE support for UAVs in March 2017 [237], and the related technical reports TR 36.777 was released in December 2017 [238], followed by a work item. To achieve reliable data transmission for either UAV-aided communication or cellular-connected UAVs, the robust, low latency, high data rate wireless links between UAVs and terrestrial networks are necessary. Due to the high mobility, the accurate real-time location of UAV is important for both safe operation and communication links maintenance. In this subsection, we first focus on the localization problems in wireless networks involving UAVs, which can be classified into two main categories, i.e., localization for UAVs and UAVs for localization, where the UAVs play a role as agent nodes and aerial ANs, respectively. Then, we elaborate the importance of real-time location information of UAVs for communication, which is referred to as location assisted UAV communication.

3.5.1 *Localization for UAVs*

We first give an overview on the conventional localization and navigation methods for UAVs, and then discuss the potential localization approaches when cellular-connected UAV is enabled in future networks.

Currently, the commonly used UAV localization and navigation systems can be mainly divided into three categories, namely, GNSS, INS, and vision-based navigation. The GNSS (e.g., GPS) can provide global coverage with meter-level localization accuracy for UAVs [239], while it is vulnerable to disruption of satellite signals due to obstacles or signal blocking. On the other hand, INS is a self-localization system by utilizing the motion information of the UAV for localization and navigation. However, the INS of UAVs is expensive and unsuitable for small aircraft. Moreover, it suffers from bias errors caused by the integral drift problem, which are continuously increasing with time, resulting in accuracy degradation [240]. To tackle this problem, the combination of GPS and INS was proposed, where the data fused from GPS and INS sensors through the navigation filter (e.g., EKF) can be used to improve localization performance. For instance, in [240], the authors proposed a new INS/GPS sensor fusion approach based on state-dependent Riccati equation nonlinear filtering, which showed better UAV localization performance compared to the method based on EKF.

However, many applications require the UAV to operate in GNSS-denied environments, like cluster urban and indoor scenarios. To this end, the methods based on image recognition, referred to as vision-based navigation, have emerged as a promising alternative to INS/GPS, which can be used both in outdoor and indoor environments. The various visual sensors (e.g., cameras) are used to acquire information of surroundings, and the visual odometry and other similar methods were proposed to localize and navigate the UAV based on computer vision [241]. The advantages of vision-based methods include that they do not rely on external signals, and the visual sensors are cheaper and easier to deploy compared to INS/GPS sensors. However, the main limitation lies in that the UAV needs to process a large amount of sensing information in real time, especially for image processing, which greatly increases the computation complexity. Therefore, vision-based localization methods are difficult to be applied for UAVs with low power consumption and limited computing resources. Moreover, since the vision-based methods rely on the visual information of the environment, the accuracy of these methods is usually poor in challenging environments with low-visibility conditions, like dusty or smoking environments [242].

For indoor UAV localization, the radio signal based methods were also proposed, like UWB-based [242–245] and WiFi-based systems [246, 247]. In [243], the authors proposed an UWB-based indoor localization system for UAV localization, which can achieve the RMSE under 10 cm in the horizontal plane and under 20 cm in 3D space for 95% cases. Furthermore, a method that combines IMU, UWB, and vision-based schemes with EKF for UAV indoor localization was proposed in [245], where the 10 cm accuracy can be achieved. In [246], the authors proposed an indoor UAV localization method based on WiFi RSSI, where the distances between the UAV and WiFi APs were measured to locate the UAV. In [247], the authors proposed an RSS fingerprint-based HiQuaLoc system for indoor UAV localization, where an RSS interpolation algorithm was proposed to reduce the overhead on training phase.

Compared with GNSS, vision-based, and short-range radio-based localization approaches, cellular-connected UAV can bring many new opportunities for UAV localization. First, the ubiquitous accessibility of cellular networks may increase the localization coverage, which can cover outdoor and indoor
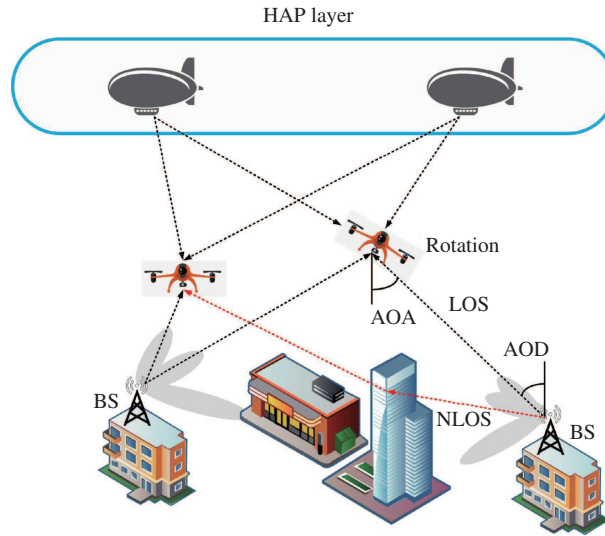
Figure 9 (Color online) An illustration of 6D localization and HAP-assisted localization approaches for UAVs.

environments. Furthermore, cellular-connected UAVs can be a good complementary to improve the GNSS performance, via techniques like differential GPS (D-GPS), to achieve more robust UAV navigation. Second, the wireless communication links between UAVs and cellular BSs, i.e., the control and non-payload communication (CNPC) link and payload data link, can be exploited and reused as the reference signals for UAV localization cost-effectively. Third, the legacy wireless localization techniques in cellular networks from 2G to 4G, like OTDOA, UTDOA, E-CID, can be extended to the sky for UAV localization. Finally, the new radio introduced by 5G, such as massive MIMO and mmWave communication, can be also exploited for UAV localization. Actually, some related studies like drone detection and tracking based on cellular networks are ongoing. For instance, in [248], the authors studied the amateur drone detection in 5G mmWave cellular networks, where the system design is outlined in terms of the density of BSs, their directional antennas, and the bandwidth, to detect the unlicensed small-sized drones. On the other hand, massive MIMO techniques with highly directional radiation pattern have also be exploited for UAV detection and tracking in [249]. In [250], the authors studied the localizability of UAVs with cellular networks, and it was concluded that the localizability for UAVs is more favorable than that for ground users since the former has better localization performance due to the higher altitudes.

Here, we envision two promising localization techniques for UAVs in future networks, referred to as 6D localization and HAP-assisted localization, as illustrated in Figure 9. For 6D localization, the goal is to precisely estimate not only the 3D spatial location of the UAV, but also its 3D orientation in terms of roll, pitch, and roll [251]. Different from the ground mobile devices for which the accurate real-time locations are typically sufficient, for the flying UAV, the accurate orientation estimation is also quite important, due to its high impact on flying gesture, power consumption, and hence the trajectory design. This leads to the important problem of the 6D localization for UAVs. The massive antenna arrays deployed on the BSs in 5G networks bring the opportunity to realize such a goal, by utilizing their high angular resolution to locate the cellular-connected UAV. Furthermore, combining with mmWave technology with the high carrier frequency and wide signal bandwidth, the antennas can be squeezed into a compact form factor, which can be deployed on the UAV. Therefore, both the AOA and AOD can be estimated in either uplink or downlink, as shown in Figure 9, and the location and the orientation of the UAV can be estimated simultaneously, leading to 6D localization. However, since the BS antennas are typically downtitled for ground users, the radio coverage in the sky may not always be guaranteed, which may degrade the localization performance of UAVs. To this end, the HAPs assisted localization approach provide a good complementary. For HAP-assisted localization, as shown in Figure 9, the quasi-static floating aerial BSs can be treated as the additional ANs to assist the localization for UAVs. Compared with terrestrial networks, HAPs can provide wide wireless coverage for very large geographic areas. Furthermore, due to the high likelihood of LoS links in high-latitude space, the UAV localization accuracy can be improved.
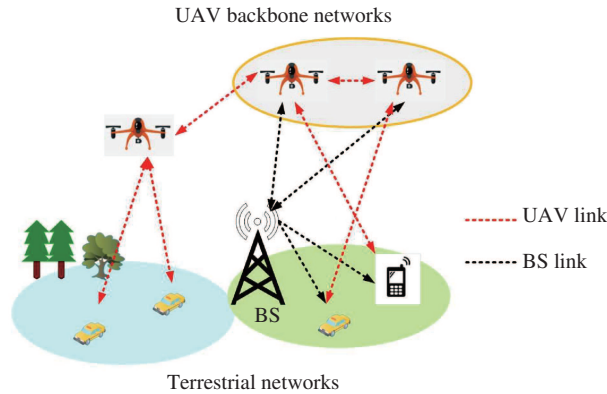
UAV backbone networks

Figure 10 (Color online) An illustration of FANETs and UAV-aided localization.

3.5.2 *UAV for localization*

Apart from estimating the locations for UAVs, the located UAVs can be also used for locating other devices in the networks, which is referred to as UAV-aided localization. UAV-aided localization can be mainly divided into two categories. One is cooperative localization among UAVs, which is quite important for the flying ad hoc networks (FANETs). The other is the UAV-assisted localization for terrestrial UEs, where the UAVs with known locations are treated as the additional ANs for providing more reference information to locate the agent node on the ground.

Cooperative localization among UAVs. Instead of developing single-UAV systems, a group of small UAVs can form FANETs in the aerial space. In such cases, as shown in Figure 10, only a subset of UAVs can communicate with the ground BSs, which is referred to as the UAV backbone networks, while other UAVs outside the coverage of the ground BSs may establish a connection with the ground through the UAV backbone networks [252]. Compared with the wireless nodes in vehicular ad hoc networks (VANETs) or MANETs on the ground, FANET nodes have much higher mobility with a more dynamic network topology. Due to the performance degradation of GPS in urban and indoor scenarios and the high mobility of UAVs, multi-UAV systems require the accurate real-time location information of each UAV in FANETs, which renders the cooperative localization among UAVs important. Intuitively, the cooperative localization approaches utilized in MANETs, VANETs, or other ad hoc networks on the ground can be similarly exploited and applied in FANETs by extending the localization schemes from 2D to 3D. For instance, the belief propagation technique which is usually used for cooperative localization in WSNs can be utilized for multi-UAVs cooperative localization. In [253], a dynamic nonparametric belief propagation method was proposed for UAVs cooperative localization, which can locate UAVs with fault GPS successfully. In [254], the authors studied the cooperative localization between two UAVs, which were equipped with heterogeneous sensors to gather more information in a limited time. In [255], the authors proposed a cooperative localization method based on the inter-UAV relative range measurements, which can locate the UAV when the GPS is unavailable. By assuming that all UAVs construct a ring communication topological structure, a cooperative localization method based on information synchronization was proposed in [256]. Typically, since there are many LoS links among UAVs, the cooperative localization performance of UAVs should be better than that of terrestrial wireless nodes, though the practical implementation is still challenging due to the high UAV mobility.

UAVs have higher altitude than ground BSs or UEs, which renders them easier to achieve wider coverage area on the ground with high probability of LoS links. Therefore, when some ANs are not available to the agent node, the flying UAV with a known location can be used as an additional AN to assist the terrestrial localization. For instance, a UAV-aided localization method for ground vehicles was proposed in [257], where each UAV first measures the RTOF of signals between the ground BSs and the UAV, and then broadcasts the measurements to the ground vehicles for localization, which can achieve decimeter-level relative position error between vehicles and meter-level absolute position accuracy. In [258], a range-based drone-aided localization method for terrestrial objects was proposed, where a flying drone was treated as a mobile anchor equipped with a GPS, and the distances between the drone and the ground objects were measured by UWB signals, then the ground objects can estimate their own position through trilateration.

3.5.3 *Location-assisted UAV communication*

Apart from navigation and tracking, the accurate real-time location information of a flying UAV is also beneficial for communication in the following aspects.

• **3D REMs modelling.** Currently, cellular networks are designed to cater for the terrestrial broadband communication, and thus BS antennas are typically downtilted to reduce the inter-cell interference [259]. With the downtilted BS antennas, the UAVs may only be served by the sidelobes, and thus the cellular coverage in the sky cannot always be guaranteed [260]. Therefore, the coverage holes exist in the sky, and when a UAV flies over these areas, it may lose wireless connection to the cellular networks. To overcome this problem, one useful approach is to construct 3D REMs about the area of interest to guide the design of UAV trajectories to avoid these coverage holes. Similar to the 2D REM as we discussed in Subsection 3.3, a 3D REM modelling also evolves a training process, where a UAV flies over the area of interest and collects radio measurements from the sampling points of the area to construct the 3D REM. Therefore, the localization accuracy of the flying UAV will affect the accuracy of the REM. For instance, in [194], the authors proposed a method named simultaneous navigation and radio mapping (SNARM), where the signal measurement of the UAV is used not only for optimizing the UAV trajectory, but also for creating a radio map which can predict the outage probabilities of communication at all locations in the area of interest.

• **Proactive management of CNPC links.** The wireless communication links of UAVs have two main categories, namely, the CNPC link and payload data link. The CNPC link is a two-way communication link between a UAV and the ground control station (GCS), or other UAVs, which is responsible for supporting the safe control from GCS to the UAV, sending reports from the UAV to GCS, and transmitting collision alert between UAVs [235]. On the other hand, the payload data link is mainly used to transmit mission-related data between the UAV and other entities, like the ground BS, mobile terminals, and other UAVs. Different from the data link which requires a higher data rate but tolerates on latency and reliability of the link, the CNPC link requires ultra-reliability, low latency, and high security to ensure safe control to the flying UAV [236]. However, due to the building blockage and shadowing effects, the CNPC links may suffer from delay and attenuation, and the interference from other UAVs and ground BSs can also degrade the CNPC performance. To enhance the reliability and robustness of the CNPC links, the combination of the 3D REMs and the predicted UAV location can provide opportunities for proactive CNPC link management. With the predicted locations of all UAVs in the area of interest, the GCS can proactively allocate the appropriate spectrum to provide CNPC link for the UAV according to the pre-built 3D REMs, enhancing the reliability and reducing the interference.

• **3D beamforming and sub-sector partition.** Another way to achieve the goal of ubiquitous wireless connectivity in 3D space is the 3D beamforming, which also relies on the accurate UAV location. Specifically, the accurate UAV location information can be used to obtain the azimuth and elevation angles of the aerial-ground links, and further help for designing the beamforming weight vectors. Furthermore, with the accurate 3D beamforming, the sectorization technique can be also extended to the 3D space, where the elevation angles are used to further partition the horizontal sector in current cellular systems to construct the sub-sectors for the aerial users [236]. With the help of these 3D network architectures, the interference in UAV-aided communication can be significantly reduced. Moreover, the accurate UAV location is also beneficial for RRM in these sub-sectors, like location-aided handover, multicasting, as we discussed in Subsection 3.3.

4 Conclusion and future working directions

In this article, we have first provided an overview on the basics of wireless localization, and then discussed the vision of future network design with ILAC towards 6G networks. In summary, an envisioned architecture of future 3D wireless networks is illustrated in Figure 11, and some related enabling technologies and promising directions of future work are discussed as follows.

As shown in Figure 11, it is envisioned that the future 6G wireless networks will be artificial intelligence (AI) enabled, heterogenous, and multi-tier networks consisting of space backbone networks (SBNs), space access networks (SANs), aerial backbone networks (ABNs), aerial mobile networks (AMNs), and terrestrial networks (TNs) [19, 261, 262].

The satellite networks are expected to play an important role in 6G networks due to their wide coverage, especially in remote and ocean areas. Since various satellite systems on different earth obits are isolated
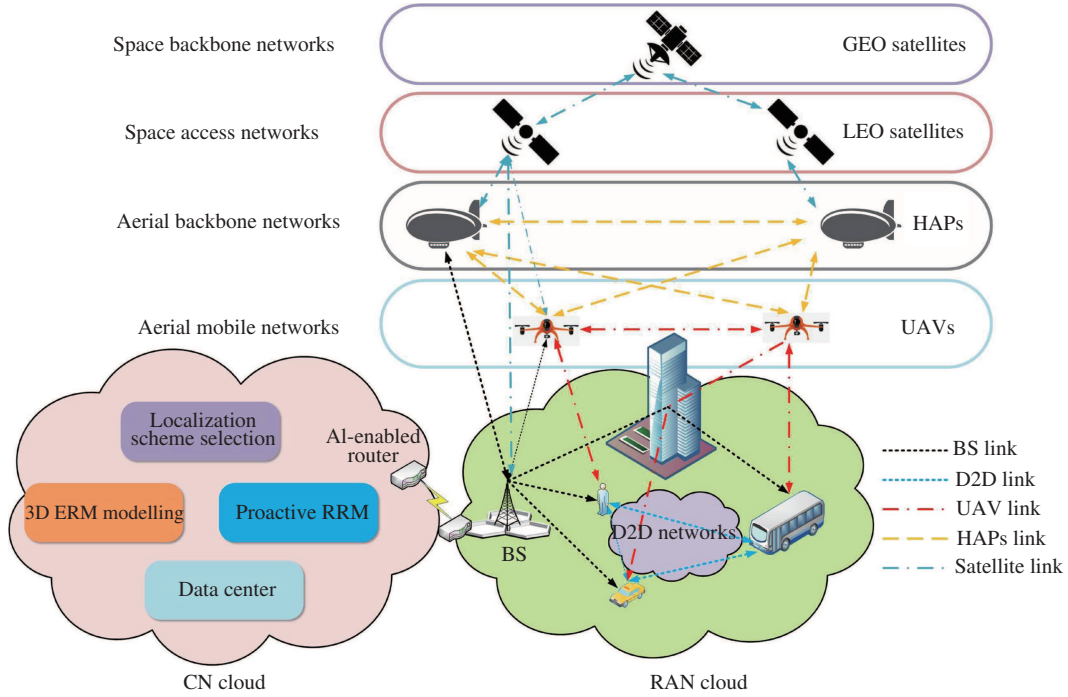
Figure 11 (Color online) An illustration of future 3D wireless networks.

and formed different autonomous systems (ASs) to provide specific services, like communications, navigation, remote sensing, the real-time information sharing across different ASs is difficult [262]. Therefore, in the future networks, the satellite networks can be divided into SBNs and SANs. The SBNs consisting of several GEO satellites have the capabilities of data storage, routing and resource management, which can function as the control and data center of the SANs consisting of LEO satellites [262]. LEO satellites may be equipped with high-gain large antenna arrays to provide ubiquitous coverage to enable UE direct communication [19]. For localization, the GNSS provided by the geostationary satellites will be still important for mobile device localization and network synchronization due to its wide coverage.

For aerial networks, on-demand UAVs with high mobility can be deployed easily and quickly, which can form the dynamic adjustable AMNs to provide temporary communication services for special missions, like emergency communication services, UAV-aided relaying, and data collection [235]. However, due to the limitations of UAVs on data storage and endurance, the ABNs consisting of floating HAPs like airships are necessary for providing reliable wireless coverage for large areas, where the airships with the computing and caching capabilities can be treated as the data center and gateways routing information to SANs and TNs. For localization, since the locations of HAPs can be determined by the GNSS or the terrestrial BSs with dedicated antennas to transmit reference signals, they can act as the aerial ANs to locate the flying UAVs. Furthermore, since UAVs with high mobility may provide short-distance LoS links to the ground devices, they can be used as the additional temporary ANs to improve the localization performance of ground devices.

In the TNs, mmWave, massive MIMO, and UDNs technologies can be combined to improve the communication and localization performance. Furthermore, as a promising trend, device-centric networks will flourish the cell-free mobile communications, which will require high accurate cooperative localization among devices for resource allocation. On the other hand, the future networks will become more heterogenous, combining with various radio access standards, like 2G/3G/LTE/5G, WiFi, as well as the Terahertz (THz) and visible light frequency bands. Therefore, the integrated resource allocation schemes and protocols are required to leverage multiple frequencies to provide seamless wireless connectivity for mobile devices. For localization, the system needs to select the appropriate localization approaches and reference signals depending on the surrounding radio environments of the mobile devices for better performance.

Finally, AI-enabled cloudization is a promising trend of future wireless networks, as illustrated in Figure 11, which can be mainly divided into RAN cloud and CN cloud. The distributed AI units

deployed at the edges of the RAN assist signal reconstruction, which is beneficial for data decoding and signal measurements to improve communication and localization performance, respectively. In the CN cloud, the centralized AI units can be used to construct the 3D REMs, which can be used to assist the localization scheme selection and proactive RRM for communications. As a vision of 6G, the AI-enabled networks are expected to autonomously optimize and manage the resource to dynamically maintain communication performance of the UE according to its accurate real-time location.

However, to achieve the above visions, there are still many challenges to be addressed. Note that while some challenges have already been discussed in Subsection 2.4, in the following, we outline some important directions for future work and highlight the promise of ILAC.

4.1 Fundamental performance analysis and design for ILAC

Achieving ultra-high spectral efficiency is critical for future networks, especially for the IIoT application scenarios, which require massive wireless connectivity to support reliable communication and accurate localization services for massive IoT devices. To achieve ultra-high spectral efficiency for ILAC, a common spectrum can be shared by localization and communication, or it can be divided into two orthogonal parts. Furthermore, for shared spectrum, the signal waveform can be jointly designed for localization and communication simultaneously, or two specific waveforms can be separately designed. Therefore, the theoretical analysis of the fundamental performance of ILAC and how to design the waveforms for ultra-high spectral efficiency deserve an in-depth investigation. On the other hand, for orthogonal spectrum usage, since localization and communication both benefit from wide signal bandwidth, a critical issue is that how to split the spectrum effectively to optimize the trade-off between localization and communication according to different QoS requirements.

4.2 Advanced signal processing for multi-tier ILAC networks

In multi-tier networks, the mobile terminals in different layers of networks can interact in different frequency bands with different links. In this case, how to maintain the wireless links dynamically and effectively is a critical issue for localization and communication. The integration of different frequency bands and dynamic resource management is a potential option to reconfigure and maintain the wireless connections, and the effective signal processing is an important factor for cross-layer information sharing. For ILAC, since the channel estimation units can be reused for location information extraction, the localization and communication systems can partly share some hardware, and the signal processing techniques for communication can be also exploited for location-related information measurements and estimation. Therefore, how to efficiently co-design the hardware architecture and signal processing techniques deserves further studies for ILAC.

4.3 Heterogenous networking

The future networks are more heterogenous than ever before, which operate on different frequency bands with different standards. A fundamental question is how to switch the protocols rapidly from one to another, while still ensuring the localization and communication performance, when the mobile terminals move quickly and suffer from different radio environments. For the distinct network architectures, each network layer employs different protocols, so a natural solution is to translate the protocols at gateways to interconnect different networks. However, such a sequence of protocol translations is inefficient, which renders the integrated protocol design that enables cross-layer, cross-module, and cross-node data transmission critical. Moreover, the network protocol design, especially in the physical and MAC layers, needs not only to consider the communication metrics, but also to be re-assessed from the localization perspective.

4.4 3D REMs and proactive RRM in complex environments

The proactive RRM is beneficial for communication in terms of cell selection, channel prediction, beam alignment and so on, but it will require high-accurate REMs. Currently, the studies on REMs modeling focus on the 2D scenarios targeting the TNs in outdoor environments. As the networks extend to 3D space, 3D REMs modeling in clutter environments like indoor and urban city is critical and challenging, which requires more accurate 3D localization in multi-path and NLoS environments. Although some effective multi-path and NLoS mitigation algorithms have been proposed in the literature, they are usually

quite complex and only feasible for remote localization systems rather than the self-localization systems where the computation and estimation need to be done by UEs. On the other hand, although some studies mainly focus on the localization in NLoS scenarios, by treating the signal reflectors as the virtual ANs, they are limited on the first-order reflection for simplicity. Therefore, in-depth investigations on localization algorithms for multi-path environments are still needed for more accurate and cost-effective localization. Furthermore, how to manage the radio resource according to the real-time location of the UE to dynamically maintain its communication performance deserves further investigations. On the other hand, although it is well-known that the accurate location information is beneficial for communications, the related fundamental metrics are needed to reveal the relation between different location accuracy and communication performance.

4.5 Intelligent ILAC

AI, or more specifically, machine learning, is one of the most promising technologies, to bring intelligence to wireless networks with complex radio conditions. Due to its capability in pattern recognition from raw data, machine learning can be effectively used for signal reconstruction, which is beneficial for both data decoding and location-related information measurements. In addition, since its learning ability in solving complex nonlinear problems, it can also be used to maintain the network performance dynamically and construct the user centric intelligent networks, which can autonomously manage resources, functions, and network control to sustain the high performance according to the real-time location of the mobile user. For ILAC, machine learning can be used in various perspectives, like waveform design, signal modulation/coding, resource allocation, to balance the performance trade-off between localization and communication, and create the 3D REMs to enhance the ILAC performance.

References

1 3rd Generation Partnership Project (3GPP). Location Services (LCS); Service description; Stage 1 (Release 15). TS 22.071. https://www.3gpp.org/ftp/Specs/archive/22_series/22.071

2 Zong B, Fan C, Wang X, et al. 6G technologies: key drivers, core requirements, system architectures, and enabling technologies. IEEE Veh Technol Mag, 2019, 14: 18–27

3 Saad W, Bennis M, Chen M. A vision of 6G wireless systems: applications, trends, technologies, and open research problems. IEEE Network, 2020, 34: 134–142

4 Strinati E C, Barbarossa S, Gonzalez-Jimenez J L, et al. 6G: the next frontier: from holographic messaging to artificial intelligence using subterahertz and visible light communication. IEEE Veh Technol Mag, 2019, 14: 42–50

5 6G Flagship. Key drivers and research challenges for 6G ubiquitous wireless intelligence (white paper). University of Oulu, 2019, http://jultika.oulu.fi/files/isbn9789526223544.pdf

6 Zhang L, Liang Y C, Niyato D. 6G visions: mobile ultra-broadband, super internet-of-things, and artificial intelligence. China Commun, 2019, 16: 1–14

7 ITU-T FG-NET-2030. Network 2030: a blueprint of technology, applications and market drivers towards the year 2030 and beyond. https://www.itu.int/en/ITU-T/focusgroups/net2030/Documents/WhitePaper.pdf

8 You X H, Wang C-X, Huang J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts. Sci China Inf Sci, 2021, 64: 110301

9 Chen M, Saad W, Yin C. Virtual reality over wireless networks: quality-of-service model and learning-based resource management. IEEE Trans Commun, 2018, 66: 5621–5635

10 Chang C M, Hsu C H, Hsu C F, et al. Performance measurements of virtual reality systems: quantifying the timing and positioning accuracy. In: Proceedings of the 24th ACM International Conference on Multimedia, 2016. 655–659

11 Tariq F, Khandaker M, Wong K K, et al. A speculative study on 6G. 2019. ArXiv:1902.06700

12 Witrisal K, Hinteregger S, Kulmer J, et al. High-accuracy positioning for indoor applications: RFID, UWB, 5G, and beyond. In: Proceedings of IEEE International Conference on RFID, 2016. 1–7

13 Witrisal K, Meissner P, Leitinger E, et al. High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend. IEEE Signal Process Mag, 2016, 33: 59–70

14 Kenney J B. Dedicated short-range communications (DSRC) standards in the United States. Proc IEEE, 2011, 99: 1162–1182

15 Chen S, Hu J, Shi Y, et al. Vehicle-to-everything (V2X) services supported by LTE-based systems and 5G. IEEE Comm Stand Mag, 2017, 1: 70–76

16 Mao Q, Hu F, Hao Q. Deep learning for intelligent wireless networks: a comprehensive survey. IEEE Commun Surv Tut, 2018, 20: 2595–2621

17 Asadi A, Wang Q, Mancuso V. A survey on device-to-device communication in cellular networks. IEEE Commun Surv Tut, 2014, 16: 1801–1819

18 Liu Y, Shi X, He S, et al. Prospective positioning architecture and technologies in 5G networks. IEEE Network, 2017, 31: 115–121

19 You H, Yin H, Wu H. On 6G and wide-area IoT. Chin J Int Things, 2020, 4: 3–11

20 Di Taranto R, Muppirisetty S, Raulefs R, et al. Location-aware communications for 5G networks: how location information can improve scalability, latency, and robustness of 5G. IEEE Signal Process Mag, 2014, 31: 102–112

21 Slock D. Location aided wireless communications. In: Proceedings of International Symposium on Communications, Control and Signal Processing, 2012. 1–6

22 Celebi H, Arslan H. Utilization of location information in cognitive wireless networks. IEEE Wireless Commun, 2007, 14: 6–13

23 Liu H, Darabi H, Banerjee P, et al. Survey of wireless indoor positioning techniques and systems. IEEE Trans Syst Man Cybern C, 2007, 37: 1067–1080

24 Yassin A, Nasser Y, Awad M, et al. Recent advances in indoor localization: a survey on theoretical approaches and applications. IEEE Commun Surv Tut, 2017, 19: 1327–1346

25 Zafari F, Gkelias A, Leung K K. A survey of indoor localization systems and technologies. IEEE Commun Surv Tut, 2019, 21: 2568–2599

26 Dardari D, Closas P, Djuric P M. Indoor tracking: theory, methods, and technologies. IEEE Trans Veh Technol, 2015, 64: 1263–1278

27 Laoudias C, Moreira A, Kim S, et al. A survey of enabling technologies for network localization, tracking, and navigation. IEEE Commun Surv Tut, 2018, 20: 3607–3644

28 Gustafsson F, Gunnarsson F. Mobile positioning using wireless networks: possibilities and fundamental limitations based on available wireless network measurements. IEEE Signal Process Mag, 2005, 22: 41–53

29 Sun G L, Chen J, Guo W, et al. Signal processing techniques in network-aided positioning: a survey of state-of-the-art positioning designs. IEEE Signal Process Mag, 2005, 22: 12–23

30 del Peral-Rosado J A, Raulefs R, Lopez-Salcedo J A, et al. Survey of cellular mobile radio localization methods: from 1G to 5G. IEEE Commun Surv Tut, 2018, 20: 1124–1148

31 Drane C, Macnaughtan M, Scott C. Positioning GSM telephones. IEEE Commun Mag, 1998, 36: 46–54, 59

32 Wax M, Kailath T. Decentralized processing in sensor arrays. IEEE Trans Acoust Speech Signal Process, 1985, 33: 1123–1129

33 Weiss A J. Direct position determination of narrowband radio frequency transmitters. IEEE Signal Process Lett, 2004, 11: 513–516

34 Bialer O, Raphaeli D, Weiss A J. Maximum-likelihood direct position estimation in dense multipath. IEEE Trans Veh Technol, 2013, 62: 2069–2079

35 Zekavat R, Buehrer R M. Handbook of Position Location: Theory, Practice and Advances. Hoboken: John Wiley & Sons, 2011

36 Yang Z, Zhou Z, Liu Y. From RSSI to CSI. ACM Comput Surv, 2013, 46: 1–32

37 Kumar P, Reddy L, Varma S. Distance measurement and error estimation scheme for RSSI based localization in wireless sensor networks. In: Proceedings of IEEE International Conference on Wireless Communication and Sensor Networks (WCSN), 2009. 1–4

38 Yang J, Chen Y. Indoor localization using improved RSS-based lateration methods. In: Proceedings of IEEE Global Telecommunications Conference, 2009. 1–6

39 Chan Y T, Tsui W Y, So H C, et al. Time-of-arrival based localization under NLOS conditions. IEEE Trans Veh Technol, 2006, 55: 17–24

40 Ke W, Wu L. Constrained least squares algorithm for TOA-based mobile location under NLOS environments. In: Proceedings of IEEE International Conference on Wireless Communications, Networking and Mobile Computing, 2009. 1–4

41 Guvenc I, Chong C C. A survey on TOA based wireless localization and NLOS mitigation techniques. IEEE Commun Surv Tutorials, 2009, 11: 107–124

42 Ledergerber A, Hamer M, D'Andrea R. A robot self-localization system using one-way ultra-wideband communication. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2015. 3131–3137

43 Wahab A A, Khattab A, Fahmy Y A. Two-way TOA with limited dead reckoning for GPS-free vehicle localization using single RSU. In: Proceedings of IEEE International Conference on ITS Telecommunications (ITST), 2013. 244–249

44 Aditya S, Molisch A F, Behairy H M. A survey on the impact of multipath on wideband time-of-arrival based localization. Proc IEEE, 2018, 106: 1183–1203

45 Kaune R, Hörst J, Koch W. Accuracy analysis for TDOA localization in sensor networks. In: Proceedings of IEEE International Conference on Information Fusion, 2011. 1–8

46 Jung S Y, Hann S, Park C S. TDOA-based optical wireless indoor localization using LED ceiling lamps. IEEE Trans Consumer Electron, 2011, 57: 1592–1597

47 Kaune R. Accuracy studies for TDOA and TOA localization. In: Proceedings of IEEE International Conference on Information Fusion, 2012. 408–415

48 Kumar S, Gil S, Katabi D, et al. Accurate indoor localization with zero start-up cost. In: Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, 2014. 483–494

49 Wang Y, Ho K C. An asymptotically efficient estimator in closed-form for 3-D AOA localization using a sensor network. IEEE Trans Wireless Commun, 2015, 14: 6524–6535

50 Hata M. Empirical formula for propagation loss in land mobile radio services. IEEE Trans Veh Technol, 1980, 29: 317–325

51 Feng C, Au W S A, Valaee S, et al. Received-signal-strength-based indoor positioning using compressive sensing. IEEE Trans Mobile Comput, 2012, 11: 1983–1993

52 Honkavirta V, Perala T, Ali-Loytty S, et al. A comparative survey of WLAN location fingerprinting methods. In: Proceedings of IEEE Workshop on Positioning, Navigation and Communication, 2009. 243–251

53 Yang C, Shao H-R. WiFi-based indoor positioning. IEEE Commun Mag, 2015, 53: 150–157

54 Vo Q D, De P. A survey of fingerprint-based outdoor localization. IEEE Commun Surv Tut, 2016, 18: 491–506

55 He S, Chan S H G. Wi-Fi fingerprint-based indoor positioning: recent advances and comparisons. IEEE Commun Surv Tut, 2016, 18: 466–490

56 Khalajmehrabadi A, Gatsis N, Akopian D. Modern WLAN fingerprinting indoor positioning methods and deployment challenges. IEEE Commun Surv Tut, 2017, 19: 1974–2002

57 Bahl P, Padmanabhan V N. RADAR: an in-building RF-based user location and tracking system. In: Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies, 2000. 775–784

58 Seco F, Jiménez A R, Prieto C, et al. A survey of mathematical methods for indoor localization. In: Proceedings of IEEE International Symposium on Intelligent Signal Processing, 2009. 9–14

59 Brida P, Duha J, Krasnovsky M. On the accuracy of weighted proximity based localization in wireless sensor networks. In: Proceedings of Personal Wireless Communications, 2007. 423–432

60 He T, Huang C, Blum B M, et al. Range-free localization schemes for large scale sensor networks. In: Proceedings of the

9th Annual International Conference on Mobile Computing and Networking, 2003. 81–95

61 Li M, Liu Y H. Rendered path: range-free localization in anisotropic sensor networks with holes. IEEE/ACM Trans Netw, 2010, 18: 320–332

62 Chan Y T, Hang H Y C, Ching P C. Exact and approximate maximum likelihood localization algorithms. IEEE Trans Veh Technol, 2006, 55: 10–16

63 Chang C, Sahai A. Estimation bounds for localization. In: Proceedings of IEEE 1st Annual Communications Society Conference on Sensor and Ad Hoc Communications and Networks, 2004. 415–424

64 Shen Y, Win M Z. Fundamental limits of wideband localization-Part I: a general framework. IEEE Trans Inform Theor, 2010, 56: 4956–4980

65 Shen Y, Wymeersch H, Win M Z. Fundamental limits of wideband localization-Part II: cooperative networks. IEEE Trans Inform Theor, 2010, 56: 4981–5000

66 Marks M, Niewiadomska-Szynkiewicz E. Two-phase stochastic optimization to sensor network localization. In: Proceedings of IEEE International Conference on Sensor Technologies and Applications, 2007. 134–139

67 Chuang P J, Wu C P. An effective PSO-based node localization scheme for wireless sensor networks. In: Proceedings of IEEE International Conference on Parallel and Distributed Computing, Applications and Technologies, 2008. 187–194

68 Saab K K, Saab S S. Application of an optimal stochastic Newton-Raphson technique to triangulation-based localization systems. In: Proceedings of IEEE/ION Position, Location and Navigation Symposium (PLANS), 2016. 981–986

69 Béjar B, Belanovic P, Zazo S. Distributed Gauss-Newton method for localization in Ad-Hoc networks. In: Proceedings of IEEE Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers, 2010. 1452–1454

70 Torrieri D J. Statistical theory of passive location systems. IEEE Trans Aerosp Electron Syst, 1984, 20: 183–198

71 Caffery J. A new approach to the geometry of TOA location. In: Proceedings of IEEE Vehicular Technology Conference Fall, 2000. 1943–1949

72 Chan Y T, Ho K C. A simple and efficient estimator for hyperbolic location. IEEE Trans Signal Process, 1994, 42: 1905–1915

73 Fang B T. Simple solutions for hyperbolic and related position fixes. IEEE Trans Aerosp Electron Syst, 1990, 26: 748–753

74 Nardone S C, Graham M L. A closed-form solution to bearings-only target motion analysis. IEEE J Ocean Eng, 1997, 22: 168–178

75 Gavish M, Weiss A J. Performance analysis of bearing-only target location algorithms. IEEE Trans Aerosp Electron Syst, 1992, 28: 817–828

76 Chen J C, Hudson R E, Yao K. Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field. IEEE Trans Signal Process, 2002, 50: 1843–1854

77 Navidi W, Murphy J W S, Hereman W. Statistical methods in surveying by trilateration. Comput Stat Data Anal, 1998, 27: 209–227

78 Lee J-Y, Scholtz R A. Ranging in a dense multipath environment using an UWB radio link. IEEE J Sel Areas Commun, 2002, 20: 1677–1683

79 Denis B, Keignart J, Daniele N. Impact of NLOS propagation upon ranging precision in UWB systems. In: Proceedings of IEEE Conference on Ultra Wideband Systems and Technologies, 2003. 379–383

80 Wu Z H, Han Y, Chen Y, et al. A time-reversal paradigm for indoor positioning system. IEEE Trans Veh Technol, 2015, 64: 1331–1339

81 Zhang F, Chen C, Wang B, et al. WiBall: a time-reversal focusing ball method for decimeter-accuracy indoor tracking. IEEE Internet Things J, 2018, 5: 4031–4041

82 Gentner C, Jost T, Wang W, et al. Multipath assisted positioning with simultaneous localization and mapping. IEEE Trans Wireless Commun, 2016, 15: 6104–6117

83 Gustafsson F. Statistical Sensor Fusion. Lund: Studentlitteratur, 2010

84 Gunnarsson F, Lindsten F, Carlsson N. Particle filtering for network-based positioning terrestrial radio networks. In: Proceedings of IET Conference on Data Fusion & Target Tracking: Algorithms and Applications, 2014. 1–7

85 Dauwels J. Computing bayesian cramer-rao bounds. In: Proceedings of IEEE International Symposium on Information Theory, 2005. 425–429

86 Weinstein E, Weiss A J. A general class of lower bounds in parameter estimation. IEEE Trans Inform Theor, 1988, 34: 338–342

87 Zeira A, Schultheiss P M. Realizable lower bounds for time delay estimation. 2. Threshold phenomena. IEEE Trans Signal Process, 1994, 42: 1001–1007

88 Gioia C, Borio D. Stand-alone and hybrid positioning using asynchronous pseudolites. Sensors, 2015, 15: 166–193

89 Duffett-Smith P J, Tarlow B. E-GPS: indoor mobile phone positioning on GSM and WCDMA. In: Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation, 2005. 2762–2768

90 Rappaport T S, Reed J H, Woerner B D. Position location using wireless communications on highways of the future. IEEE Commun Mag, 1996, 34: 33–41

91 Spirito M A, Poykko S, Knuuttila O. Experimental performance of methods to estimate the location of legacy handsets in GSM. In: Proceedings of IEEE Vehicular Technology Conference Fall, 2001. 2716–2720

92 Lin D B, Juang R T. Mobile location estimation based on differences of signal attenuations for GSM systems. IEEE Trans Veh Technol, 2005, 54: 1447–1454

93 Borkowski J, Niemela J, Lempiainen J. Enhanced performance of Cell ID+ RTT by implementing forced soft handover algorithm. In: Proceedings of IEEE Vehicular Technology Conference Fall, 2004. 3545–3549

94 Ghosh A, Ratasuk R, Mondal B, et al. LTE-advanced: next-generation wireless broadband technology . IEEE Wireless Commun, 2010, 17: 10–22

95 Sadowski J. TDOA navigation using CDMA2000 signals–experimental results. In: Proceedings of IEEE Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), 2014. 265–271

96 del Peral-Rosado J A, López-Salcedo J A, Seco-Granados G, et al. Joint maximum likelihood time-delay estimation for LTE positioning in multipath channels. EURASIP J Adv Signal Process, 2014, 2014: 33

97 Driusso M, Babich F, Knutti F, et al. Estimation and tracking of LTE signals time of arrival in a mobile multipath environment. In: Proceedings of International Symposium on Image and Signal Processing and Analysis (ISPA), 2015. 276–281

98 Cherian S S, Rudrapatna A N. LTE location technologies and delivery solutions. Bell Labs Tech J, 2013, 18: 175–194

99 3rd Generation Partnership Project (3GPP). Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Positioning Protocol (LPP) (Release 9). TS 36.355. https://www.3gpp.org/ftp/Specs/archive/36_series/36.355

100 Ericsson. Positioning with LTE. White paper. http://sharetechnote.com/Docs/WP-LTE-positioning.pdf

101 Centenaro M, Vangelista L, Zanella A, et al. Long-range communications in unlicensed bands: the rising stars in the IoT and smart city scenarios. IEEE Wireless Commun, 2016, 23: 60–67

102 Liu H, Gan Y, Yang J, et al. Push the limit of WiFi based localization for smartphones. In: Proceedings of the 18th Annual ACM International Conference on Mobile Computing and Networking, 2012. 305–316

103 Liu H H, Yang Y N. WiFi-based indoor positioning for multi-floor environment. In: Proceedings of IEEE Region 10 Conference, 2011. 597–601

104 Davidson P, Piche R. A survey of selected indoor positioning methods for smartphones. IEEE Commun Surv Tut, 2017, 19: 1347–1370

105 Zou H, Jiang H, Lu X, et al. An online sequential extreme learning machine approach to WiFi based indoor positioning. In: Proceedings of IEEE World Forum on Internet of Things (WF-IoT), 2014. 111–116

106 Kotaru M, Joshi K, Bharadia D, et al. Spotfi: decimeter level localization using WiFi. In: Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication, 2015. 269–282

107 Oppermann I, Hämäläinen M, Iinatti J. UWB: Theory and Applications. Hoboken: John Wiley & Sons, 2005

108 Sahinoglu Z, Gezici S, Guvenc I. Ultra-wideband Positioning Systems. Cambridge: Cambridge University Press, 2008

109 Gezici S, Tian Z, Giannakis G B, et al. Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks. IEEE Signal Process Mag, 2005, 22: 70–84

110 Mekonnen Z W, Slottke E, Luecken H, et al. Constrained maximum likelihood positioning for UWB based human motion tracking. In: Proceedings of IEEE International Conference on Indoor Positioning and Indoor Navigation, 2010. 1–10

111 Alarifi A, Al-Salman A M, Alsaleh M, et al. Ultra wideband indoor positioning technologies: analysis and recent advances. Sensors, 2016, 16: 707

112 Hassan N U, Naeem A, Pasha M A, et al. Indoor positioning using visible LED lights. ACM Comput Surv, 2015, 48: 1–32

113 Zhuang Y, Hua L, Qi L, et al. A survey of positioning systems using visible LED lights. IEEE Commun Surv Tut, 2018, 20: 1963–1988

114 Kail G, Maechler P, Preyss N, et al. Robust asynchronous indoor localization using LED lighting. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014. 1866–1870

115 Zhu B, Cheng J, Wang Y, et al. Three-dimensional VLC positioning based on angle difference of arrival with arbitrary tilting angle of receiver. IEEE J Sel Areas Commun, 2018, 36: 8–22

116 Durrant-Whyte H, Bailey T. Simultaneous localization and mapping: part I. IEEE Robot Automat Mag, 2006, 13: 99–110

117 Bailey T, Durrant-Whyte H. Simultaneous localization and mapping (SLAM): part II. IEEE Robot Automat Mag, 2006, 13: 108–117

118 Cadena C, Carlone L, Carrillo H, et al. Past, present, and future of simultaneous localization and mapping: toward the robust-perception age. IEEE Trans Robot, 2016, 32: 1309–1332

119 Patwari N, Ash J N, Kyperountas S, et al. Locating the nodes: cooperative localization in wireless sensor networks. IEEE Signal Process Mag, 2005, 22: 54–69

120 Mao G, Fidan B, Anderson B D O. Wireless sensor network localization techniques. Comput Networks, 2007, 51: 2529–2553

121 Wymeersch H, Lien J, Win M Z. Cooperative localization in wireless networks. Proc IEEE, 2009, 97: 427–450

122 Niculescu D, Nath B. Ad hoc positioning system (APS). In: Proceedings of IEEE Global Telecommunications Conference, 2001. 2926–2931

123 Shang Y, Rumi W, Zhang Y, et al. Localization from connectivity in sensor networks. IEEE Trans Parallel Distrib Syst, 2004, 15: 961–974

124 Vivekanandan V, Wong V W S. Concentric anchor beacon localization algorithm for wireless sensor networks. IEEE Trans Veh Technol, 2007, 56: 2733–2744

125 Ihler A T, Fisher J W, Moses R L, et al. Nonparametric belief propagation for self-localization of sensor networks. IEEE J Sel Areas Commun, 2005, 23: 809–819

126 Radnosrati K, Gunnarsson F, Gustafsson F. New trends in radio network positioning. In: Proceedings of IEEE International Conference on Information Fusion (Fusion), 2015. 492–498

127 de Angelis A, Fischione C. Mobile node localization via pareto optimization: algorithm and fundamental performance limitations. IEEE J Sel Areas Commun, 2015, 33: 1288–1303

128 Bullock J B, Chowdhary M, Rubin D, et al. Continuous indoor positioning using GNSS, Wi-Fi, and MEMS dead reckoning. In: Proceedings of the 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012), 2012. 2408–2416

129 ITU-R M.2083 report. IMT vision framework and overall objectives of the future development of IMT for 2020 and beyond. https://www.itu.int/rec/R-REC-M.2083-0-201509-I/en

130 Keating R, Säily M, Hulkkonen J, et al. Overview of positioning in 5G new radio. In: Proceedings of International Symposium on Wireless Communication Systems (ISWCS), 2019. 320–324

131 3rd Generation Partnership Project (3GPP). Study on Positioning Use Cases; Stage 1 (Release 16). TR 22.872. https://www.3gpp.org/ftp/Specs/archive/22_series/22.872

132 3rd Generation Partnership Project (3GPP). Study on NR positioning support (Release 16). TR 38.855. https://www.3gpp.org/ftp/Specs/archive/38_series/38.855

133 3rd Generation Partnership Project (3GPP). Service requirements for next generation new services and markets; Stage 1 (Release 15). TS 22.261. https://www.3gpp.org/ftp/Specs/archive/22_series/22.261

134 3rd Generation Partnership Project (3GPP). Study on Communication for Automation in Vertical domains (CAV) (Release 15). TR 22.804. https://www.3gpp.org/ftp/Specs/archive/22_series/22.804

135 Heath R W, Gonzalez-Prelcic N, Rangan S, et al. An overview of signal processing techniques for millimeter wave MIMO systems. IEEE J Sel Top Signal Process, 2016, 10: 436–453

136 Lemic F, Martin J, Yarp C, et al. Localization as a feature of mmWave communication. In: Proceedings of IEEE International Wireless Communications and Mobile Computing Conference (IWCMC), 2016. 1033–1038

137 Larsson E G, Edfors O, Tufvesson F, et al. Massive MIMO for next generation wireless systems. IEEE Commun Mag, 2014, 52: 186–195

138 Rusek F, Persson D, Lau B K, et al. Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Signal Process Mag, 2013, 30: 40–60

139 Lu L, Li G Y, Swindlehurst A L, et al. An overview of massive MIMO: benefits and challenges. IEEE J Sel Top Signal Process, 2014, 8: 742–758

140　Jungnickel V, Manolakis K, Zirwas W, et al. The role of small cells, coordinated multipoint, and massive MIMO in 5G. IEEE Commun Mag, 2014, 52: 44–51

141　Shahmansoori A, Garcia G E, Destino G, et al. 5G position and orientation estimation through millimeter wave MIMO. In: Proceedings of IEEE Globecom Workshops (GC Wkshps), 2015. 1–6

142　Abu-Shaban Z, Zhou X, Abhayapala T, et al. Error bounds for uplink and downlink 3D localization in 5G millimeter wave systems. IEEE Trans Wireless Commun, 2018, 17: 4939–4954

143　Sur S, Venkateswaran V, Zhang X, et al. 60 GHz indoor networking through flexible beams: a link-level profiling. In: Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, 2015. 71–84

144　Perahia E, Cordeiro C, Park M, et al. IEEE 802.11 ad: defining the next generation multi-Gbps Wi-Fi. In: Proceedings of IEEE Consumer Communications and Networking Conference, 2010. 1–5

145　Baykas T, Sum C S, Lan Z, et al. IEEE 802.15.3c: the first IEEE wireless standard for data rates over 1 Gb/s. IEEE Commun Mag, 2011, 49: 114–121

146　Singh S, Mudumbai R, Madhow U. Interference analysis for highly directional 60-GHz mesh networks: the case for rethinking medium access control. IEEE/ACM Trans Netw, 2011, 19: 1513–1527

147　Deng H, Sayeed A. Mm-wave MIMO channel modeling and user localization using sparse beamspace signatures. In: Proceedings of IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2014. 130–134

148　Saloranta J, Destino G. On the utilization of MIMO-OFDM channel sparsity for accurate positioning. In: Proceedings of European Signal Processing Conference (EUSIPCO), 2016. 748–752

149　Qingling Z, Li J. Rain attenuation in millimeter wave ranges. In: Proceedings of IEEE International Symposium on Antennas, Propagation & EM Theory, 2006. 1–4

150　Zeng Y, Zhang R. Millimeter wave MIMO with lens antenna array: a new path division multiplexing paradigm. IEEE Trans Commun, 2016, 64: 1557–1571

151　Lee Y U, Choi J, Song I, et al. Distributed source modeling and direction-of-arrival estimation techniques. IEEE Trans Signal Process, 1997, 45: 960–969

152　Astely D, Ottersten B. The effects of local scattering on direction of arrival estimation with MUSIC. IEEE Trans Signal Process, 1999, 47: 3220–3234

153　Wan L, Han G, Jiang J, et al. DOA estimation for coherently distributed sources considering circular and noncircular signals in massive MIMO systems. IEEE Syst J, 2017, 11: 41–49

154　Hu A, Lv T, Gao H, et al. An ESPRIT-based approach for 2-D localization of incoherently distributed sources in massive MIMO systems. IEEE J Sel Top Signal Process, 2014, 8: 996–1011

155　Wen F, Wymeersch H, Peng B, et al. A survey on 5G massive MIMO localization. Digital Signal Process, 2019, 94: 21–28

156　Haardt M, Pesavento M, Roemer F, et al. Subspace methods and exploitation of special array structures. Academic Press Library Signal Process, 2014, 3: 651–717

157　Tsai C R, Liu Y H, Wu A Y. Efficient compressive channel estimation for millimeter-wave large-scale antenna systems. IEEE Trans Signal Process, 2018, 66: 2414–2428

158　Valaee S, Champagne B, Kabal P. Parametric localization of distributed sources. IEEE Trans Signal Process, 1995, 43: 2144–2153

159　Shahbazpanahi S, Valaee S, Bastani M H. Distributed source localization using ESPRIT algorithm. IEEE Trans Signal Process, 2001, 49: 2169–2178

160　Lee J, Song I, Kwon H, et al. Low-complexity estimation of 2D DOA for coherently distributed sources. Signal Process, 2003, 83: 1789–1802

161　Trump T, Ottersten B. Estimation of nominal direction of arrival and angular spread using an array of sensors. Signal Process, 1996, 50: 57–69

162　Ottersten B, Stoica P, Roy R. Covariance matching estimation techniques for array signal processing applications. Digital Signal Processing, 1998, 8: 185–210

163　Zoubir A, Wang Y, ChargÉ P. Efficient subspace-based estimator for localization of multiple incoherently distributed sources. IEEE Trans Signal Process, 2008, 56: 532–542

164　Tapio M. Direction and spread estimation of spatially distributed signals via the power azimuth spectrum. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002. 3005–3008

165　Yang J, Jin S, Han Y, et al. 3-D position and velocity estimation in 5G mmWave CRAN with lens antenna arrays. In: Proceedings of IEEE Vehicular Technology Conference Fall, 2019. 1–6

166　Yang J, Jin S, Wen C K, et al. 3-D positioning and environment mapping for mmWave communication systems. 2019. ArXiv:1908.04142

167　Vari M, Cassioli D. mmWaves RSSI indoor network localization. In: Proceedings of IEEE International Conference on Communications Workshops (ICC), 2014. 127–132

168　Wei T, Zhang X. mtrack: high-precision passive tracking using millimeter wave radios. In: Proceedings of the 21st Annual International Conference on Mobile Computing and Networking, 2015. 117–129

169　Shahmansoori A, Garcia G E, Destino G, et al. Position and orientation estimation through millimeter-wave MIMO in 5G systems. IEEE Trans Wireless Commun, 2018, 17: 1822–1835

170　Li J, Conan J, Pierre S. Position location of mobile terminal in wireless MIMO communication systems. J Commun Netw, 2007, 9: 254–264

171　Savic V, Larsson E G. Fingerprinting-based positioning in distributed massive MIMO systems. In: Proceedings of IEEE Vehicular Technology Conference Fall, 2015. 1–5

172　Garcia N, Wymeersch H, Larsson E G, et al. Direct localization for massive MIMO. IEEE Trans Signal Process, 2017, 65: 2475–2487

173　Koivisto M, Costa M, Werner J, et al. Joint device positioning and clock synchronization in 5G ultra-dense networks. IEEE Trans Wireless Commun, 2017, 16: 2866–2881

174　Koivisto M, Hakkarainen A, Costa M, et al. High-efficiency device positioning and location-aware communications in dense 5G networks. IEEE Commun Mag, 2017, 55: 188–195

175　Wymeersch H, Seco-Granados G, Destino G, et al. 5G mmWave positioning for vehicular networks. IEEE Wireless Commun, 2017, 24: 80–86

176　Boccardi F, Heath R W, Lozano A, et al. Five disruptive technology directions for 5G. IEEE Commun Mag, 2014, 52: 74–80

177 Lin Y D, Hsu Y C. Multihop cellular: a new architecture for wireless communications. In: Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies, 2000. 1273–1282

178 Lei L, Zhong Z D, Lin C, et al. Operator controlled device-to-device communications in LTE-advanced networks. IEEE Wireless Commun, 2012, 19: 96–104

179 Pratas N K, Popovski P. Low-rate machine-type communication via wireless device-to-device (D2D) links. 2013. ArXiv:1305.6783

180 Bao X, Lee U, Rimac I, et al. Dataspotting: offloading cellular traffic via managed device-to-device data transfer at data spots. SIGMOBILE Mob Comput Commun Rev, 2010, 14: 37–39

181 Tehrani M N, Uysal M, Yanikomeroglu H. Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions. IEEE Commun Mag, 2014, 52: 86–92

182 Baldemair R, Irnich T, Balachandran K, et al. Ultra-dense networks in millimeter-wave frequencies. IEEE Commun Mag, 2015, 53: 202–208

183 Kamel M, Hamouda W, Youssef A. Ultra-dense networks: a survey. IEEE Commun Surv Tut, 2016, 18: 2522–2545

184 Ding M, López-Pérez D, Mao G, et al. Will the area spectral efficiency monotonically grow as small cells go dense? In: Proceedings of IEEE Global Communications Conference (GLOBECOM), 2015. 1–7

185 Werner J, Costa M, Hakkarainen A, et al. Joint user node positioning and clock offset estimation in 5G ultra-dense networks. In: Proceedings of IEEE Global Communications Conference (GLOBECOM), 2015. 1–7

186 Koivisto M, Talvitie J, Costa M, et al. Joint cmWave-based multiuser positioning and network synchronization in dense 5G networks. In: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), 2018. 1–6

187 Dammann A, Agapiou G, Bastos J, et al. WHERE2 location aided communications. In: Proceedings of European Wireless Conference, 2013. 1–8

188 Kela P, Costa M, Turkka J, et al. Location based beamforming in 5G ultra-dense networks. In: Proceedings of IEEE Vehicular Technology Conference Fall, 2016. 1–7

189 Garcia N, Wymeersch H, Ström E G, et al. Location-aided mm-wave channel estimation for vehicular communication. In: Proceedings of IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2016. 1–5

190 Abdelreheem A, Mohamed E M, Esmaiel H. Millimeter wave location-based beamforming using compressive sensing. In: Proceedings of IEEE 28th International Conference on Microelectronics (ICM), 2016. 213–216

191 Abdelreheem A, Mohamed E M, Esmaiel H. Location-based millimeter wave multi-level beamforming using compressive sensing. IEEE Commun Lett, 2018, 22: 185–188

192 Yilmaz H, Tugcu T, Alagöz F, et al. Radio environment map as enabler for practical cognitive radio networks. IEEE Commun Mag, 2013, 51: 162–169

193 Bi S, Lyu J, Ding Z, et al. Engineering radio maps for wireless resource management. IEEE Wireless Commun, 2019, 26: 133–141

194 Zeng Y, Xu X, Jin S, et al. Simultaneous navigation and radio mapping for cellular-connected UAV with deep reinforcement learning. 2020. ArXiv:2003.07574

195 Üreten S, Yongaçoğlu A, Petriu E. A comparison of interference cartography generation techniques in cognitive radio networks. In: Proceedings of IEEE International Conference on Communications (ICC), 2012. 1879–1883

196 Yilmaz H B, Tugcu T. Location estimation-based radio environment map construction in fading channels. Wirel Commun Mob Comput, 2015, 15: 561–570

197 Nevat I, Peters G W, Collings I B. Location-aware cooperative spectrum sensing via Gaussian processes. In: Proceedings of Australian Communications Theory Workshop (AusCTW), 2012. 19–24

198 Sand S, Tanbourgi R, Mensing C, et al. Position aware adaptive communication systems. In: Proceedings of IEEE Conference Record of the 43rd Asilomar Conference on Signals, Systems and Computers, 2009. 73–77

199 Hakkarainen A, Werner J, Costa M, et al. High-efficiency device localization in 5G ultra-dense networks: prospects and enabling technologies. In: Proceedings of IEEE Vehicular Technology Conference Fall, 2015. 1–5

200 Irmer R, Droste H, Marsch P, et al. Coordinated multipoint: concepts, performance, and field trial results. IEEE Commun Mag, 2011, 49: 102–111

201 Ko Y B, Vaidya N H. Location-aided routing (LAR) in mobile ad hoc networks. Wireless Networks, 2000, 6: 307–321

202 Sun M T, Feng W, Lai T H. Location aided broadcast in wireless ad hoc networks. In: Proceedings of IEEE Global Telecommunications Conference, 2001. 2842–2846

203 Bai Y, An J, Zhang H B. Location aided probabilistic broadcast algorithm for mobile ad-hoc network routing. J China Universities Posts Telecommun, 2017, 24: 66–71

204 Ko Y B, Vaidya N H. Geocasting in mobile ad hoc networks: location-based multicast algorithms. In: Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, 1999. 101–110

205 Shih T F, Shih C C, Chen C L. Location-based multicast routing protocol for mobile ad hoc networks. WSEAS Trans Comput, 2008, 7: 1270–1279

206 Zhao H, Su W. Cooperative wireless multicast: performance analysis and power/location optimization. IEEE Trans Wireless Commun, 2010, 9: 2088–2100

207 Jiang J, Zhang J, Zhang L, et al. Passive location resource scheduling based on an improved genetic algorithm. Sensors, 2018, 18: 2093

208 Girgis A M, El-Keyi A, Nafie M, et al. Proactive location-based scheduling of delay-constrained traffic over fading channels. In: Proceedings of IEEE Vehicular Technology Conference Fall, 2016. 1–6

209 Zhu C, Leung V C M, Yang L T, et al. Collaborative location-based sleep scheduling for wireless sensor networks integratedwith mobile cloud computing. IEEE Trans Comput, 2015, 64: 1844–1856

210 Al Nuaimi K, Mohamed N, Al Nuaimi M, et al. A survey of load balancing in cloud computing: challenges and algorithms. In: Proceedings of IEEE Second Symposium on Network Cloud Computing and Applications, 2012. 137–142

211 Cybenko G. Dynamic load balancing for distributed memory multiprocessors. J Parallel Distrib Comput, 1989, 7: 279–301

212 Wang Y, Basnayaka D A, Wu X, et al. Optimization of load balancing in hybrid LiFi/RF networks. IEEE Trans Commun, 2017, 65: 1708–1720

213 Wang Y L, Haas H. Dynamic load balancing with handover in hybrid Li-Fi and Wi-Fi networks. J Lightwave Technol, 2015, 33: 4671–4682

214 Ma W, Zhang L, Wu Z. Location information-aided load balancing design for hybrid LiFi and WiFi networks. In: Proceedings of IEEE International Conference on Computing, Networking and Communications (ICNC), 2019. 413–417

215 Yanmaz E, Tonguz O K. Location dependent dynamic load balancing. In: Proceedings of IEEE Global Telecommunications Conference, 2005. 5

216 Cadger F, Curran K, Santos J, et al. A survey of geographical routing in wireless ad-hoc networks. IEEE Commun Surv Tut, 2013, 15: 621–653

217 Kuhn F, Wattenhofer R, Zollinger A. An algorithmic approach to geographic routing in ad hoc and sensor networks. IEEE/ACM Trans Netw, 2008, 16: 51–62

218 Shah S H, Nahrstedt K. Predictive location-based QoS routing in mobile ad hoc networks. In: Proceedings of IEEE International Conference on Communications, 2002. 1022–1027

219 Stojmenovic I, Datta S. Power and cost aware localized routing with guaranteed delivery in unit graph based ad hoc networks. Wirel Commun Mob Comput, 2004, 4: 175–188

220 Abou-Zeid H, Hassanein H. Toward green media delivery: location-aware opportunities and approaches. IEEE Wireless Commun, 2014, 21: 38–46

221 Yao J, Kanhere S S, Hassan M. Improving QoS in high-speed mobility using bandwidth maps. IEEE Trans Mobile Comput, 2012, 11: 603–617

222 Wang X, Chen M, Taleb T, et al. Cache in the air: exploiting content caching and delivery techniques for 5G systems. IEEE Commun Mag, 2014, 52: 131–139

223 Gautam N, Petander H, Noel J. A comparison of the cost and energy efficiency of prefetching and streaming of mobile video. In: Proceedings of the 5th Workshop on Mobile Video, 2013. 7–12

224 Ghatak G, Koirala R, de Domenico A, et al. Positioning data-rate trade-off in mm-wave small cells and service differentiation for 5G networks. In: Proceedings of IEEE Vehicular Technology Conference Spring, 2018. 1–5

225 Jeong S, Simeone O, Haimovich A, et al. Beamforming design for joint localization and data transmission in distributed antenna system. IEEE Trans Veh Technol, 2015, 64: 62–76

226 Destino G, Wymeersch H. On the trade-off between positioning and data rate for mm-wave communication. In: Proceedings of IEEE International Conference on Communications Workshops (ICC Workshops), 2017. 797–802

227 Destino G, Saloranta J, Wymeersch H, et al. Impact of imperfect beam alignment on the rate-positioning trade-off. In: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), 2018. 1–5

228 Kumar D, Saloranta J, Destino G, et al. On trade-off between 5G positioning and mmWave communication in a multi-user scenario. In: Proceedings of International Conference on Localization and GNSS (ICL-GNSS), 2018. 1–5

229 Proakis J G, Salehi M. Digital Communications. New York: McGraw-hill, 2001

230 Dardari D, Conti A, Ferner U, et al. Ranging with ultrawide bandwidth signals in multipath environments. Proc IEEE, 2009, 97: 404–426

231 Dammann A, Jost T, Raulefs R, et al. Optimizing waveforms for positioning in 5G. In: Proceedings of IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2016. 1–5

232 Raulefs R, Dammann A, Jost T, et al. The 5G localization waveform. In: Proceedings of ETSI Workshop on Future Radio Technologies Focusing on Air Interfaces, 2016

233 Chong C C, Watanabe F, Win M Z. Effect of bandwidth on UWB ranging error. In: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), 2007. 1559–1564

234 Zeng Y, Wu Q, Zhang R. Accessing from the sky: a tutorial on UAV communications for 5G and beyond. Proc IEEE, 2019, 107: 2327–2375

235 Zeng Y, Zhang R, Lim T J. Wireless communications with unmanned aerial vehicles: opportunities and challenges. IEEE Commun Mag, 2016, 54: 36–42

236 Zeng Y, Lyu J, Zhang R. Cellular-connected UAV: potential, challenges, and promising technologies. IEEE Wireless Commun, 2019, 26: 120–127

237 3rd Generation Partnership Project (3GPP). Study on Enhanced LTE Support for Aerial Vehicles. RP-170779. https://www.3gpp.org/ftp/tsgran/tsgran/TSG_RAN/TSGR_75/Docs//RP-170779.zip

238 3rd Generation Partnership Project (3GPP). Technical specication group radio access network: study on enhanced LTE support for aerial vehicles. TR 36.777. https://www.3gpp.org/ftp/Specs/archive/36_series/36.777

239 Yoo C S, Ahn I K. Low cost GPS/INS sensor fusion system for UAV navigation. In: Proceedings of IEEE Digital Avionics Systems Conference, 2003

240 Nemra A, Aouf N. Robust INS/GPS sensor fusion for UAV localization using SDRE nonlinear filtering. IEEE Sens J, 2010, 10: 789–798

241 Lu Y, Xue Z, Xia G S, et al. A survey on vision-based UAV navigation. Geo-spatial Inf Sci, 2018, 21: 21–32

242 Queralta J P, Almansa C M, Schiano F, et al. UWB-based system for UAV localization in GNSS-denied environments: characterization and dataset. 2020. ArXiv:2003.04380

243 Tiemann J, Schweikowski F, Wietfeld C. Design of an UWB indoor-positioning system for UAV navigation in GNSS-denied environments. In: Proceedings of IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2015. 1–7

244 Tiemann J, Wietfeld C. Scalable and precise multi-UAV indoor navigation using TDOA-based UWB localization. In: Proceedings of IEEE International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2017. 1–7

245 Benini A, Mancini A, Longhi S. An IMU/UWB/vision-based extended kalman filter for mini-UAV localization in indoor environment using 802.15.4a wireless sensor network. J Intell Robot Syst, 2013, 70: 461–476

246 Stojkoska B R, Palikrushev J, Trivodaliev K, et al. Indoor localization of unmanned aerial vehicles based on RSSI. In: Proceedings of the 17th International Conference on Smart Technologies, 2017. 120–125

247 Tian X, Song Z, Jiang B, et al. HiQuadLoc: a RSS fingerprinting based indoor localization system for quadrotors. IEEE Trans Mobile Comput, 2017, 16: 2545–2559

248 Solomitckii D, Gapeyenko M, Semkin V, et al. Technologies for efficient amateur drone detection in 5G millimeter-wave cellular infrastructure. IEEE Commun Mag, 2018, 56: 43–50

249 Ezuma M, Ozdemir O, Anjinappa C K, et al. Micro-UAV detection with a low-grazing angle millimeter wave radar. In: Proceedings of IEEE Radio and Wireless Symposium (RWS), 2019. 1–4

250 Meer I A, Ozger M, Cavdar C. On the localization of unmanned aerial vehicles with cellular networks. 2020. ArXiv:2003.04889

251 Björnson E, Sanguinetti L, Wymeersch H, et al. Massive MIMO is a reality–What is next?: Five promising research directions for antenna arrays. Digital Signal Process, 2019, 94: 3–20

252 Bekmezci I, Sahingoz O K, Temel Ş. Flying ad-hoc networks (FANETs): a survey. Ad Hoc Networks, 2013, 11: 1254–1270

253 Wan J, Zhong L, Zhang F. Cooperative localization of multi-UAVs via dynamic nonparametric belief propagation under GPS

signal loss condition. Int J Distributed Sens Networks, 2014, 10: 562380

254 Lee W, Bang H, Leeghim H. Cooperative localization between small UAVs using a combination of heterogeneous sensors. Aerospace Sci Tech, 2013, 27: 105–111

255 Qu Y, Zhang Y. Cooperative localization against GPS signal loss in multiple UAVs flight. J Syst Eng Electron, 2011, 22: 103–112

256 Qu Y, Zhang Y, Zhou Q. Cooperative localization of UAV based on information synchronization. In: Proceedings of IEEE International Conference on Mechatronics and Automation, 2010. 225–230

257 Liu Y, Shen Y. UAV-aided high-accuracy relative localization of ground vehicles. In: Proceedings of IEEE International Conference on Communications (ICC), 2018. 1–6

258 Sorbelli F B, Das S K, Pinotti C M, et al. Range based algorithms for precise localization of terrestrial objects using a drone. Pervasive Mobile Comput, 2018, 48: 20–42

259 Lin X, Yajnanarayana V, Muruganathan S D, et al. The sky is not the limit: LTE for unmanned aerial vehicles. IEEE Commun Mag, 2018, 56: 204–210

260 Qualcomm. LTE Unmanned Aircraft Systems. 2017. Trial Report, https://www.qualcomm.com/documents/lteunmanned-aircraft-systems-trial-r\eport

261 Zhang Z, Xiao Y, Ma Z, et al. 6G wireless networks: vision, requirements, architecture, and key technologies. IEEE Veh Technol Mag, 2019, 14: 28–41

262 Yao H, Wang L, Wang X, et al. The space-terrestrial integrated network: an overview. IEEE Commun Mag, 2018, 56: 178–185