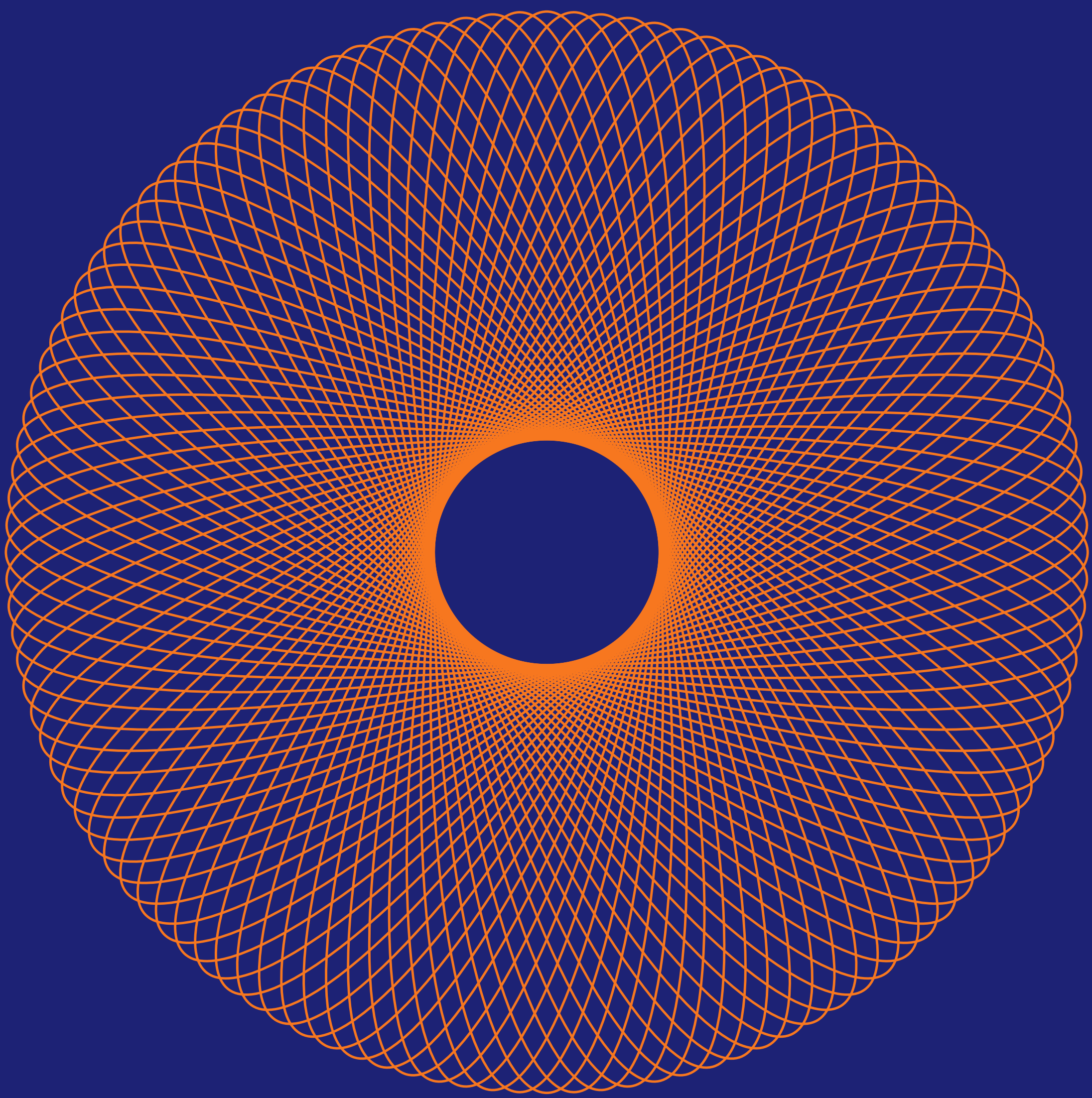
SAMSUNG

6G

The Next Hyper——Connected Experience for All.



Issued by Samsung Research

Table of Contents

Preface

Following the commercialization of 5G technologies, both academia and industry are initiating research activities to shape the next-generation communication system, namely 6G. Considering the general trend of successive generations of communication systems introducing new services with more stringent requirements, it is reasonable to expect 6G to satisfy unprecedented requirements and expectations that 5G cannot meet.

We expect that 6G will provide ultimate experience for all through hyper-connectivity involving humans and everything. In this white paper, we aim to provide readers with a comprehensive overview of various aspects related to 6G, including technical and societal trends, services, requirements, and candidate technologies. The rest of this white paper is organized as follows:

Section 1 introduces megatrends driving technology evolution towards 6G.

Section 2 discusses major services that have to be taken into account in developing 6G technologies.

Section 3 describes requirements to realize the expected services for 6G. They consist of performance requirements, architectural requirements, and trustworthiness requirements.

Section 4 introduces candidate technologies that will be essential to satisfy the requirements for 6G, which currently include support of the terahertz band, novel antenna technologies, evolution of duplex technology, evolution of network topology, spectrum sharing, comprehensive AI, split computing, and high-precision network.

Section 5 provides an initial expectation of the 6G timeline. We anticipate that the earliest commercialization could occur as early as 2028 while massive commercialization may emerge around 2030.

Section 6 provides concluding remarks.

1 Megatrends toward 6G

Applications that take advantage of wireless communications are expanding from connecting humans to connecting various things. Wireless communication is becoming an important part of social infrastructure and people's daily lives. In addition, today's exponential growth of advanced technologies such as artificial intelligence (AI), robotics, and automation will usher in unprecedented paradigm shifts in the wireless communication. These circumstances lead to four major megatrends advancing toward 6G: connected machines, use of AI for the wireless communication, openness of mobile communications, and increased contribution for achieving social goals. The rest of this section discusses details of these four megatrends.

Connected Machines

Machine as a Main User

It is envisaged that the number of connected devices will reach 500 billion by 2030 [1], which is about 59 times larger than the expected world population (8.5 billion [2]) by that time. Mobile devices will take various form-factors, such as augmented reality (AR) glasses, virtual reality (VR) headsets, and hologram devices. Increasingly, machines will need to be connected by means of wireless communications. Examples of connected machines include vehicles, robots, drones, home appliances, displays, smart sensors installed in various infrastructures, construction machineries, and factory equipment. Figure 1 illustrates this trend of mobile devices and connected machines.

Figure 1

Evolution of mobile devices and connected machines.

As the number of connected machines grows exponentially, those machines will become dominant users of 6G communications. Looking back at the history of wireless communications, technologies have been developed assuming services for humans as the major driving applications. In 5G, machines were also considered in defining requirements and developing technologies. We expect new 6G technologies have to be developed specifically to connect hundreds of billions of machines taking into account what is required for machines.

To provide an initial insight into the performance targets needed for connected machines, Table 1 compares the perception capability of humans and machines. For example, the capability of human eye is limited to a maximum resolution of 1/150° and view angle of 200° in azimuth and 130° in zenith. On the other hand, machine vision capability is not constrained by such limitations, since it can take advantage of many cameras with various functions. Considering such high capabilities of machines, the performance requirements for the 6G system could be extremely high for relevant service scenarios.

Table 1

Comparison of the perception capability of humans and machines.

	Human	Machine
Maximum Resolution	1/150° (Smartphone display 290 ppi at 30 cm)	
Latency Perception	<100 ms	
Audible Frequency	250-2,000 Hz	Exceeds Human Limitations!
Visible Wavelength	280-780 nm	
Viewing Angle	Azimuth 200°, Zenith 130°	

AI
New Tool for Wireless Communications

In recent years, the rise of AI has pervaded various areas such as finance, health care, manufacturing, industry, and wireless communication systems. Application of AI in wireless communications holds the potential to improve performance and reduce capital expenditure (CAPEX) and operational expenditure (OPEX). For example, AI can

- Improve performance of handover operation taking into account network deployments and geographical environments
- Optimize network planning involving base station (BS) location determination
- Reduce network energy consumption
- Predict, detect, and enable self-healing of network anomalies

The potential benefits of AI applied to wireless communications are promising. On the other hand, there is a limit to what is achievable today, as use of AI in communication networks was not considered when developing existing communication systems such as 5G. In the case of 6G, knowing that AI technologies are available for practical applications, we can develop a system that takes into account the possibility of embedding AI in various entities comprising wireless networks and services. A tremendous amount of data associated with hundreds of billions of connected machines and humans needs to be collected and utilized in 6G systems.

Considering AI from the initial phase of developing concepts and technologies for 6G will give us more opportunities to take advantage of AI for improvement of overall network operation in terms of performance, cost, and ability to provide various services.

Openness of Mobile Communications

Substantial improvements of the computing power of general purpose processors such as central processing units (CPUs) and graphic processing units (GPUs) enabled software-based implementations of network entities including core networks and BSs. This trend also makes open source software an attractive option to realize network functions, as it can lower barriers to market entry, promote interoperability, and expedite development cycles based on shared knowledge. An example of the related industry activities is the open radio access network (O-RAN) alliance that aims to provide an open and intelligent radio access network (RAN). Another example is the open network automation platform (ONAP), which develops a platform for network management and its automation through an open-sourced shared architecture.

Another noteworthy trend is the utilization of personal yet possibly anonymized user information to improve personalized quality-of-service (QoS) and quality-of-experience (QoE) of the services provided by mobile network operators (MNOs). Use of AI in various services increases the need for utilizing user information.

Social Goals and Mobile Communications

With the growing importance of mobile communications as social infrastructure, governments and international organizations expect 5G to play a pivotal role in ameliorating many social issues such as climate change, hunger, and education inequality [3]. For example, the combination of 5G and digitization is expected to help reduce greenhouse gas emission by up to 15% by 2030 [4]. 5G networks can enable remote learning, which could provide a means to improve education equality. Despite help from 5G and other measures, regional and social disparities continue to widen. For fundamental resolution of social issues, the UN adopted the Sustainable Development Goals (SDGs) in the Agenda 2030 [5].

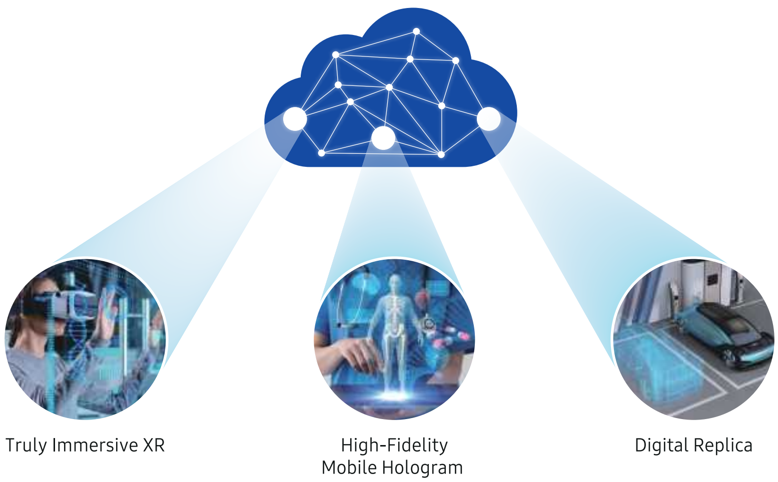
We anticipate a societal need for 6G mobile communications, to contribute even more to address social issues and achieve the SDGs. Hyper-connectivity and ultimate experience delivered by 6G mobile communications will improve and enable access to required information, resources (both virtual and physical), and social services without constraints of time and physical location. A wide deployment of 6G will reduce differences in regional and social infrastructure and economic opportunities and thereby provide alternatives to rural exodus, mass urbanization and its attendant problems. We expect, in summary, that 6G mobile communications will play an important role in achievement of UN's SDGs, and tremendously contribute to the quality and opportunities of human life.

2 6G Services

Representative categories of 5G services, i.e., enhanced mobile broadband (eMBB), ultra-reliable and low latency communications (URLLC), and massive machine-type communications (mMTC) will continue to improve moving towards 6G. In this section, we focus on new 6G services that will emerge due to advances in communications as well as other technologies such as sensing, imaging, displaying, and AI. Those new services will be introduced through hyper-connectivity involving humans and everything and provide ultimate multimedia experience. In the rest of this section, we highlight three key 6G services, namely, truly immersive extended reality (XR), high-fidelity mobile hologram, and digital replica as illustrated in Figure 2.

Figure 2

Three key 6G services: truly immersive XR, high-fidelity mobile hologram, and digital replica.



Truly Immersive XR

XR is a new term that combines VR, AR, and mixed reality (MR). It has attracted great attention and opened new horizons in various fields including entertainment, medicine, science, education, and manufacturing industries. Technical development to realize XR is still in progress, and new innovative technologies are constantly appearing. The critical obstacle between the potential and reality of XR is hardware. In particular, these technologies require advanced device form-factors, such as hand-held components, to support mobile and active software content. Current mobile devices lack sufficient stand-alone computing capability. Unfortunately, progress in hardware performance, especially mobile computing power and battery capacity, cannot keep pace with what the boom of XR requires. This discrepancy could severely deter market expansion. In our view, these challenges can be overcome by offloading computing to more powerful devices or servers.

Another challenge is sufficient wireless capacity. Note that current AR technology requires 55.3 megabits per second (Mbps) to support 8K display (with one million points), which can provide enough user experience on a mobile display. However, in order to provide truly immersive AR, the density should be largely improved and it will require 0.44 gigabits per second (Gbps) throughput (with 16 million points). In addition, XR media streaming may have similar demands to 16K UHD (Ultra High Definition) quality video. For example, 16K VR requires 0.9 Gbps throughput (with compression ratio of 1/400). The current user experienced data rate of 5G is not sufficient for seamless streaming. It is expected that the market sizes for VR and AR will reach $44.7 billion [6] and $87 billion [7], respectively, by 2030.

Figure 3

Truly immersive XR.

High-Fidelity Mobile Hologram

With the unprecedented rate of advances in high-resolution rendering, wearable displays, and wireless networks, mobile devices will be able to render media for 3D hologram displays. Hologram is a next-generation media technology that can present gestures and facial expressions by means of a holographic display. The content to display can be obtained by means of real-time capture, transmission, and 3D rendering techniques. In order to provide hologram display as a part of real-time services, extremely high data rate transmission, hundreds of times greater than current 5G system, will be essential. For example, 19.1 Gigapixel requires 1 terabits per second (Tbps) [8]. A hologram display over a mobile device (one micro meter pixel size on a 6.7 inch display, i.e., 11.1 Gigapixel) form-factor requires at least 0.58 Tbps. Moreover, support of a human-sized hologram requires a significantly large number of pixels (e.g., requiring several Tbps) [9]. The peak data rate of 5G is 20 Gbps. 5G cannot possibly support such an extremely large volume of data as required for hologram media in real-time. To reduce the magnitude of data communication required for hologram displays and realize it in the 6G era, AI can be leveraged to achieve efficient compression, extraction, and rendering of the hologram data. The market size for the hologram displays is expected to be $7.6 billion by year 2023 [10].

Figure 4

3D hologram display over mobile devices.

Digital Replica

With the help of advanced sensors, AI, and communication technologies, it will be possible to replicate physical entities, including people, devices, objects, systems, and even places, in a virtual world. This digital replica of a physical entity is called a digital twin. In a 6G environment, through digital twins, users will be able to explore and monitor the reality in a virtual world, without temporal or spatial constraints. Users will be able to observe changes or detect problems remotely through the representation offered by digital twins.

Users will be even able to go beyond observation, and actually interact with the digital twins, using VR devices or holographic displays. A digital twin could be a representation of a remotely controlled set of sensors and actuators. In this manner, a user's interaction with a digital twin can result in actions in the physical world. For example, a user could physically move within a remote site by controlling a robot in that space entirely via real-time interactions with a digital twin representation of that remote site.

With the help of AI, digital replication, management of real world and problem detection and mitigation can be done efficiently without the presence or even detailed supervision by a human being. For instance, if a problem is detected in the digital twin representation, AI can invoke required actions in the real world.

The technical challenges are significant. In order to, for example, duplicate 1 m x 1 m area, we need a Tera-pixel, which requires 0.8 Tbps throughput assuming periodic synchronization of 100 ms and a compression ratio of 1/300. The expected market size for digital replica is estimated to be $26 billion in 2025 [11].

Figure 5

Digital replica: bridge the real and virtual worlds.

3 Requirements

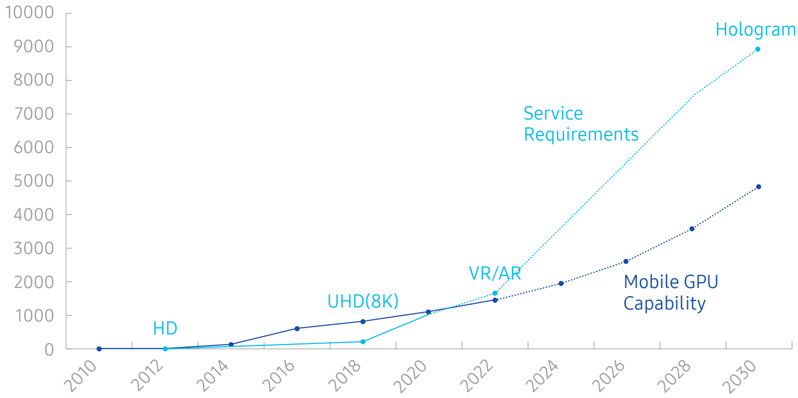
There will be new advanced services in 6G era, which require a tremendous amount of real-time data processing, a hyper-fast data rate, and extremely low latency, as discussed in Section 2. These new services can be characterized as providing ultimate experience through hyper-connectivity.

Relying only on improvement of the communication link performance does not suffice to realize new 6G services. This is because the progress of increase in mobile computing power is not keeping pace with the growth rate of the computing power requirements as shown in Figure 6. Also, the improvement of mobile devices' battery life is not fast enough given the extensive demand for multimedia services and their rapid evolution.

Figure 6

Computing power: requirement versus capability.

Computing Capability Trend (Gigaflops/sec)



The limitation on computing power and battery life of mobile devices highlights the need for taking a holistic view on the overall architecture of 6G system including network entities as well as mobile devices. In addition, the openness of mobile communications as a new megatrend toward 6G (see Section 1) gives rise to the need for defining new requirements to maintain security and keep the trust of subscribers.

In the rest of this section, we describe our view on 6G requirements for the key performance indices, overall architecture, and trustworthiness.

Performance Requirements

In order to realize advanced multimedia services such as truly immersive XR, mobile hologram, and digital replica, 6G needs to provide a much higher data rate than 5G. While 5G was designed to achieve 20 Gbps peak data rate, in 6G, we aim to provide the peak data rate of 1,000 Gbps and a user experienced data rate of 1 Gbps. To provide advanced multimedia services to a large number of people, the overall network performance needs to be improved, e.g., we can aim to have 2 times higher spectral efficiency than 5G.

To provide the ultimate experience of delay-sensitive real-time applications such as interactive tactile internet, latency-related performance needs to significantly improve. Performance targets include air latency less than 100 µs, end-to-end (E2E) latency less than 1 ms, and extremely low delay jitter in the order of microseconds. With these requirements satisfied, the user experienced latency can be less than 10 ms, which is the motion-to-photon latency requirement for XR services [12]. The user experienced latency requirement applies to the aggregate of all latency components in wireless links, wireline links, and the computation on both client and server sides.

For support of latency-sensitive services requiring extreme reliability, e.g., industrial automation, emergency response, and remote surgery, we intend to improve reliability by 100 times compared to 5G so that the error rate is 10^{-7}.

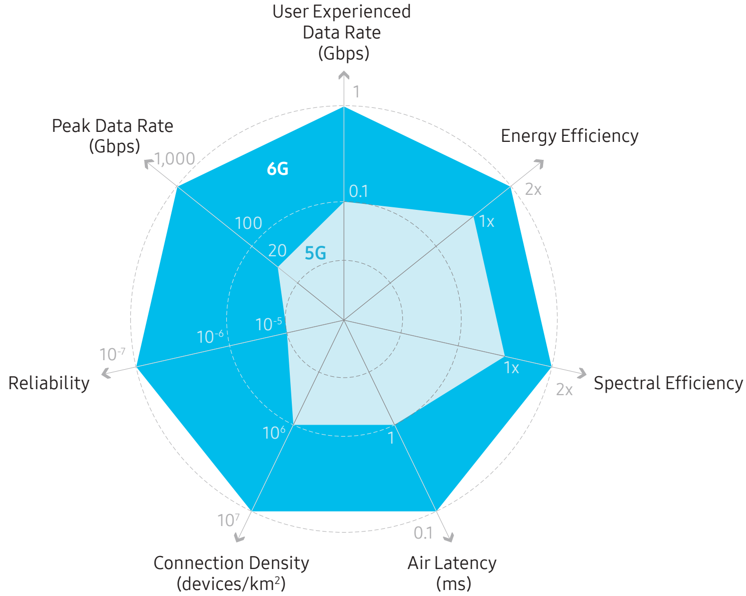
Network coverage has always been important over past generations and will remain very important in 6G. We aim to support larger coverage than 5G. The maximum supported speed of the mobile device improved from 350 km/h in 4G to 500 km/h in 5G. It may need to further improve in 6G depending on the evolution of transportation systems. The explosive growth in the number of connected machines will require 6G to support about 10^7 devices per square kilometer. This is ten times larger than the connection density requirement of 5G.

In the 6G era, users will expect seamless high-end services in their everyday lives, ideally with improved battery life. Considering the growing concern about environmental sustainability, the energy consumption of 6G networks should be minimized. We intend to improve the energy efficiency of both devices and networks by at least two times.

Figure 7 illustrates the enhancement of key requirements from 5G to 6G.

Figure 7

Comparison of key performance requirements between 5G and 6G.



Architectural Requirements

The architecture of 6G communication network should be developed so that it can resolve the issues arising from the limited computation capability of mobile devices. A possible way to achieve this is to offload computation tasks to more powerful devices or servers. In order to support offloading of real-time intensive computation tasks, hyper-fast data rate and extremely low latency communications are required. Understanding that there must be practical limits on the achievable data rate and latency, the communication network should be designed in a holistic manner, so as to best utilize computation power that can be made available by various entities in the network. We term this joint design "communications and computing convergence."

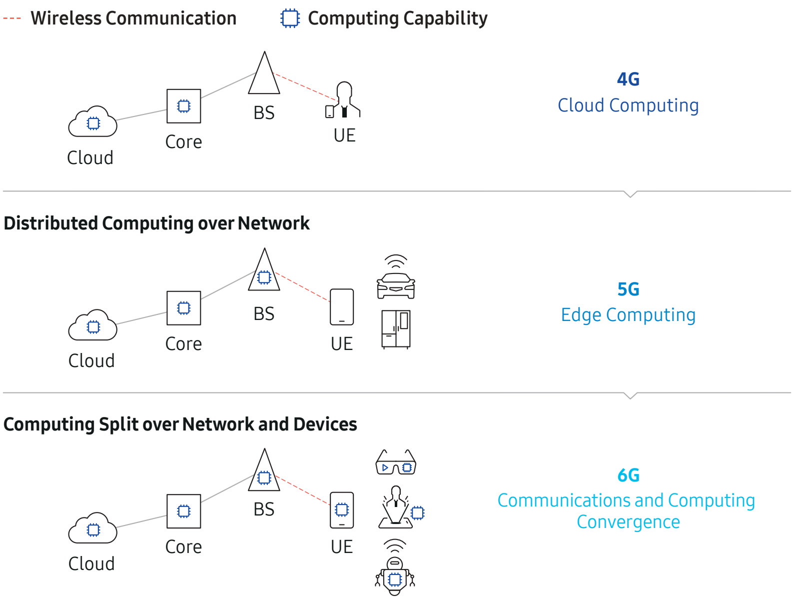
As illustrated in Figure 8, computation for IT services was performed in the cloud for 4G communications. One of the important advances of 5G is low latency. With the commercial deployment of 5G, there is increasing interest to move the computing capability closer to mobile devices, e.g., to position a multi-access edge computing (MEC) server in the 5G core or in the BS, so that the overall latency experienced by end user equipment (UE) can be further reduced. 6G should be designed to achieve true convergence of communications and computing so that an end user's various devices can seamlessly utilize the computing power available in the network.

On top of the communications and computing convergence, AI will need to be embedded in all system components in 6G networks. We call this design approach "native AI," which allows all system components to obtain and evaluate a massive amount of real-time information. In this way, system can handle complex optimization tasks across layers to optimize the system parameters and overall system performance.

The 6G network should be designed to support terrestrial components, e.g., fixed BSs or moving BSs, as well as non-terrestrial components, e.g., airplanes, urban air mobility (UAM) systems, low earth orbit (LEO) and geostationary orbit (GEO) satellites, and high altitude platform stations (HAPS).

Figure 8

Architectural evolution for convergence of communications and computing.



Trustworthiness Requirements

As discussed as a megatrend prefiguring 6G, the use of open source software and personal user information will increase the openness of communication systems and hence increase the attack surface. This could make the whole system more vulnerable to security and privacy threats as described in the following examples. First, there may not be enough verification of open source software codes against possible security attacks. Second, the service provider's access to user information will expose attack surfaces for leaking confidential user information and poses a severe threat to user privacy. In addition, user devices can be hacked unless these devices provide a sufficiently secure trusted environment. Compromised user devices decrease the security of the whole telecommunication system and the services accessed by users.

Considering the increasing risk of security threats, we expect that trustworthiness will become an essential requirement. Although it is currently difficult to define this requirement concretely, at least the following has to be considered.

- Hardware-based secure environment that provides secure operation of software code and protection of credential
- Secure-by-design approach to guarantee that any hardware/software can be trusted
- Transparency to ensure that the system identifies how and when the AI system accesses any code, training data, etc. related to personal information as well as how securely the AI system operates against adversarial machine learning
- Mechanisms to securely utilize an unprecedented amount of information concerning business and human users and to strictly maintain the privacy of such information

4 Candidate Technologies

The services and requirements for 6G presented in the previous sections pose various challenges on the development of future wireless system. In this section, we introduce candidate technologies that could be key enablers to realize 6G. While we find these technologies quite promising and relevant, we will consider additional technologies in the future and our view regarding the importance and usefulness of different technologies will naturally evolve as we proceed with our research.

Terahertz Technologies

It is inspiring that in March 2019, the Federal Communications Commission (FCC) opened the spectrum between 95 GHz and 3,000 GHz for experimental use and unlicensed applications to encourage the development of new wireless communication technologies [13]. Moreover, discussions on use cases and deployment scenarios for 5G new radio (NR) systems operating at bands beyond 52.6 GHz have begun [14]. Following this trend, it is inevitable that mobile communications will utilize the terahertz (THz) bands (i.e., 0.1-10 THz [15]) in future wireless systems. The THz band includes enormous amount of available bandwidth, which will enable extremely wideband channels with tens of GHz-wide bandwidth. This could potentially provide a means to meet the 6G requirement of Tbps data rate. Considering the advance of related technologies, we expect that 6G would need to be designed to utilize up to 3,000 GHz as shown in Figure 9.

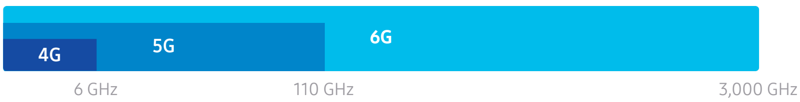


Figure 9

Spectrum usage for different generations.

While the availability of wideband spectrum is the main driver for THz communications, other benefits can also be realized. For example, the communication in THz band can potentially provide high-precision positioning capability for the following reasons: 1) Extremely wideband waveforms in the THz band would enable accurate ranging between transmitter and receiver (possibly with sub-centimeter-scale accuracy) [16][17]. 2) Links between transmitter and receiver will most likely be line of sight (LoS), as discussed further in detail later in this paper. This will greatly improve the accuracy of distance-based positioning systems. 3) The use of pencil-point sharp beams steered in both azimuth and elevation will greatly improve angular resolution and triangulation accuracy of 3D position estimation.

However, to realize stable THz communications in practice, a handful of fundamental and technical challenges need to be overcome. In the following, we highlight some of these challenges, mostly from the physical layer's perspective.

THz Challenges

- *Severe path-loss and atmospheric absorption:* Free-space path-loss is proportional to the square of the signal frequency. For example, a link at 280 GHz has 20 dB additional path-loss compared to 28 GHz. Nevertheless, the severe path-loss in THz band can be overcome, for example, by utilizing very large antenna arrays at BSs, namely *ultra-massive multiple-input multiple-output (MIMO).* In addition, the effect of atmospheric absorption (i.e., absorption by molecules in air) in the THz band is in general severer than in lower frequencies as the absorption lines for oxygen and water are mostly located in the THz band [18]. In order to design efficient THz communication systems in practice, accurate yet tractable THz multipath channel models need to be developed for both indoor and outdoor environments.
- *RF front-end, photonics and data conversion:* The THz band is often referred to as the THz gap due mainly to the lack of existing efficient devices, which can generate and detect signals in these frequencies. In these bands, the device dimensions are significantly large relative to the wavelength, and it results in high power loss or equivalently low efficiency. On the positive side, during the last decade, researchers put great efforts for developing chip-scale THz technologies. As a result, nowadays semiconductor technologies based on InP, GaAs, SiGe, and even CMOS are capable of generating power in the mW range with acceptable efficiency [19][20][21] at low THz band. However, further development of solid-state electronics is required for operation in high THz band.

  While the research on generating and improving the output power of the THz signal is important, researchers need to address many other challenges: 1) transporting the signal within the integrated system and to the antenna with low loss; 2) packaging of the integrated system without significant loss, and maintaining proper heat dissipation; 3) lowering the mixer phase-noise; 4) low power multi-Giga-samples-per-second analog-to-digital converters (ADCs) and digital-to-analog convertors (DACs); and lastly 5) low power digital input/output (IO) to DACs and ADCs to transfer data at Tbps data rate with acceptable power consumption.

- *Antenna, lens, and beamforming architecture:* Moving up to THz-range frequency means drastic increase in path-loss. Consequently, unprecedentedly massive antenna arrays are necessary to compensate for the path-loss. Designing such arrays that will operate with high efficiency at THz frequency poses many challenges to designing the feed network and the antenna elements to support GHz-wide bandwidth. Moreover, the use of ultra-massive arrays results in very focused beams, similar to laser beams. As a result, communication links at these frequencies will depend on LoS and focused-reflected paths, not on scattering and diffracting paths. It is of great importance to optimize the beamforming architecture to provide high dynamic-range and high flexibility at a reasonable cost and energy consumption. Novel antenna technologies are discussed in the next section.
- *New waveforms, signals, channels, and protocols:* Efficient operation of a wireless system is highly dependent on the design of a proper waveform. While OFDM remains a strong candidate waveform for 6G THz systems, it is necessary to explore alternative waveforms to support GHz-wide channels, reduce PAPR, and withstand the THz hardware limitations. Moreover, it is necessary to develop proper design of signals, channels, and protocols, which are effective yet of low complexity for THz operation.

Novel Antenna Technologies

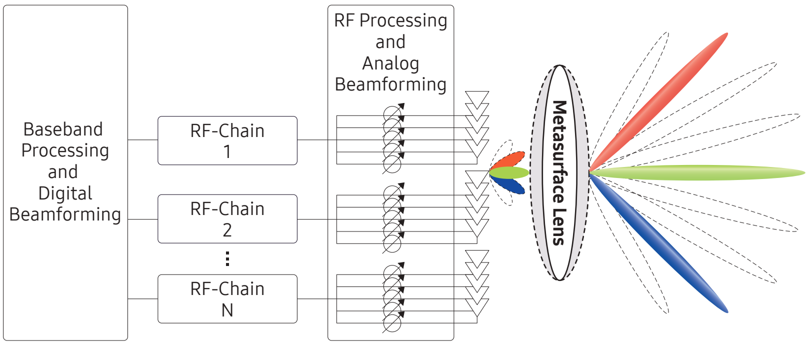
To cope with the difficult propagation characteristics of THz band, it may be natural to enhance the massive MIMO technology that was introduced to support millimeter wave (mmWave) band in 5G. Since the THz band requires much more antennas than the mmWave band, there may be significantly more practical difficulties. In this section, we briefly review novel antenna technologies as possible alternatives.

Metamaterial based Antenna and RF Front-End

A metamaterial is usually constructed by arranging multiple tunable elements (PIN diodes, varactor diodes, etc.) in repeating patterns, at scales that are smaller than the wavelengths [22]. Its precise shape, geometry, size, orientation, and arrangement enable smart properties capable of manipulating electromagnetic waves, e.g., blocking, absorbing, enhancing, or bending waves, to achieve benefits that go beyond what is possible with conventional materials. In addition, each element constituting a metamaterial can be controlled independently to achieve desirable characteristics of the electromagnetic waves such as the direction of propagation and reflection. There are three outstanding approaches for utilizing metamaterial as follows.

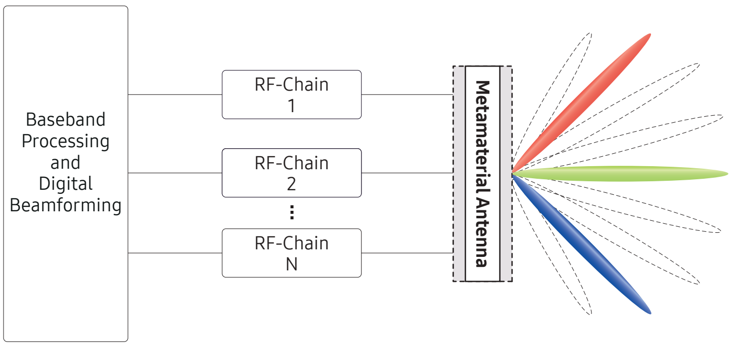
- *Metasurface lens* as a phase shifting structure is applied to the signal radiated from an antenna array [23]. It can adjust a beam direction by applying DC bias to its constituting elements as shown in Figure 10. The metasurface lens has potential to help sharpen a beam shape.

Figure 10
Metasurface lens.



- *Metamaterial antenna* acts as a resonant antenna to radiate directive beams by itself as shown in Figure 11 [24]. In contrast to the metasurface lens, it does not require a separate antenna array with phase shifters.

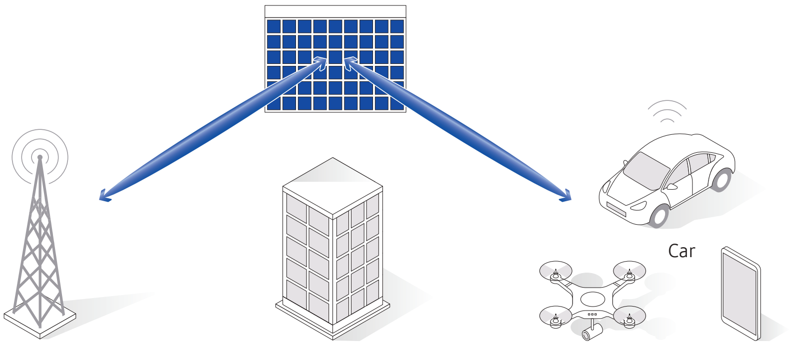
Figure 11
Metamaterial antenna.



- *Reconfigurable intelligent surface (RIS)* can be used to provide a propagation path where no LoS link exists [25]. An example of signal reflection via RIS is illustrated in Figure 12.

Figure 12

RIS-aided communication between a BS and a mobile user, where the LoS path is blocked.



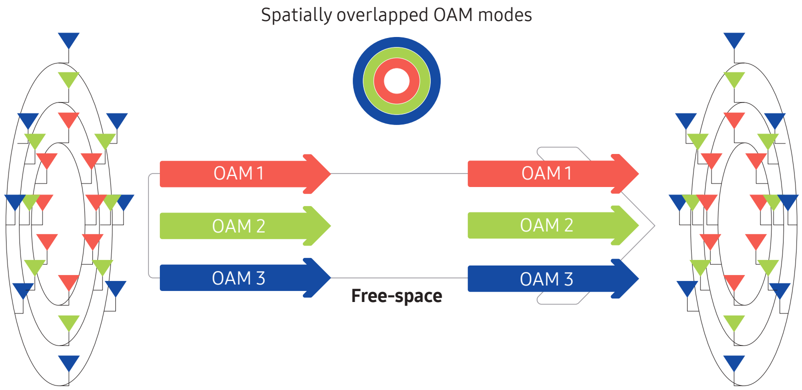
Orbital Angular Momentum

When light travels through space, electric and magnetic fields composing the light have their own oscillating axes perpendicular to each other. There are two types of rotations related to these axes and these rotations can be observed as two types of momentum, namely, spin angular momentum (SAM) and orbital angular momentum (OAM). OAM theoretically could have multiple orthogonal modes depending on how fast the light can rotate around its propagation direction.

In case of the electromagnetic waves, it has been proven by experiments that different OAM modes can be generated simultaneously by using a transmit antenna array [26]. This property can be used to multiplex multiple signals (or layers) by using different OAM modes to increase data rate. A study has shown that OAM is a special case of traditional spatial multiplexing in terms of capacity and overall antenna occupation [27]. Figure 13 shows a specific design of the OAM multiplexing proposed in [28] for wireless backhaul channels, where the LoS path is dominant. Through OAM, high-order spatial multiplexing can be achieved in the environment where it is impossible with conventional MIMO technologies, i.e., LoS channel. In that context, this technology is quite promising, yet there remain many issues for the real implementation and operation in practical environments.

Figure 13

OAM multiplexing with different OAM modes.

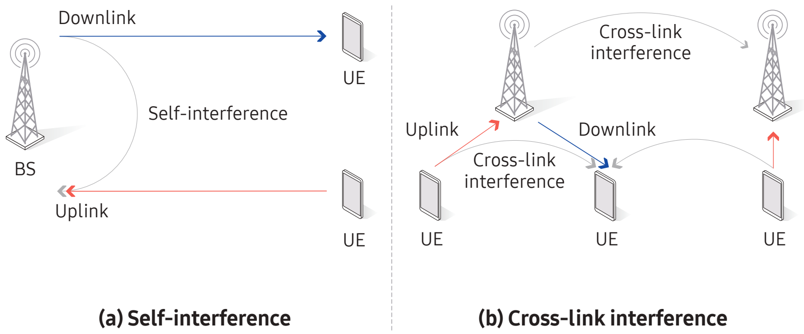


Evolution of Duplex Technology

In conventional communication systems, downlink and uplink transmissions occur in a mutually exclusive manner either in time domain (i.e., TDD) or in frequency domain (i.e., FDD). Typically, the downlink and uplink receive fixed allocations of time-frequency resources in practical systems. In 5G NR, dynamic TDD was introduced to improve the duplex flexibility, thus making it possible to adjust the ratio between downlink and uplink time slots depending on traffic demand. While this is an improvement over earlier generations, there is still active research [29][30][31] into how to remove the restriction that downlink and uplink must use mutually exclusive time-frequency resources. We refer to this restriction as the "mutually exclusive" principle hereafter.

Figure 14

Main obstacles in deviating from the "mutually exclusive" principle.



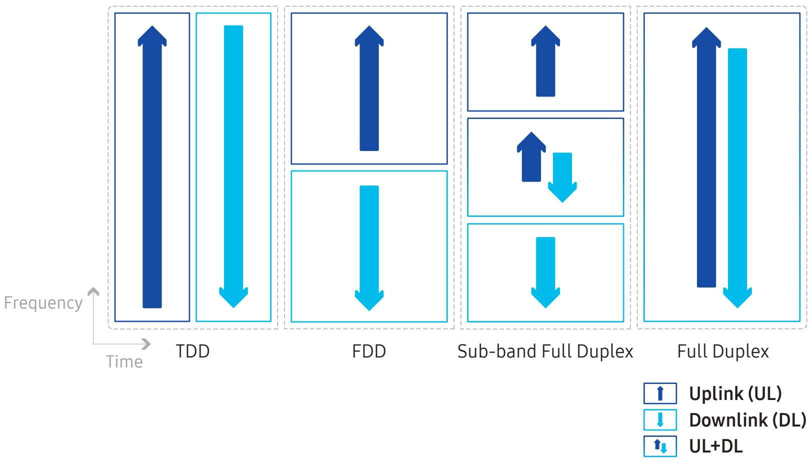
Allowing overlap between downlink and uplink over the entire time-frequency resource (a.k.a. "full duplex") can increase system capacity by two times, in theory. The main obstacles encountered upon deviating from the "mutually exclusive" principle include self-interference and cross-link interference. Self-interference experienced by a BS receiver is illustrated in Figure 14(a). The BS transmits downlink signal using the same time-frequency resource as used for the uplink signal from UEs. Since the BS's transmit and receive antennas are located in close proximity, self-interference is much stronger than the desired signals from the UEs. Therefore, to evolve duplex technology by departing from the "mutually exclusive" principle, it is crucial to be able to remove self-interference. There has been relevant research on self-interference cancellation (SIC) techniques, which typically require both analog and digital domain cancellation [32][33].

Cross-link interference (CLI) is the interference between UEs or between BSs. Figure 14(b) illustrates an example, where UEs maintain the "mutually exclusive" principle between its transmission (i.e., uplink) and reception (i.e., downlink) while BS does not. CLI between UEs is caused if the same time-frequency resource is allocated for the uplink transmission from a UE and the downlink transmission to another UE. The UE-to-UE CLI can be mitigated if a BS chooses a set of UEs that do not cause severe interference to each other. The CLI between BSs occurs when the aggressor BS's downlink uses the same time-frequency resource as the victim BS's uplink. The BS-to-BS CLI can be mitigated via close coordination between BSs.

In today's practical cellular systems, a band has a fixed duplex scheme, i.e., either FDD or TDD. If deviating from the "mutually exclusive" principle becomes a reality, it would be possible to adapt the duplex scheme in a dynamic manner as shown in Figure 15. This would improve the spectral efficiency as well as the system operation flexibility.

Figure 15

Dynamic operation of duplex modes.



Evolution of Network Topology

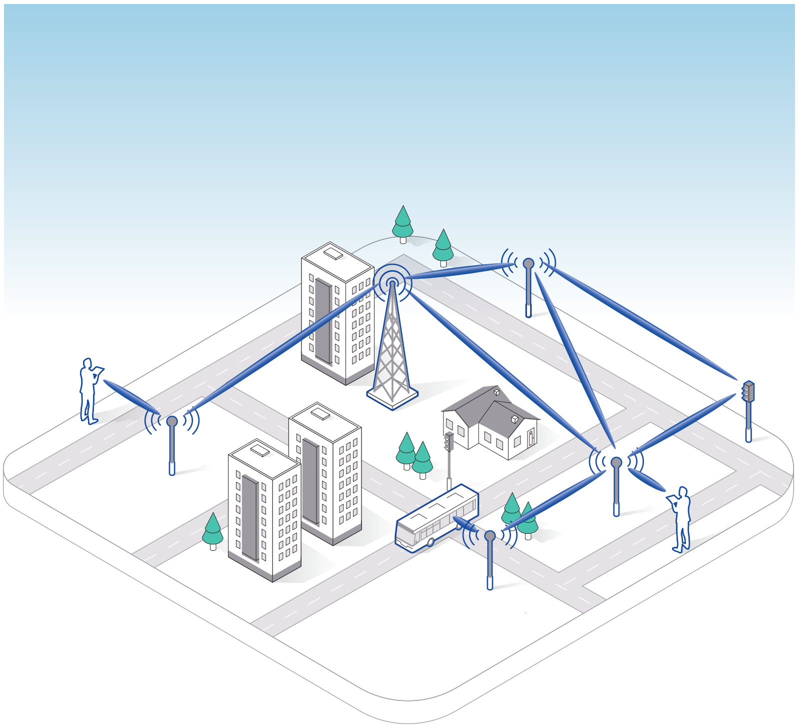
Cellular BSs have typically been deployed with fixed locations and connected by fixed networks. With such static network topology and wireline backhaul and fronthaul, it is difficult and costly to establish additional BSs to accommodate an increase of data traffic or to fill holes in coverage. To enable flexible network deployments, the mobile industry introduced support for network entities to connect to BSs via wireless connections, such as relay in 4G and integrated access and backhaul (IAB) in 5G. In addition to mobility support for individual mobile devices in the cellular network, there has been interest in the concept of group mobility (also known as a mobile relay or mobile BS) for efficient support of mobile devices that are moving as a group on a bus, a train, or even an airplane.

Moving toward 6G, we expect that the technologies related to the above trend will further advance to achieve the following.

- Automated addition, configuration, and optimization of new network entities connected to existing BSs via wireless connections. This will significantly reduce the effort for network planning, and hence, the mesh type network topology as illustrated in Figure 16 can become a major technology for flexible and adaptive network deployment.
- Enhanced mobility support for mobile network entities taking into account the speed of transportation systems that can be a part of the cellular network.
- Enhanced service continuity for user devices served by the network entities, which themselves may be moving and are connected to the cellular network through wireless connections.

Figure 16

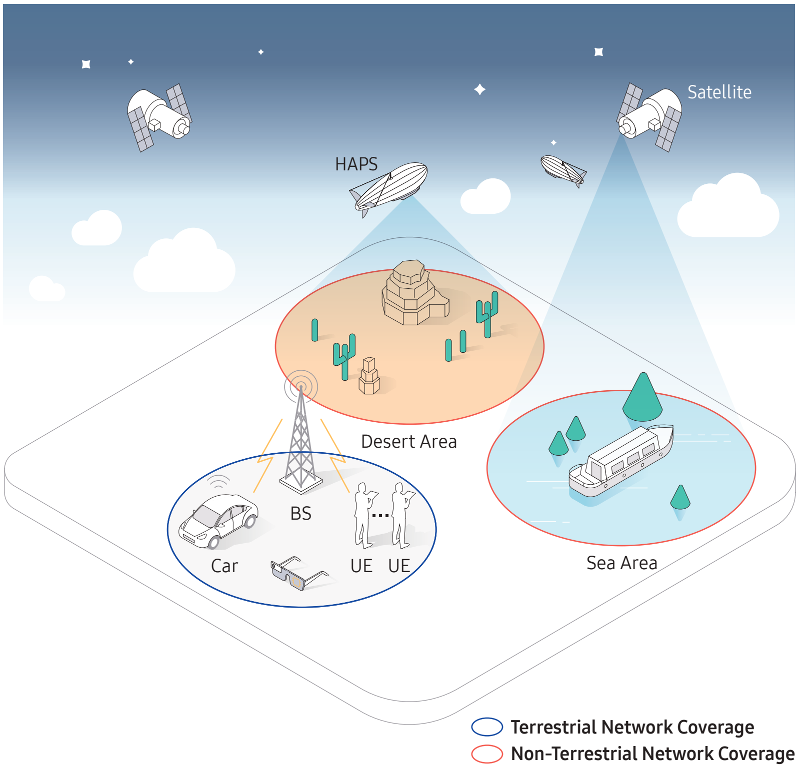
Mesh type network topology.



Another trend that continues to make progress in network topology evolution is the use of non-terrestrial network (NTN) components, e.g., satellite and HAPS, to provide coverage even in locations where there is no terrestrial network, as illustrated in Figure 17. Realization of NTN technology necessitates consideration of new aspects absent from terrestrial networks, including support of moving cells, large cell sizes as great as hundreds of kilometers, long propagation delays, large Doppler shift due to the high speed of NTN components, and large path-loss. Additional aspects, as yet unrecognized, may arise and need to be considered, since the mobile industry is at the initial stage for developing technologies to support NTN. As NTN components become widely deployed, investigation of technologies will proceed, to improve the overall performance of communications involving NTN components and to provide tight integration of NTN components in overall operation of mobile communication systems.

Figure 17

Inclusion of non-terrestrial components in mobile communications.



Spectrum Sharing

Spectrum sharing technology enables the use of spectrum by multiple entities. Exclusive licensees often underutilize licensed spectrum because they do not actively utilize it all the time. Allowing opportunistic use of the underutilized spectrum by others can make the best use of the limited and precious spectrum resources, especially those at low frequency ranges, e.g., below 6 GHz. These resources are critically important for guaranteeing the seamless coverage of mobile communications, but are scarce. We also observe that regulatory bodies begin to consider deviating from the traditional exclusive spectrum licensing approach to achieve better utilization of the limited spectrum. Considering such trends, spectrum sharing technology is worth paying attention to.

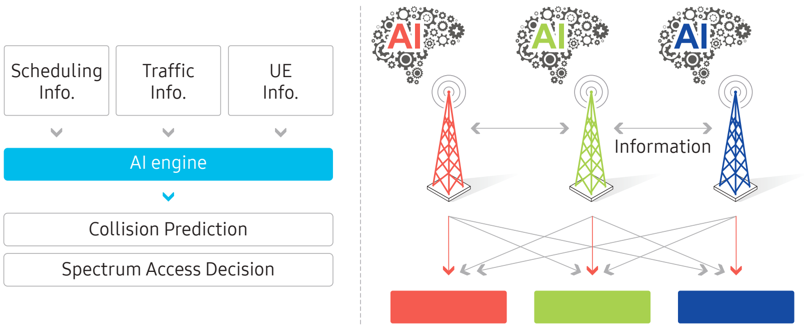
In U.S., the Citizens Broadband Radio Service (CBRS) band (3.55-3.7 GHz) has been opened for shared access by FCC. The sharing occurs according to a unique three-tiered access model: incumbents, i.e., federal government and fixed satellite service users, priority access licensees (PALs) and general authorized access (GAA) users, in descending priority order [34]. The Wireless Innovation Forum (WinnForum) [35] has defined the functionality and architecture for Spectrum Access Systems (SAS). The SAS framework provides access to a database and the ability to determine the availability of CBRS channels at a given location. In addition, Environmental Sensing Capability (ESC) is a functionality used to detect whether incumbent users occupy CBRS channels. Together, these two mechanisms maintain and enforce the hierarchical use of the spectrum. WinnForum also defines the interface between SAS and CBRS Devices (CBSDs), i.e., BSs, as well as the framework for testing and certification. At the same time, the CBRS Alliance has been developing the so-called Coexistence Manager (CxM) between SAS and GAA CBSDs to enable the sharing of the spectrum between GAA CBSDs in a semi-static manner for the allowed spectrum indicated by SAS [36]. In addition to the CBRS band, authorities consider making the 37-37.6 GHz band available for coordinated co-primary shared access between federal and non-federal users [37].

Traffic and spectrum demand are mostly concentrated in urban areas. In addition, traffic demand tends to follow people's activity cycle, e.g., high traffic during the daytime and so on. Given that traffic demand is highly correlated among the networks of different operators both temporally and geographically, there remains limited opportunity for sharing spectrum when and where the spectrum is the most needed. On the other hand, different operators' networks may exhibit quite non-uniform traffic patterns in short-term scales. To increase the spectrum sharing opportunity in response to this phenomenon, one can consider relying on more dynamic spectrum sharing rather than relying on a semi-static method based on database.

The main challenge of the dynamic spectrum sharing is avoiding (or minimizing) collision of spectrum usage among different entities while allowing them to access spectrum in a dynamic manner. Theoretically, to prevent such collisions, network operators could exchange all relevant spectrum access information. In practice, however, this would not be possible because acquiring all required information for every entity in real time would impose an enormous communication overhead. AI could avoid collisions by predicting the spectrum usage of other entities with a limited amount of information exchanged, as illustrated in Figure 18.

Figure 18

Intelligent spectrum sharing.



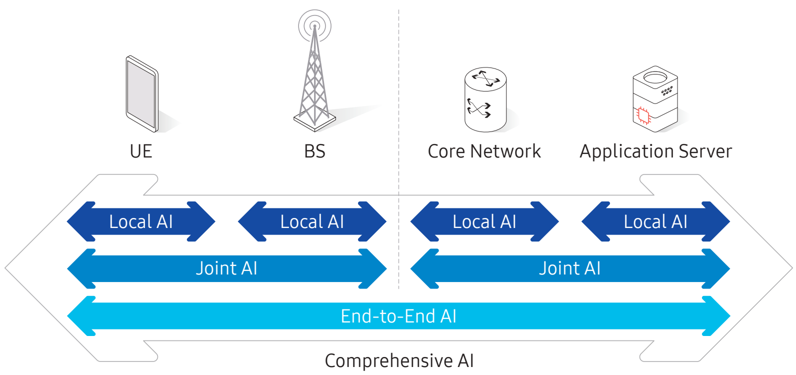
Comprehensive AI

AI receives much attention as a tool to solve problems that were previously deemed intractable due to their tremendous complexity or the lack of the necessary model and algorithm. In this section, we discuss a comprehensive AI system to optimize the overall system performance and network operation.

In general, an overall network architecture consists of four tiers of entities: UE, BS, core network, and application server. Application of AI can be categorized into three levels as shown in Figure 19: 1) local AI, 2) joint AI, and 3) E2E AI.

Figure 19

Comprehensive AI system model.



Local AI is implemented in each entity. An example is the use of AI for optimization of modulation, source coding, and channel coding [38][39]. Joint AI can optimize the joint operation of UEs and BSs or the joint operation of core networks and application servers. An example opportunity for joint optimization is handover optimization based on prediction of future network conditions in complex wireless environments [40]. E2E AI optimizes the entire communication system. With the E2E AI, it becomes possible to identify or predict anomalies in network operation and suggest corrective actions [41].

To obtain benefit from application of AI on various components of cellular systems, there are ongoing efforts to introduce support for AI in standards. The third generation partnership project (3GPP) has standardized network data analytics function (NWDAF) for data collection and analytics in automated cellular networks [42]. In addition to 3GPP, leading mobile network operators have established the O-RAN Alliance in 2018, with the intention to usher in an open and efficient RAN leveraging AI technologies [43]. Leading network vendors have joined the alliance. We agree that this is the right time to work towards use of AI in wireless communications. This effort, as we progress towards 6G, will result in native support of a comprehensive AI system to realize more efficient, more reliable, and low cost communication systems.

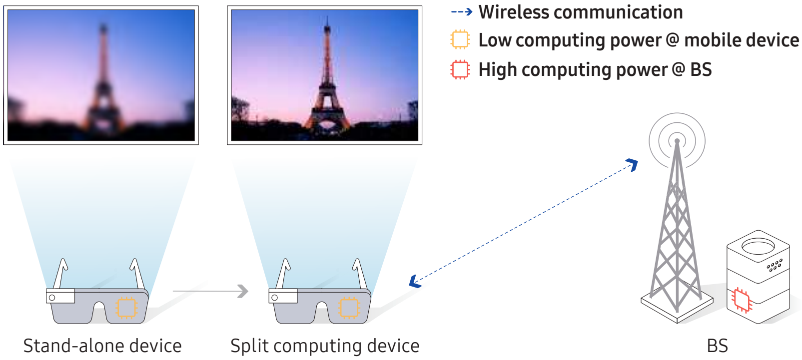
Split Computing

Future applications, such as truly immersive XR, mobile holograms, and digital replica, require extensive computation capabilities to deliver real-time immersive user experience. However, it would be challenging to meet such computational requirements solely with mobile devices, especially, given that many of future mobile devices will tend to become thinner and lighter. For example, AR glasses should be as light, thin, and small as regular glasses to meet the user's expectations.

In order to overcome the limits of the computing power of mobile devices, we consider the concept of split computing that makes use of reachable computing resources over the network. These computing resources could be available on various entities of networks, e.g., mobile devices, BSs, MEC servers and cloud servers. With split computing, mobile devices can effectively achieve higher performance even as they extend their battery life, as devices offload heavy computation tasks to computation resources available in the network. An example of split computing is illustrated in Figure 20.

Figure 20

Split computing.



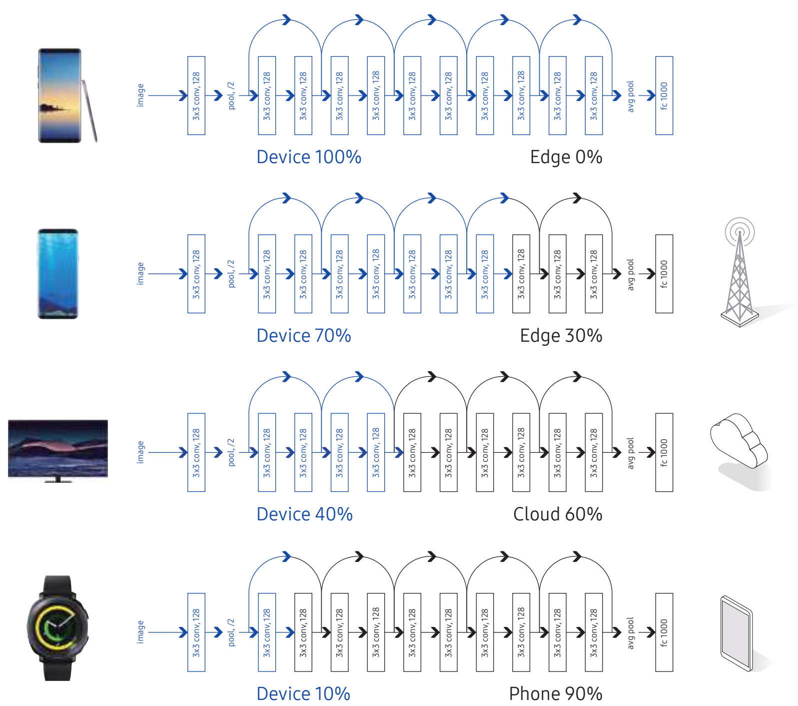
In order to realize the split computing concept, the following factors have to be considered.

- *Software platform:* In general, conventional distributed computing conforms to a client-server model, in which the implementation of each client and server is specific to a given developer. In order to accommodate various mobile devices with different hardware and software from diverse developers, it would be beneficial to develop a split computing platform as open source or as a standard, to the extent that this is possible.
- *Low power and low latency wireless communication:* To support extreme services on a lightweight device such as AR glasses, the device needs low latency wireless communication with low device power consumption.
- *Data synchronization:* A split computing platform partitions the computation of an application between mobile devices and servers. This requires synchronization of a large amount of data, context, and the program itself among network entities.

Figure 21 presents a number of split computing examples. A high-end mobile device may be capable of fully performing the necessary computation by itself. Middle-end devices may be able to support only a part, e.g., 70% or 40%, of the computation, and need to offload the rest of the computation. There may also be low-end devices, such as smart watches, that may need adjacent devices to enhance their performance.

Figure 21

Examples of split computing by various devices.



High-Precision Network

To guarantee high QoE for interactive services with high data rate and low latency requirements, it is important to maintain deterministic E2E latency and to minimize jitter at the microsecond level. High-precision network (HPN) is a solution to achieve this, when paired with massive connectivity supported by both radio link protocols and protocols above radio link. IEEE's time-sensitive networking (TSN) defines mechanisms for the transmission of time-sensitive data over Ethernet. Another solution for implementing HPN is IETF's deterministic networking (DetNet), which specifies a mechanism defined on Internet Protocol (IP) and transport layers. These existing technologies have constraints, since TSN was not designed for mobile networks and DetNet operates on top of TSN. Integrating them with the mobile network is quite a difficult task due to fundamental mismatches between wireless and wired networks. For example, device mobility in mobile network causes changes in the data path far more frequently than would be necessary in wired networks.

In order to realize HPN, the following key features should be considered: multi-pathing, multi-homing, and dynamic mobility. Multi-pathing enables usage of multiple alternate network paths and thereby offers improved reliability and optimal usage of bandwidth resources. Multi-homing is the capability of the end device to connect via multiple interfaces simultaneously. Multi-homing requires support for multi-pathing. Non-IP solutions such as information-centric networking (ICN) could be candidates to provide these features and mitigate shortcomings of the present IP suite. ICN changes the focus of internet architecture from host-centric to data/content-centric. ICN, being content-centric, enables multi-pathing, end-user mobility, and optimal usage of network bandwidth, since data and content can be cached and served by intermediate routers.

In addition, the layered structure of the present IP suite with duplicate lower-layer functionality needs to be enhanced to guarantee deterministic throughput and low latency requirements. Therefore, a new cross-layer protocol design that merges the traditional network and transport layers and combines the functionalities of both layers can be considered. This combined layer may ensure better control over IP and transport layers by taking into account application requirements as well as the network state.

5 6G Timeline

Mobile communication systems have evolved over multiple generations from 2G to 5G approximately every 10 years. Each generation has taken a big step and introduced new technologies. However, we note, as shown in Figure 22, the time spent for defining vision and developing technical standards for each successive generation has shortened from 15 years for 3G to 8 years for 5G. This can be attributed to accelerating growth of technologies and market needs for mobile communications over the past decades.

For 6G, we expect ITU-R will begin their work to define a 6G vision in 2021. Taking into account the trend of speeding up of development of technical standards for each new generation, we expect that the completion of the 6G standard and its earliest commercialization could happen as early as 2028, while massive commercialization may occur around 2030.

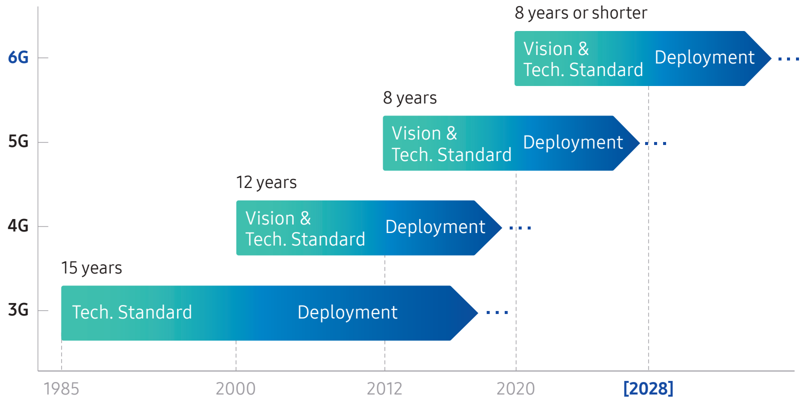


Figure 22

Timeline of different generations.

6 Concluding Remarks

The mobile industry has achieved great successes, from 2G to 4G. While it is still quite important to work to ensure commercial success of 5G in coming years, we believe it is the right time to start preparing for 6G. Shaping 6G will require many years, as we have seen with previous generations in the past. In this spirit, we have presented our initial view of various aspects of 6G including the megatrends, services, requirements, candidate technologies, and timeline for standardization and commercialization. Our view will naturally be updated as we proceed with our research for 6G in the future.

7 References

[1] Cisco, Cisco Edge-to-Enterprise IoT Analytics for Electric Utilities Solution Overview, Available: https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/big-data/solution-overview-c22-740248.html

[2] UN Projects World Population to Reach 8.5 Billion by 2030, Driven by Growth in Developing Countries, Available: https://news.un.org/en/story/2015/07/505352-un-projects-world-population-reach-85-billion-2030-driven-growth-developing

[3] GSMA, 2019 Mobile Industry Impact Report: Sustainable Development Goals, Sep. 2019, Available: https://www.gsmaintelligence.com/research/?file=a60d6541465e86561f37f0f77ebee0f7&download

[4] https://exponentialroadmap.org/wp-content/uploads/2019/09/ExponentialRoadmap_1.5_20190919_Single-Pages.pdf

[5] UN SDG Report 2019, Available: https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019.pdf

[6] M&M, Virtual Reality Market, Available: https://www.marketsandmarkets.com/Market-Reports/reality-applications-market-458.html?gclid=CjwKCAjw7-P1BRA2EiwAXoPWA17McofYdIRbzxQwtSHg-0M9nWNuD09joYOUiYA4N7cI_xTXs0djAhoCK-rUQAvD_BwE

[7] Digi-Capital, AR to Approach $90bn Revenue by 2022, Available: https://advanced-television.com/2018/01/29/digi-capital-ar-to-approach-90bn-revenue-by-2022/

[8] Xuewu Xu *et al.*, "3D Holographic Display and Its Data Transmission Requirement," in Proc. Int'l Conf. Info, Photonics and Optical Commun., pp. 1-4, Oct. 2011.

[9] Pierre-Alexandre Blanche *et al.*, "Holographic Three-Dimensional Telepresence Using Large-Area Photorefractive Polymer," Nature, vol. 468, pp. 80-83, Nov. 2010.

[10] Mordor, Global Holographic Display Market - Segmented by Technology, Available: https://www.mordorintelligence.com/industry-reports/holographic-display-market

[11] GrandViewResearch, Digital Twin Market Size Worth $26.07 Billion By 2025 with

CAGR 38.2%, Available: https://www.grandviewresearch.com/press-release/global-digital-twin-market

[12] https://medium.com/@DAQRI/motion-to-photon-latency-in-mobile-ar-and-vr-99f82c480926

[13] FCC Docket 18-21, "FCC Opens Spectrum Horizons for New Services and Technologies," Mar. 2019.

[14] 3GPP TR 38.807, "Study on Requirements for NR beyond 52.6 GHz," Mar. 2019.

[15] Roger D. Pollard, "Guest Editorial," IEEE Transactions on Microwave Theory and Techniques, vol. 48, no. 4, pp. 625-625, Apr. 2000.

[16] Eirini Karapistoli *et al.*, "An Overview of the IEEE 802.15.4a Standard," IEEE Communications Magazine, vol. 48, no. 1, pp. 47-53, Jan. 2010.

[17] Hadi Sarieddeen *et al.*, "Next Generation Terahertz Communications: A Rendezvous of Sensing, Imaging, and Localization," Available: https://arxiv.org/pdf/1909.10462.pdf

[18] Gustavo A. Siles *et al.*, "Atmospheric Attenuation in Wireless Communication Systems at Millimeter and THz Frequencies," IEEE Antennas and Propagation Magazine, vol. 57, no. 1, pp. 48-61, Feb. 2015.

[19] Kang Ning *et al.*, "A 140-GHz Power Amplifier in a 250-nm InP Process with 32% PAE," Available: https://www.src.org/library/publication/p098807/p098807.pdf

[20] Arda Simsek *et al.*, "A 140 GHz MIMO Transceiver in 45 nm SOI CMOS," in Proc. IEEE BCICTS '18, pp. 231-234, Oct. 2018.

[21] Kaushik Sengupta *et al.*, "Terahertz Integrated Electronic and Hybrid Electronic-Photonic Systems," Nature Electronics, vol. 1, no. 12, pp. 622-635, Dec. 2018.

[22] Ricardo Marqués *et al.*, Metamaterials with Negative Parameters: Theory, Design, and Microwave Applications, John Wiley & Sons, 2007.

[23] John Brian Pendry, "Negative Refraction Makes a Perfect Lens," Physical review letters, vol. 85, no. 18, pp. 3966-3969, Oct. 2000.

[24] Richard W. Ziolkowski *et al.*, "Metamaterial-Based Efficient Electrically Small Antennas," IEEE Transactions on Antennas and Propagation, vol. 54, no. 7, pp. 2113-2130, Jul. 2006.

[25] Chongwen Huang *et al.*, "Reconfigurable Intelligent Surfaces for Energy Efficiency in Wireless Communication," IEEE Transactions on Wireless Communications, vol. 18, no. 8, pp. 4157-4170, Aug. 2019.

[26] Alison M. Yao *et al.*, "Orbital Angular Momentum: Origins, Behavior and Applica-

tions," Advances in Optics and Photonics, vol. 3, no. 2, pp. 161-204, Jun. 2011.

[27] Ove Edfors *et al.*, "Is Orbital Angular Momentum (OAM) Based Radio Communication an Unexploited Area?," IEEE Transactions on Antennas and Propagation, vol. 60, no. 2, pp. 1126–1131, Feb. 2012.

[28] Doohwan Lee *et al.*, "Orbital Angular Momentum (OAM) Multiplexing: An Enabler of a New Era of Wireless Communications," IEICE Transactions on Communications, vol. 100, no. 7, pp. 1044-1063, Jul. 2017.

[29] Zhongshan Zhang *et al.*, "Full-Duplex Wireless Communications: Challenges, Solutions, and Future Research Directions," Proceedings of the IEEE, vol. 104, no. 7, pp. 1369-1409, Jul. 2016.

[30] Kenneth E. Kolodziej *et al.*, "In-Band Full-Duplex Technology: Techniques and Systems Survey," IEEE Transactions on Microwave Theory and Techniques, vol. 67, no. 7, pp. 3025-3041, Jul. 2019.

[31] Meng Ma *et al.*, "A Prototype of Co-Frequency Co-Time Full Duplex Networking," IEEE Wireless Communications, vol. 27, no. 1, pp. 132-139, Feb. 2020.

[32] Min Soo Sim *et al.*, "Nonlinear Self-Interference Cancellation for Full-Duplex Radios: From Link-Level and System-Level Performance Perspectives," IEEE Communications Magazine, vol. 55, no. 9, pp. 158-167, Sep. 2017.

[33] Ruozhu Li *et al.*, "Self-Interference Cancellation with Nonlinearity and Phase-Noise Suppression in Full-Duplex Systems," IEEE Transactions on Vehicular Technology, vol. 67, no. 3, pp. 2118-2129, Mar. 2018.

[34] FCC GN Docket No. 17-258, "Promoting Investment in the 3550-3700 MHz Band," Oct. 2018.

[35] https://www.wirelessinnovation.org/

[36] Citizens Broadband Radio Service, Available: https://www.cbrsalliance.org/

[37] FCC 16-89 Report and Order and Further Notice of Proposed Rulemaking, "Use of Spectrum Band Above 24 GHz for Mobile Radio Services," Jul. 2016.

[38] Junghyun Kim *et al.*, "Deep Learning-Assisted Multi-Dimensional Modulation and Resource Mapping for Advanced OFDM Systems," in Proc. IEEE GLOBECOM '18 Workshops, pp. 1-6, Dec. 2018.

[39] Timothy O'Shea *et al.*, "An Introduction to Deep Learning for the Physical Layer," IEEE Transactions on Cognitive Communications and Networking, vol. 3, no. 4, pp. 563–575, Dec. 2017.

[40] Rubayet Shafin *et al.*, "Self-Tuning Sectorization: Deep Reinforcement Learning Meets Broadcast Beam Optimization," IEEE Transactions on Wireless Communica-

tions, vol. 19, no. 6, pp. 4038-4053, Jun. 2020.

[41] Rubayet Shafin *et al.*, "Artificial Intelligence-Enabled Cellular Networks: A Critical Path to Beyond-5G and 6G," IEEE Wireless Communications, vol. 27, no. 2, pp. 212-217, Apr. 2020.

[42] 3GPP TR 23.791, "Study of Enablers for Network Automation for 5G," Jun. 2019.

[43] O-RAN Alliance White Paper, "O-RAN: Towards an Open and Smart RAN," Oct. 2018.

Samsung Research

research.samsung.com