Communications in the 6G Era

White paper

The focus of wireless research is increasingly shifting toward 6G as 5G deployments get underway. At this juncture, it is essential to establish a vision of future communications to provide guidance for that research. In this paper, we attempt to paint a broad picture of communication needs and technologies in the timeframe of 6G. The future of connectivity is in the creation of digital twin worlds that are a true representation of the physical and biological worlds at every spatial and time instant, unifying our experience across these physical, biological and digital worlds. New themes are likely to emerge that will shape 6G system requirements and technologies, such as: (i) new man–machine interfaces created by a collection of multiple local devices acting in unison; (ii) ubiquitous universal computing distributed among multiple local devices and the cloud; (iii) multi-sensory data fusion to create multi-verse maps and new mixed-reality experiences; and (iv) precision sensing and actuation to control the physical world. With rapid advances in artificial intelligence, it has the potential to become the foundation for the 6G air interface and network, making data, compute and energy the new resources to be exploited for achieving superior performance. In addition, in this paper we discuss the other major technology transformations that are likely to define 6G: (i) cognitive spectrum sharing methods and new spectrum bands; (ii) the integration of localization and sensing capabilities into the system definition, (iii) the achievement of extreme performance requirements on latency and reliability; (iv) new network architecture paradigms involving sub-networks and RAN–Core convergence; and (v) new security and privacy schemes.

Contents

Introduction

With the deployment of 5G systems in full swing, the research focus toward 6G mobile cellular systems has begun. Keeping up with the tradition of a new generation of cellular system once every ten years or so, there is an expectation that a 6G system will be standardized with deployments starting before 2030. Since it often takes more than ten years for a novel technology to see the commercial daylight, it is time to begin research on novel technology components for 6G.

It is essential to establish a vision of future communications to provide guidance for research, which is the purpose of this paper. We attempt to paint a broad picture of communication needs and technologies in the timeframe of 6G. It is possible that some of these requirements can already be met by incorporating new technologies within the 5G framework. In general, we expect to see as part of the 5G evolution the introduction of any modifications that can be introduced in a backward-compatible fashion at a reasonable cost within the 5G framework to meet new performance requirements. On the other hand, modifications that are a fundamental shift and are incompatible with the existing 5G framework or can only be incorporated with high cost to the network or devices will be part of the next generation.

Besides enhanced mobile broadband for consumers, 5G is widely expected to enable the Fourth Industrial Revolution, or Industry 4.0, through the digitalization and connectivity of all things big and small. Digital twins of various objects created in edge clouds will form the essential foundation of the future digital world.

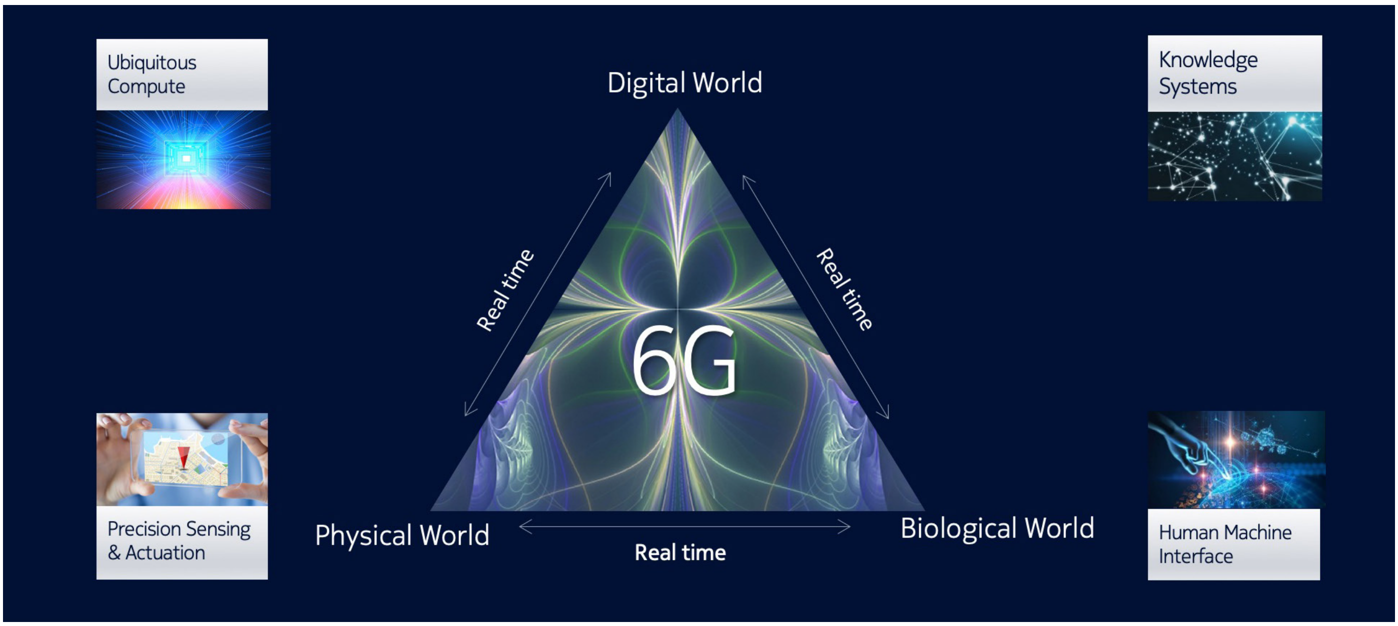
Digital twin worlds of both physical and biological entities will be an essential platform for the new digital services of the future. The realization of a comprehensive digital world that is a complete and true representation of the physical world at every spatial and time instant will require an enormous amount of capacity at low latency. Digitalization will also pave the way for the creation of new virtual worlds with digital representations of imaginary objects that can be blended with the digital twin world to various degrees to create a mixed-reality, super-physical world. As smart watches and heart rate monitors transform into skin patchables, ingestables, body implants, body armor skeleton and brain activity detectors, the biology of humans will be mapped accurately every instant and integrated into the digital and virtual worlds, enabling new super-human capabilities. Augmented reality user interfaces will enable efficient and intuitive human control of all these worlds, whether physical, virtual or biological.

The connectivity of the future is therefore about enabling the seamless integration of these different worlds, illustrated in Figure 1, to create a unified experience for humans, or should we say create an internet of cyborgs. When considering such a future, the following major new themes emerge in addition to the new communication needs: (i) end devices extending from being single entities to a collection of multiple local entities acting in unison to create the new man–machine interface; (ii) ubiquitous universal computing distributed among multiple local devices and the cloud; (iii) knowledge systems that store, process and convert data into actionable knowledge and (iv) precision sensing and actuation to control the physical world.

Recently, several publications have espoused their views on 6G. We take a unique and broader perspective, focusing not only on the technologies but also the human transformation we expect in the 6G era, which helps to provide a view of the performance requirements and design principles for 6G. Our view on the technology transformations starts from where the current 5G systems are, moving to how they are evolving, and then to what may become fundamentally different beyond that. We also address transformations likely to happen in the nature of standardization needed in a world with open platforms.

The paper is organized as follows. In Section 2 we dive deeper into what the world might become in the 2030s, and from that draw relevant new use cases for 6G. This leads to a description in Section 3 of the potential requirements and performance indicators that will distinguish 6G. In Section 4 we discuss some new fundamental dimensions to consider in the design of 6G radio interfaces, in addition to the traditional dimensions of space, spectrum and spectral efficiency. In Section 5 we elaborate on the new technologies that may form the basis for a new generation of cellular networks. In Section 6 we discuss the open platform approach to mobile networking as it seeks to address increasingly specialized requirements, and the consequent implications on the evolution of standardization. We conclude Section 6 with a brief summary.

Figure 1. 6G for the inter-connection of physical, biological and digital worlds



What will 6G be used for?

What will life and our digital society on the other side of the 2030s look like? We begin with devices that humans may use to connect to the network. While the smartphone and the tablet will still be around, we are likely to see new man–machine interfaces that will make it substantially more convenient for us to consume and control information. We expect that:

- Wearable devices, such as earbuds and devices embedded in our clothing, will become common, and skin patches and bio-implants may not be so uncommon. We might even become reliant on new brain sensors to actuate machines. We will have multiple wearables that we carry with us and they will work seamlessly with each other, providing natural, intuitive interfaces. Figure 2 shows the potential evolution in devices.
- Touchscreen typing will likely become outdated. Gesturing and talking to whatever devices we use to get things done will become the norm.
- The devices we use will be fully context-aware, and the network will become increasingly sophisticated at predicting our needs. This context awareness combined with new human–machine interfaces will make our interaction with the physical and digital world much more intuitive and efficient.

The computing needed for these devices will likely not all reside in the devices themselves because of form factor and battery power considerations. Rather, they may have to rely on locally available computing resources to complete tasks, beyond the edge cloud. Networks will thus play a significant role in the man–machine interface of tomorrow.

Figure 2. Likely evolution in devices

As consumers, we can expect that:

- The self-driving concept cars of today will be available to the masses by the 2030s. They will be self-driving most of the time but will still likely need at least a remote driver or the passenger to take control under certain conditions. This will substantially increase the time available for us to consume data from the internet in the form of more entertainment, rich communications or education. The cars themselves will also consume significantly more data: vehicle sensor data will be uploaded in real-time to the network, high-resolution maps will be downloaded and cars will link directly to one another. .
- There will be a massive deployment of wireless cameras as sensors. With advances in AI and machine vision and their capacity to recognize people and objects (or more generally, automatically gather information from images and videos), the camera will become a universal sensor that can be used everywhere. Privacy concerns will be addressed by limiting access to data and anonymizing information. Also, radio and other sensing modalities like acoustics will be used to gather information on the environment.
- Advanced techniques will be used in security-screening procedures to eliminate security lines. A combination of various sensing modalities will be used to screen people as they move through crowded areas rather than only at entrances. Radio sensing will be an essential component of achieving this, supported by the communication systems of the future.
- Digital cash and keys may become the norm, with transactions in both the physical and digital worlds being conducted through the plethora of devices that we will have. The network of the future should provide the security and privacy that is fundamental to such a transformation.
- Numerous domestic service robots will complement the vacuum cleaners and lawn mowers we know today. These may take the form of a swarm of smaller robots that work together to accomplish tasks. The robots will be equipped with video cameras streaming to a local compute server for real-time processing. Thus, we will see an increase in the number of devices and higher capacity requirements within our home networks.
- Health care will be substantially transformed, with 24/7 monitoring of vital parameters for both the healthy and the sick through numerous wearable devices. Health monitoring will also include in-body devices that communicate with wearables outside, which in turn can transport the data to the internet.

The transformation to Industry 4.0 and the first wave of wireless-enabled automation will already have happened before the 2030s. 5G networks providing ultra-reliable low-latency communications (URLLC) will have facilitated real-time processing in the cloud. However, industrial use cases relying on much more extreme requirements for wireless communication will require 6G.

- Holographic telepresence will become the norm for both work and social interaction. It will be possible to make it appear as though one is in a certain location while really being in a different location – for example, appearing to be in the office while actually being in the car. We will have systems that combine current facial expressions with a virtual self within the digital representation of any physical world.
- We will see massive use of mobile robot swarms and drones in various verticals such as hospitality, hospitals, warehouses and package delivery.
- Dynamic digital twins in the digital world with increasingly accurate, synchronous updates of the physical world will be an essential platform for augmenting human intelligence.

Based on the above vision of the future, we can extrapolate the following key use cases. These are a combination of what 5G will enable but with adoption at scale in the timeframe of 6G using new technologies, plus a set of new use cases enabled by the new 6G technologies. See Table 1.

Table 1. Use case for 6G

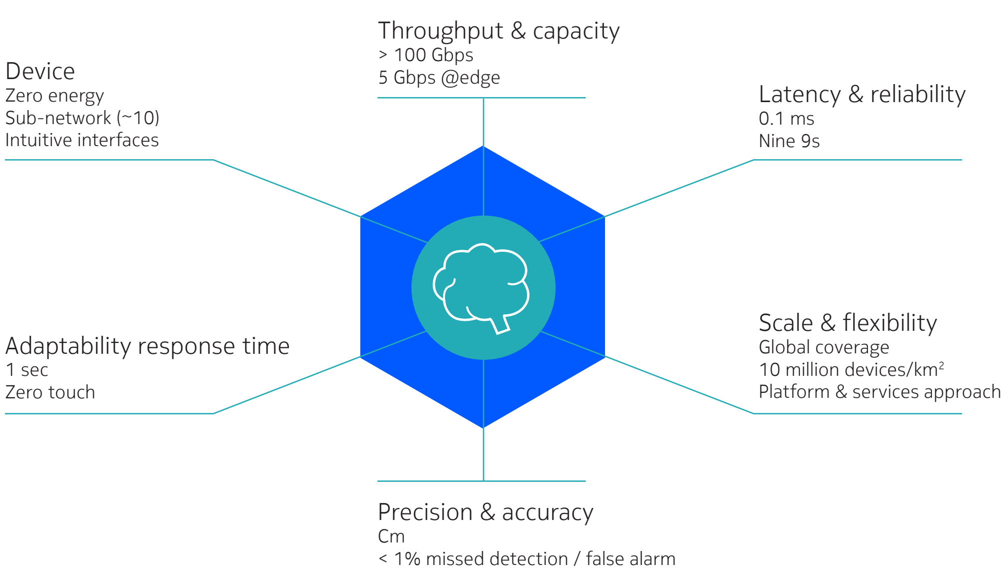
Use case (capability)	5G	6G
Augmented Reality for Industry	Low resolution / high level tasks	High resolution, multi-sensory / detailed tasks, co-design
Telepresence (capacity)	High video quality, limited scale	Mixed reality / Holographic
Security surveillance, defect detection (positioning & sensing)	External sensing, limied automation	Integrated radio sensing, fully automated
Distributed computing, Automation (time synchronization)	Microsecond-level tasks	Higher precision nanosecond-level tasks
Dynamic digital twins and virtual worlds (real-time, multi-sensory mapping and rendering)	No	Yes
Wireless in Data Center (peak rate and capacity)	No	Yes
Zero Energy Devices (back scatter communications)	No	Yes
Swarms of robots or drones	Maybe	Yes
Bio sensors and AI	Limited	Yes

Six requirements and key performance measures for 6G

The array of new use cases we expect to see by 2030 and beyond will drive the new requirements that need to be achieved by 6G. The 5G key performance indicators (KPIs) of data rate/throughput/capacity, latency, reliability, scale and flexibility will continue to be important measures for 6G performance. Several new characteristics will also become important for 6G given the potential use cases described in the previous section. In Figure 2, we group the requirements for 6G into six categories – three categories with KPIs similar to 5G and three new categories:

1. Localization and sensing using the communication network will be an important feature of 6G. We identify precision and accuracy as the corresponding performance measures for localization and sensing, respectively. We expect that centimeter-level precision will be achieved. Object sensing accuracy can be measured in terms of missed detection (MD) and false alarm (FA) probabilities and parameter estimation errors.
2. The network will be engineered with distributed AI/ML techniques embedded in various nodes, and how quickly they adapt to new conditions in the network is an important measure. Network automation will be the norm, and thus how close a network is to complete automation with zero manual intervention will be another criterion.
3. Finally, we expect a major revolution in the end device in the timeframe of 6G. Hence, we introduce a few characteristics under a device category to point out the major transitions that we expect. First, we believe that the end device will evolve in many scenarios to be a network of devices or a sub-network. As examples, we can imagine a machine-area network or a robot-area network involving connecting multiple parts of a machine such as a controller and its drives. Another hallmark of the device in the timeframe of 6G will be that we will have much more intuitive interfaces, with access through gesturing rather than typing, for example. Finally, another possibility for a certain class of device is one that will be extremely low-power and potentially battery-less, relying on the network to power the device.

Figure 3. Key requirements and characteristics of 6G. Object sensing is characterized through missed detection (MD) and false alarm (FA)

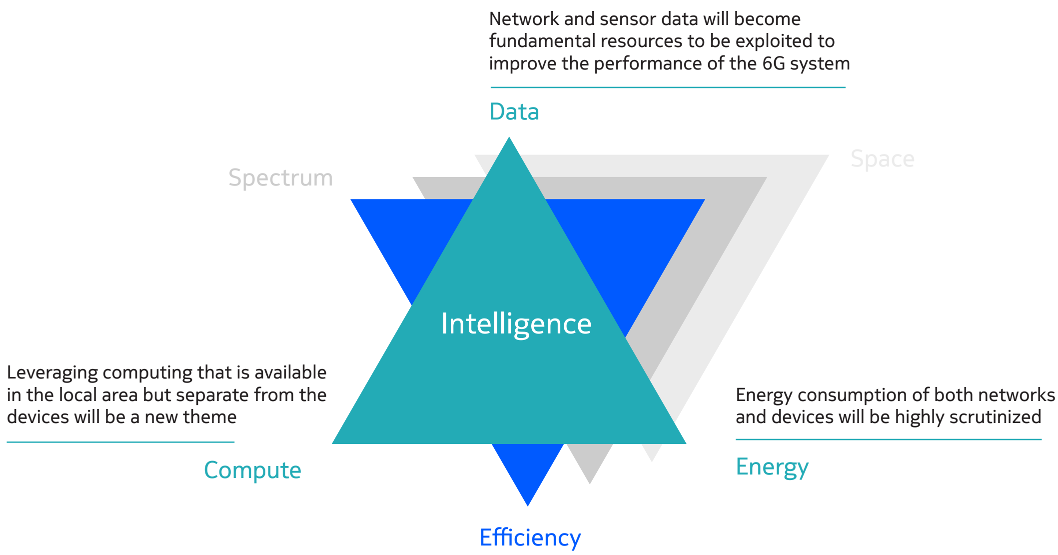


Six fundamental dimensions to design 6G

In every generation until 5G, the three fundamental dimensions of spectrum, spectral efficiency and spatial reuse have dictated how we can grow capacity, as represented by the black triangle in Figure 3. It will continue to be the case for 6G. RF technology will advance to power and cost-effectively use spectrum in even higher bands. There is the opportunity of at least a tenfold increase in the amount of spectrum by going to terahertz-frequency bands. Spectral efficiency will improve through the use of massive multi-user MIMO not only in centimeter wave (cmWave) but also millimeter wave (mmWave) bands as we transition from analog to hybrid/digital beamforming in these lower mmWave bands. As the cost of massive MIMO falls, even larger arrays may be deployed to further increase spectral efficiency. Network densification will undoubtedly continue to increase – not only for capacity reasons but simply to provide increased coverage at higher-frequency bands, at higher data rates and with higher reliability. Furthermore, more pervasive spectrum access will come into play; the sharing between operators of even licensed spectrum powered by software-defined radio (SDR) and AI/ML will allow much higher reuse of spectrum. Efficient spectrum reuse is especially important in the lower bands as they have good non-line-of-sight (NLOS) propagation properties and spectrum resources in those bands are scarce.

We suggest that 6G will fundamentally differ from the previous generations in that three new fundamental dimensions will come into play in addition to the above three traditional dimensions, as shown by the reverse cyan triangle in Figure 3. These dimensions represent the fundamental resources of data, compute and energy. As is well known, AI/ML techniques are data-driven, and whoever has access to large volumes of domain-specific data will be successful in applying these techniques. Application of AI/ML to the design of 6G systems will be fundamental, and similar to various other domains, network and sensor data will become fundamental resources to be exploited to improve the performance of the system. While computing power has always been an important resource for cellular systems, two major trends point in the direction of that becoming a limited resource, and hence how that is exploited in 6G will become significant. The first trend we observe is the emerging saturation in the number of transistors that can be packed into a unit volume, which limits the computing power of devices. The second trend is that we will adopt multiple end devices to augment human sensing capabilities, such as glasses, earbuds and other wearables, which all have very small form factors and hence will have limited computing capability. The current approach of computing offloaded to the edge cloud is unlikely to be sufficient to meet the synchronous computing needs across the different devices. Leveraging computing that is available in the local area but separate from the devices will be a new theme in the 6G timeframe. In this sense, we treat compute as another essential dimension driving the design of the new communication system. Finally, available energy at every element of the network will determine the achievable performance. This ranges from near-to-zero energy at some types of devices, to power supply limits at radio base stations, and to power constraints in data centers. In addition, climate change solutions will become a major focus everywhere in the world by the 2030s, and the growing energy consumption of networks and devices will be highly scrutinized. Thus, energy becomes another important dimension for the design of 6G.
6G will therefore have six fundamental dimensions that research will need to explore to achieve flexible performance targets, as illustrated in Figure 4.

Figure 4. Three new fundamental dimensions for system design



Six key technologies for 6G

A new generation is ultimately characterized by the number of novel, essential technologies that shape the communication system. Truly fundamental new technologies typically take a decade or more to become realized in practice. In view of this, the truly novel technologies forming 6G must be research concepts today. Keeping with the theme of "six" for 6G, we have identified six new potential technology transformation that we expect to be part of shaping the 6G system: (i) AI/ML-driven air interface design and optimization; (ii) expansion into new spectrum bands and new cognitive spectrum sharing methods; (iii) the integration of localization and sensing capabilities into system definition; (iv) the achievement of extreme performance requirements on latency and reliability; (v) new network architecture paradigms involving sub-networks and RAN–core convergence; and (vi) new security and privacy schemes. Each of these is described in the following subsections.

The expanding role of artificial intelligence and machine learning

AI and ML techniques, especially deep learning, have rapidly advanced over the last decade and are now central to several domains involving image classification and computer vision, ranging from social networks to security. They are applied in problem areas where significant amounts of data are readily available for training. Reinforcement learning is beginning to be applied in a variety of robotic control applications following various demonstrations of its prowess in gaming environments, such as AlphaGo.

Recently there has been much exploration of the application of deep-learning techniques to wireless systems. Over the next few years, we expect AL/ML to be applied to 5G systems in at least three different ways. First, they have the potential to replace some of the model-based Layer 1 and Layer 2 algorithms such as channel estimation, preamble detection, equalization and user scheduling, either because they perform better or are less complex. Second, they are likely to be applied extensively in deployment optimization, for example for configuring an optimal subset of beams with which to illuminate the coverage area, taking cell traffic patterns into account. Given the complexity of 5G systems in terms of the sheer number of parameters to be configured at the time of deployment, AI/ML techniques will play an important role in the vision of zero human touch network optimization. Finally, we can expect some other use cases such as localization of end devices using 5G technology to exploit learning techniques for improved accuracy.

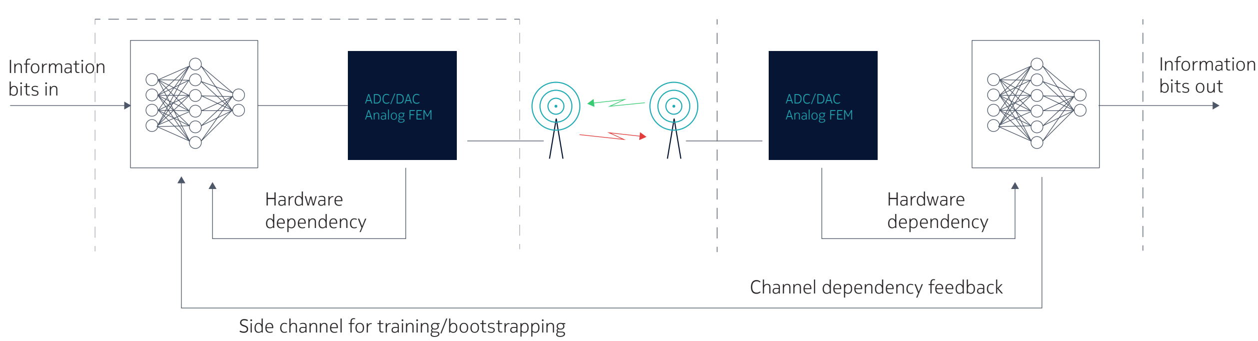
In addition to the use of AI/ML in the RAN, AI/ML will become essential for the 5G end-to-end network automation dealing with the complexity of orchestration across multiple network domains and layers. This will allow for dynamic adaptation of network and cloud resources according to changing demands, rapid deployment of new services and fast mitigation of failures, while significantly reducing operational expenditures.

We envision 6G systems to employ AI/ML in a more fundamental way than the above 5G approach. We expect to go from AI as an enhancement to AI as a foundation for air interface design and optimization – self-optimizing transmitters and receivers, cognitive spectrum use and context awareness.

Self-optimizing transmitters and receivers

Ongoing research has demonstrated that deep-learning systems can learn to communicate over quasi-static links more efficiently than model-based system designs. No explicit design of waveform, constellation or reference signals is required. Through extensive training, a single deep-learning network at the transmitter and one at the receiver learn to pick the best design for these parameters, as illustrated in Figure 5. While such an end-to-end learning approach may be unfeasible for complex, dynamically changing multi-user environments, the 6G communication framework will be designed in such a way as to allow learning in the field to make some design choices. This will enable optimization of the air interface characteristics based on the choice of spectrum, environment, hardware deployed and target requirements. One important shift will be to include the capabilities of the hardware in the optimization of the communication framework. In the current approach, the air interface is designed taking into account some practical limits on implementation. But after the design phase, it is expected that all implementations will have the hardware required for the chosen air interface design. In the future, we can expect the air interface to adapt to the capabilities of the hardware. For example, a certain implementation may have a limited number of analog-to-digital conversion (ADC) or digital-to-analog conversion (DAC) resolutions, which can be taken into account by the learning systems to determine the optimal signaling choice.

Figure 5. End-to-end learning systems adapting to hardware and channels



NOKIA BELL LABS

Cognitive spectrum use

Low-frequency spectrum will continue to be of paramount importance for wide-area coverage due to the superior propagation properties in NLOS compared to higher-frequency bands. Over the next decade, substantial amounts of new spectrum will be allocated to 5G and its evolutions, and that is likely to lead to near-exhaustion of spectrum in bands below 6 GHz. Thus, in the timeframe of 6G, new spectrum-use methods will be required even within the licensed spectrum regime to allow better local access to the spectrum and coexistence with other users. Operators may need to share spectrum among themselves and with other private dedicated networks. And even within a single operator, multiple generations of technologies will coexist and share spectrum. With advances in radio technology enabling multi-band operation and learning techniques such as deep reinforcement learning, efficient autonomous sharing of spectrum can alleviate major spectrum sharing hurdles \cite{FuhuiZhou,PTilghman}. With increasing use of advanced beamforming techniques and densification, use of spectrum becomes highly local, facilitating more reuse spectrum and hence allowing various forms of coexistence among cognitive sharing systems that will be highly beneficial.

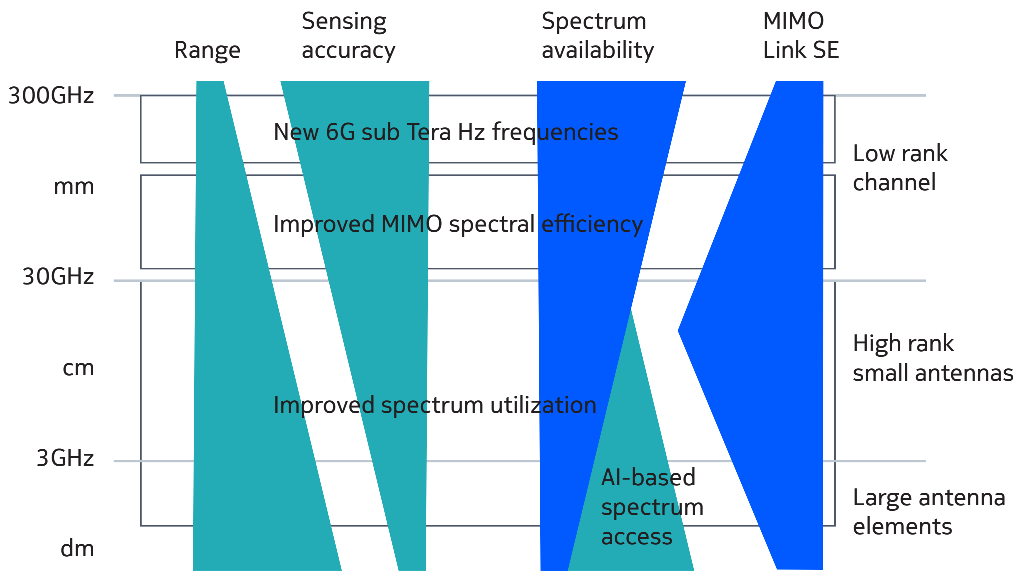
Context awareness

Another major development we can expect in the timeframe of 6G is seamless integration of awareness of the environment, traffic patterns, mobility patterns and location into the optimization of communication schemes aided by new AI/ML techniques. For example, in environments such as factory floors, video cameras will be able to capture the presence and movement of various machines and devices that can be processed in real time through deep-learning networks to predict changes to the propagation environment, which in turn can be used to optimize communication. Essentially, new data acquisition and processing techniques integrated into the communication system can reduce the randomness in the communication links. Long-term mobility patterns can be derived in indoor and outdoor settings that can then be used to optimize the service experience by establishing connectivity to the right technology at the right time. Another important element of future systems may be the use of digitally controlled passive elements such as large-scale meta-surfaces. These are likely to be distributed opportunistically, especially in indoor environments, and new methods are required to exploit this for improving communications. Determining optimal control of these elements using model-based optimization methods may be intractable. It will be challenging to exactly model signal propagation incorporating their collective effects, which in turn depend on how they are controlled. AI/ML techniques will likely be used to solve such complex problems in the 6G era.

Higher-level semantic knowledge of how the communication is being used, for example whether it is for robot control or augmented reality in a factory or for gaming, can be learned from traffic patterns and device characteristics, and appropriate services can be provided automatically. Accurate service personalization down to the lower layers of the communication can be achieved through learning techniques.

Moving from AI for 5G to AI for 6G, we expect that various forms of learning will be employed to realize the above applications. Transfer learning and federated learning will play critical roles. Systems will have to be trained offline in simulation environments to a sufficient extent first so that basic communications can be established, and then be subsequently trained in the field to optimize performance. So there will be transfer of learning from the simulation to the field environment. Devices and network infrastructure have to co-learn to incorporate end-to-end operations, and here, federated learning will play a role. Rather than sharing large data sets between various devices and the network, models will be shared. At the higher layers, deep reinforcement learning will be necessary for optimization of resource allocation and control of various parameters. Hierarchical and multi-agent reinforcement learning will need to be used across different nodes.

Figure 6. Spectrum options for 6G



Exploiting new spectrum bands

The need for higher peak rates and capacity has continuously driven mobile communication systems to utilize higher-spectrum bands. New spectrum bands between 3GHz and 6GHz and from 24GHz to 50GHz have been allocated for 5G in various regions. Entirely new physical layer designs with transition to single-carrier waveforms may be possible within the highly flexible 5G framework. The major challenge to using these high bands has been the realization of high-output power devices at reasonable cost. Massive antenna arrays are employed to form narrow-beam-width transmitters to increase the effective isotropic radiated power (EIRP) and range. Signal propagation in these high bands also poses a challenge, since signals are easily blocked owing to their small wavelength. Diffraction around objects is limited and signal absorption by water is significant. On the other hand, in dense urban and indoor environments, reflections from buildings and walls allow NLOS coverage along the same street or corridor as the access point. Despite these challenges and idiosyncrasies of high-band propagation, a significant amount of research and development is leading to viable deployments.

The trend of using ever-higher band spectrum will continue. In the time of 6G systems, we expect sub-terahertz bands from 114GHz to 300GHz (see Figure 6) to become available and practical for use in cellular systems in specific scenarios. An obvious use case for sub-terahertz spectrum is for backhaul in integrated access and backhaul networks of the future. Narrow beam point-to-point communication in these bands can free up spectrum for access in mmWave bands. Other potential use cases include short-range communications across display and compute devices and rack-to-rack communications in rapidly deployable edge data centers.

A significant amount of ongoing research to improve mmWave systems will also be beneficial for the sub-terahertz systems. Initial product design in those bands will follow the approach established in mmWave systems today. New radio frequency integrated circuits (RFICs) with on-chip or on-board antenna arrays and with phase shifters capable of forming narrow beams will be implemented. New component technologies such as antenna-on-glass could emerge that help reduce the cost of devices. Hybrid beamforming will be needed to achieve massive capacities using single-user or multiple-user MIMO. New receiver architectures with pre-combing before the low-noise amplifier could be introduced. To reduce power consumption, new waveforms designed using AI/ML techniques suitable for single-bit converters are being explored. The propagation delay across the antenna array becomes comparable to the symbol duration with large bandwidth signals, and hence new signal processing schemes to handle this beam squint will be required. New channel measurements and models for using these spectrum bands for access will also be required.

In addition to exploiting sub-terahertz spectrum bands, lower-cost massive MIMO techniques will enable much better use of spectrum in mmWave and cmWave bands. Early mmWave systems relied on analog beamforming and thus are restricted in terms of number of users that can be served simultaneously from a single panel. As network density increases and massive MIMO technologies are cost-reduced, multi-user MIMO will be widely applied in mmWave bands to enable massive-scale, multi-user massive MIMO to exploit the available spectrum. What are considered high bands today will essentially become mid bands in the time of 6G. In the lowest-frequency bands for 6G, namely the lower cmWave, the applicability of massive MIMO becomes gradually restricted by the large size of the antenna elements.

The basic path loss and material penetration properties are much better toward the sub-gigahertz frequencies, and the lower-frequency bands will remain essential for wide-area coverage at the time of 6G.

Spectrum availability is scarce and research on improved spectrum utilization for the lower-frequency bands is important. Spectrum assignment will move from a static split between operators and services toward much more dynamic AI-based spectrum access in time, frequency and space (see Figure 6). Visible light communication is likely to be used in limited scenarios but is unlikely to become a mainstream 6G technology, since radio communications will be cheaper to achieve at the same data rates.

The network with the sixth sense

One of the critical requirements for industrial automation is high-accuracy localization. While real-time kinematics global navigation satellite system (RTK GNSS) can provide highly accurate localization under the conditions of good satellite visibility, many of the automation use cases are indoors, where that is not possible. The current approach to solving the localization problem is to rely on specialized systems based on ultra-wideband (UWB) or Bluetooth Low Energy (BLW) and requires additional access points and devices to be installed. A separate system for localization in addition to the communication system incurs additional infrastructure expense and ongoing maintenance costs that can be avoided if the communication system is also able to perform accurate localization. As a result, 5G includes capabilities to improve localization accuracy and can become the single system for both URLLC and localization in industrial automation environments.

As we move toward 6G, we expect the network to perform various sensing tasks in addition to high-precision localization. Localization solutions will be enhanced to achieve centimeter-level accuracy indoors over a larger area where line-of-sight visibility to a large number of access points is limited. New channel charting methods based on AI/ML techniques applied to large antenna array systems as well as data fusion across RF, camera and other sensors on robots will improve accuracy of sensing even with a limited number of visible access points.

6G systems will be used for imaging of passive objects. System design will not only be optimized for communication but will also incorporate special capabilities for sensing. For example, waveforms suitable for sensing such as chirp signals can be multiplexed with waveforms optimized for communications. Large antenna arrays deployed for massive MIMO communications can be leveraged for forming narrow beams that can be periodically swept for sensing. Multiple transmitters and receivers can coordinate to enhance the sensing capabilities of the network. The evolution to sub-terahertz and terahertz bands, with the associated large signaling bandwidth, increases the opportunity for precision sensing. Millimeter-precision imaging using terahertz-band infrastructure will enable a significant number of new use cases in industry automation and health care, such as fault detection in extrusion manufacturing processes or detection of cancerous tissue and tooth cavities. There are plenty of applications that will benefit from radio points being turned into sensors, such as food quality control in supermarkets, or invisible metal detectors in airport or event infrastructures to replace the current security gate systems.

Combining the multi-modal sensing capabilities with the cognitive technologies enabled by the 6G platform will allow for analyzing behavioral patterns and people's preferences and even emotions, hence creating a sixth sense that anticipates user needs. It will allow for interactions with the physical world in a much more intuitive way.

Extreme networking

The new Industrial Internet of Things (IIoT) use cases targeted by 5G rest on achieving ultra-low latency of 1ms over the air with five-nines reliability. The main techniques introduced to achieve these performance targets are mini-slots and grant-free faster channel access for low latency and multiply connected links involving multiple access points, carriers and packet duplication for reliability.

Even 1ms latency over the air is insufficient for many use cases, for example in replacing traditional industrial wireline connectivity solutions such as Sercos or EtherCAT. For such use cases, substantially lower radio latencies of the order of 100μs at gigabit-per-second data rates are required. Furthermore, the actual requirement on reliability is based on the actual downtime of the equipment, the onset of which is triggered by multiple consecutive packet losses. The target for such a reliability measure can be of the order of nine-nines for some industrial automation use cases.

6G will be designed to achieve these extreme requirements cost-effectively. Recent measurements show that, contrary to conventional wisdom, the blocking of signals in mmWave is not that severe in factory floors. By relying on wider bandwidth available in mmWave spectrum, it will possible to achieve such extreme low latencies at the desired high data rates. Reliability can be enhanced through simultaneous transmission through multiple paths involving multiple wireless hops. Cooperative relaying through device-to-device connections can be used to create separate paths from the network to the device. Predictive beam management using AI/ML prediction techniques can also substantially reduce uncertainties in link quality.

In 4G, low-power wireless access was introduced through Narrowband IoT (NB-IoT) technology. Another extreme form of networking that we can expect in 6G is going from low-power to zero-energy devices for IoT. Wireless zero-energy devices are well known as passive radio frequency identification (RFID) tags, with low-cost active RFIDs supporting sensors typically having a battery life of three to five years, but these are generally restricted to very short ranges. There are several use cases where sustainable sensor devices and extremely long operation times are desired in a wider-area context. For construction inspection of bridges or tunnels, as an example, it is desirable to have wireless sensor devices that can be fully embedded in the construction and that can operate on the order of 100 years without human intervention. Potential solutions will likely include a mixture of low-power communication, extreme low-idle current, energy harvesting (potentially from the communication network) and reachable energy storage.

6G network architecture concepts

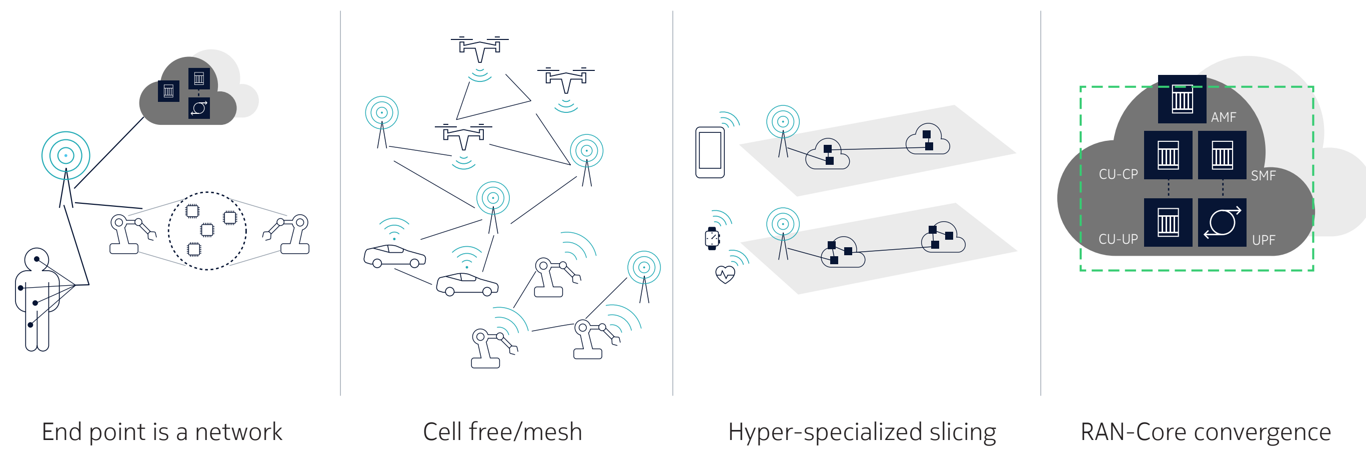
Sub-networks

The cellular network architecture in previous generations has been designed primarily for extending the voice and data internet to individual, mobile end points. 5G is the first system designed to make inroads into the industrial environment, meeting the challenging requirements through new architectural evolutions such as supporting time-sensitive networking (TSN) bridge functionality. To build on the road paved by 5G, and become truly entrenched in the industrial environment and replace wired connectivity everywhere, 6G should provide deterministic wire-grade reliability for a variety of connectivity scenarios (Figure 7), from static, isolated devices, to inter-related locally interacting devices, to rapidly moving swarms of robots and drones that need to inter-connect but also connect directly to the network when separated from the swarm. To ensure both high time and spatial domain reliability and determinism, we see the need for semi-autonomous 6G sub-networks, where at least the most critical services in the sub-network will continue uninterrupted despite poor or no connectivity to the wider network. Multiple path connectivity employing infrastructure and opportunistic device-to-device connections will be required for the ultra-reliability, potentially leading to truly cell-less architectures. Integration of these sub-networks to 6G as one holistic architecture has some advantages:

- The 6G sub-network will ensure high data rates, extreme low latency, reliability and resilience
- 6G security and resilience features are enforced to the lowest level of devices in the sub-network
- 6G service execution can dynamically be split between execution in the edge cloud or in the device that is part of the sub-network.

Time-sensitive communications (TSC), through integration of the 5G network and TSN, with the 5G network acting as a TSN bridge, will evolve to 6G to provide native TSN, including over wider areas with mobility.

Figure 7. 6G architecture themes



Hyper-specialized slicing

Beyond the extension of the traditional connectivity architecture into a variety of sub-network and multiply connected scenarios, we expect further advances in slicing and virtualization. Slices can become highly specialized, potentially with separate software stacks in each slice for different functional treatment of the flows, as illustrated in Figure 7. The current trend in virtualization of the higher layers of the RAN will lead to further disaggregation of RAN functions into modular micro-services that can be flexibly composed into slice-specific RAN implementations. For example, one can envision slice specializations for a video service slice incorporating specific video optimization micro-services included in that RAN slice but not necessary in other slices. Similarly, low-throughput IoT slices can incorporate functions allowing connectionless access while other slices are based on a traditional access approach. In addition, we can expect flexible slice-specific function placement in gateway devices, relays, cell sites, far edges, edges and regional clouds across a variety of different hardware platforms according to the needs of the slice. New innovations in service management and orchestration are needed to create and manage such highly specialized slices.

RAN–Core convergence

In 5G, the base station has been compartmentalized into the distributed unit (DU) and centralized unit (CU). The DU includes the lower layers of the user and control plane protocol stack, namely the physical Layer 1 and real-time Layer 2, while the CU includes the non-real-time Layer 2 and Layer 3 functions. The CU is further split into the control plane and user plane, with a well-defined interface between the two. The CU is typically implemented as a virtualized function in the edge or metro cloud and can serve multiple DUs. The 5G core functions, on the other hand, are becoming more decentralized as the amount traffic through the core increases substantially. The various core functions are also being virtualized and implemented in regional or metro clouds, or even edge clouds for low-latency services. With increasing centralization of the higher-layer RAN functions and the distribution of the core functions, simplification can be achieved by combining some of the RAN and core functions into single entities. Thus, in the timeframe of 6G, we expect a reduced set of functional blocks implementing a combination of 5G RAN and core, resulting in a “coreless” RAN, especially on the user plane.

New security, privacy and trust paradigms

Wireline-grade reliability also implies that the network must be designed with new security and privacy measures. Jamming in industrial networks is a new threat that the networks will have to be protected against. Attackers could attempt to jam networks from outside the industrial facility, and so physical security will be insufficient. In the future, jamming may also take the form of simply delaying packet delivery by creating interference only sporadically. This can seriously impact industrial operations relying on time-sensitive networks. 6G networks will be designed to protect against such new threats.

The definition of sub-networks in networks requires a change to the authorization strategy. It is no longer authorization by the network, but by the sub-network. When analyzing a body area network (BAN), the assets in the sub-network belong to the sub-network, and therefore the authorization and asset management must be handled in that trust boundary. The network will connect sub-networks, and a second level of authorization might be needed at network level. Different sub-networks might belong to mutual untrusted entities, which calls for clear separation between sub-networks but also between the network and sub-networks. It will be crucial for the sub-networks to act as an independent network, empowered as an authorization authority and responsible for sub-network asset management. Due to the dynamic behavior of devices joining/leaving the sub-network, maintaining sub-network privacy and potential anonymity will be a challenge to be solved in the 6G network architecture.

When the physical and biological worlds are mirrored in the digital world with high precision and when new mixed-reality worlds combining digital representations of real and virtual objects are created, privacy solutions of today are unlikely to be sufficient. Although multi-modal sensing will capture nearly everything in their surroundings, users will want to restrict what others are allowed to experience in the content that they share. Users must be able to set a rich set of preferences in a simple fashion on what they wish to share, and data processing should automatically ensure that. A variety of new signal processing techniques are emerging to address security in the mixed-reality world of tomorrow that will become an integral part of the 6G network. Physical layer security mechanisms typically depend on the uniqueness of the wireless channel to establish authentication, confidentiality and key exchange, and may become more mature in the timeframe of 6G, addressing new issues such as jamming. Trust in the network is critical for the success of 6G.

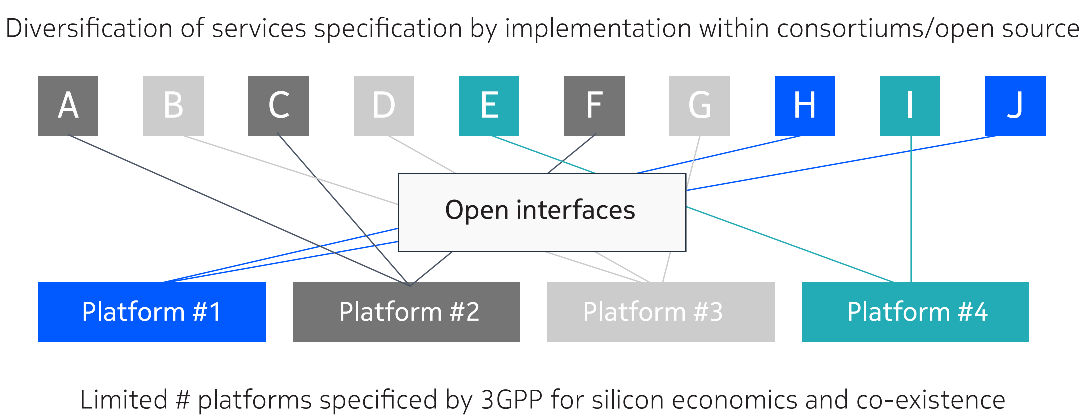
6G as an open platform for the creation of highly specialized solutions

One important factor for the huge success of mobile communication has been global open standards. Until recently, mobile communication services have been dominated by human-centric services like voice, short messaging service (SMS) and best-effort broadband data. Tremendous volumes have been achieved, with over a billion smartphones in use and tens of millions of mobile network base stations deployed.

During the last five years, mobile communication services have expanded to now include IIoT. First, low-power wide-area (LPWA) connectivity solutions such as NB-IoT and Cat-M were standardized. Next, Cellular V2X connectivity for cars followed. UAV connectivity for drones has been specified. Most recently, in 5G, URLLC and TSC for real-time control are being standardized. LPWA and C-V2X connectivity devices have the potential to achieve volumes in the order of billions.

As the applicability of the cellular network expands into new IIoT and future home and enterprise environments, we foresee the need for much more specialized connectivity solutions, optimized for the specific requirements. The natural consequences are that there will be significantly smaller volumes for each specialized connectivity solution. To efficiently support this long-tailed distribution of new wireless connectivity solutions, we foresee the need to specify and introduce mobile network specifications as a platform with a few core capabilities at the lower layers. This will be needed for the different scenarios, which will then be utilized by several different interest groups to specify their own higher-layer specifications to achieve a complete connectivity solution for their own use case family. The latter may in some cases be realized through joint software development within the interest group, or open source software. This is illustrated in Figure 8. A major advantage of this platform approach is the open interfaces from the platform to the specialized layers of connectivity. We believe that the lower-layer communication protocols typically implemented in silicon still need to be specified by a global standardization body to ensure economy of scale and coexistence among connectivity services.

Figure 8. Illustration of 6G as a solution platform model



The current trend toward open RANs with multiple vendors providing different pieces of the network will be further facilitated by the above network as a platform approach in 6G. Apart from the radio and some of the processing-intensive functions that are best realized in custom hardware, the rest will be software functions running on any commercial compute hardware, and the interface specifications will also become open. The intelligent control layers of the mobile network can thus be optimized to suit the connectivity needs of the specific use case family and can come from any vendor. Note that this is also in line with the notion of the next generation of slicing that was described earlier, where each slice can have its own specialized function.

Mobile networks and terminals as a platform as such is not a 6G technology, but rather a new way of efficiently and rapidly specifying solutions for the long-tailed distribution of specialized industrial and other indoor use cases.

Summary

The tradition of deploying a new cellular generation approximately every decade will continue into the future, with 6G becoming a reality in the 2030s. 6G will be optimized and cost-reduced for the new use cases introduced in 5G, driving their adoption at scale. At the same time, it will enable new use cases that we cannot yet imagine or describe in detail. The expansion of mobile cellular to verticals that began with the introduction of low-cost IoT technologies in 4G and ultra-reliable low latency IIoT in 5G will continue, becoming both broader and deeper in 6G. The rapid advance of AI/ML technology and its effectiveness in solving problems in several domains points toward a 6G system that will fundamentally exploit these new capabilities to improve performance by better adapting to the operational environment.

The inexorable demand for higher capacity and peak rates points toward technologies that will exploit ever-higher bands. As the density of infrastructure increases, coupled with the use of wider bandwidth signals at high-band spectrum, especially indoors, new opportunities to utilize this for localization and sensing will encourage a 6G design that is not only optimized for communication but also for perception and understanding of the physical world and people's needs, thus augmenting human existence in the most intuitive way. In this paper, we identified the following key technology transformations as having the highest potential to be defining for the 6G system: (i) AI/ML-driven air interface design and optimization; (ii) expansion into new spectrum bands and new cognitive spectrum sharing methods; (iii) the integration of localization and sensing capabilities into system definition; (iv) the achievement of extreme performance requirements on latency and reliability; (v) new network architecture paradigms involving sub-networks and RAN–core convergence; and (vi) new security and privacy schemes. Finally, the expansion into many varied use cases calls for a shift to a platform approach to the network, with decoupling of the air interface from networking.

Authors

Harish Viswanathan and Preben Mogensen

Sources

The content of the white paper mirrors an existing IEEE paper "Communications in the 6G Era", published in March 2020.

Acknowledgments

We wish to thank a number of colleagues at Nokia Bell Labs with whom we had numerous discussions about the future of wireless. In particular, we would like to thank Harri Holma, Peter Merz, Mikko Uusitalo, Peter Vetter, Thorsten Wild and Volker Ziegler for their comments on an early draft of this paper. Special thanks also to Amanda Hu, Kevin Fritchard and Ramoni Adeogun for helping with the preparation of the manuscript.

Nokia OYJ
Karakaari 7
02610 Espoo
Finland
Tel. +358 (0) 10 44 88 000

Document code: CID207766 (February)