1

Introduction

Shilpa Talwar and Rath Vannithamby
Intel Corporation, USA

1.1 Evolution of Cellular Systems through the Generations

The first large-scale commercial cellular communications systems were deployed in the 1980s and these became known as first-generation (1G) systems. 1G systems were built on narrowband analog technology, and provided a basic voice service. These were replaced by second-generation (2G) cellular telecom networks by the early 1990s. 2G networks marked the start of the digital voice communication era, and provided a secure and reliable communication channel. 2G systems use either time division multiple access (TDMA) or code division multiple access (CDMA) technologies, and provided higher rates. The European Global System for Mobile Communications system is based on TDMA technology while IS-95 (also known as CDMA One) is based on CDMA technology. These 2G digital technologies provide expanded capacity, improved sound quality, better security and unique services such as caller ID, call forwarding, and short messaging. A critical feature was seamless roaming, which let subscribers move across provider boundaries.

The third-generation (3G) – International Mobile Telecommunications-2000 (IMT-2000) – is a set of standards for mobile phones and mobile telecommunications services fulfilling the recommendations of the International Telecommunication Union-Radio (ITU-R). 3G mobile networks became popular due to ability of users to access the Internet over mobile devices and laptops. The speed of data transmission on a 3G network is up to 2 Mbps, and therefore the network enables voice and video calling, file transmission, internet surfing, online TV, playing of games and much more. 3G uses CDMA technology in various forms. Wideband CDMA and High Speed Packet Access technologies were developed as part of the Third Generation Partnership Project (3GPP) organization, and CDMA2000 was developed as part of the 3GPP2 organization.

Fourth-generation (4G) requirements – the International Mobile Telecommunications Advanced (IMT-Advanced) specification – were specified by ITU-R in March 2008. The key requirements specified 4G peak service speeds of 100 Mbps for high-mobility communication (such as from trains and cars) and 1 Gbps for low-mobility communication (such as pedestrians and stationary users). A 4G system not only provides voice and other 3G services but also provides ultra-broadband network access to mobile devices. Applications vary from IP telephony, HD mobile television, video conferencing to gaming services and cloud computing. There are two 4G technologies: Long-Term Evolution (LTE) and Worldwide Interoperability for Microwave Access (WiMAX). LTE was developed as part of 3GPP and WiMAX was developed as part of IEEE. LTE uses orthogonal frequency division multiple access (OFDMA) in the downlink and single carrier frequency division multiple access in the uplink whereas WiMAX uses OFDMA in both uplink and downlink.

1.2   Moving Towards 5G

4G standards were completed in 2011 and networks are currently being deployed. The attention of the mobile research community is now shifting towards what will be the next set of innovations in wireless communication technologies, which we will refer to collectively as 5G (fifth-generation technologies). Given a historical 10-year cycle for every generation of cellular advancement, it is expected that networks with 5G technologies will be deployed around 2020. Similar to 3G/4G, where ITU-R issued a recommendation for IMT-2000/IMT-Advanced [1], ITU-R has recently released a recommendation for the framework and overall objectives of the future development of systems for 2020 and beyond [2]. This highlights the emerging consensus on the use cases and requirements that systems deployed in 2020 must address. These include requirements for new services such as smart grids, e-health, autonomous transport, augmented reality, wireless industry automation, remote tactile control and so on, which cannot be met by IMT-2000 systems.

The usage scenarios envisioned for IMT for 2020 and beyond can be broadly classified as follows:

Enhanced Mobile Broadband   The dramatic growth in the number of smartphones, tablets, wearables, and other data-consuming devices, coupled with the advent of enhanced multimedia applications, has resulted in a tremendous increase in the volume of mobile data traffic. According to industry estimates, this increase in data traffic is expected to continue in the coming years and around 2020 cellular networks might need to deliver as much as 100–1000 times the capacity of current commercial cellular systems [3, 4]. While the roll-out of 4G technologies with their expected enhancements will address some of capacity demands of future mobile broadband users, a mobile broadband user in 2020 will expect to be seamlessly connected all the time, at any location, to any device. This poses stringent requirements on the 5G network, which must provide users with a uniform and seamless connectivity experience regardless of where they are and what device/network they connect to.

Massive Machine-type Communications   This use case refers to the growing interest in the area of machine-to-machine (M2M) communications and the Internet-of-Things (IoT). Together, these represent a future in which billions of everyday objects are connected and

managed through wireless networks and management servers [5]. One can envisage creating an immensely rich set of applications by connecting the thousands of objects surrounding us. Examples include:

- smart homes, in which intelligent appliances autonomously minimize energy use and cost
- remote monitoring of expensive industrial or medical equipment
- remote sensing of environmental metrics such as water pressure, air pollution and so on.

These applications and services demand communication architectures and protocols that are different from traditional human-based networks. The integration of human and machine-type traffic in a single 5G network is therefore a challenge. In addition, IoT traffic can be quite diverse, from low to high bandwidth, from delay-sensitive to delay-tolerant, from error-tolerant to high reliability, which poses additional complexity. This use case focuses on applications where a very large number of connected devices transmit relatively low volumes of non-delay-sensitive data. The devices are typically low-cost and low-complexity, and require a very long battery life.

Ultra-reliable and Low-latency Communications.   This use case addresses IoT applications that have stringent requirements for reliability, latency, and network availability. Examples include:

- connected cars, which react in real time to prevent accidents
- body area networks, which track vital signs and trigger an emergency response when life is at risk
- wireless control of industrial manufacturing or production processes.

As evidenced by diverse set of usages anticipated by 2020, the 5G system will require enhancements to performance metrics beyond the "hard" metrics of 3G/4G, which included peak rate, coverage, spectral efficiency, and latency. The 5G system will see expanded performance metrics centered on the user's quality of experience (QoE), including factors such as ease of connectivity with nearby devices, connection density, area traffic capacity, and improved energy efficiency. The eight parameters in Table 1.1 are considered to be key capabilities of IMT-2020 systems. Their target values are also summarized. These are currently recommendations, and subject to further research and technological development [2].

1.3   5G Networks and Devices

As it can be seen from the description above, 5G networks will have to accommodate diverse types of traffic, spectrum, and devices. The network itself is anticipated to consist of hierarchical nodes of various characteristics and capacities. The 5G network will support multiple radio access technologies (RATs), such as 3G/4G/5G, WiFi, and WiGig, and also multiple modes ranging from ultradense small cells, device-to-device (D2D) communications, and new sub-networks oriented toward wearable devices. Inevitably, the user experience and quality will need to be maintained as users move along various networks and get connected to the various types of node. 5G networks will likely use a multi-layer network

Table 1.1   Key parameters of IMT-2020 systems.

Parameter	Details	Target
Peak data rate	Maximum achievable data rate under ideal conditions per user/device	10–20 Gbps
User-experienced data rate	Achievable data rate that is available ubiquitously across the coverage area to a mobile user/device	100 Mbps–1 Gbps, depending on wide-area or hotspot coverage
Latency	Time contribution by the radio network from the time from when the source sends a packet to when the destination receives it	1 ms over-the-air latency
Mobility	Maximum speed at which a defined QoS and seamless transfer between radio nodes which may belong to different layers and/or radio access technologies (multi-layer/-RAT) can be achieved	To provide high mobility up to 500 km/h with acceptable QoS
Connection density	Total number of connected and/or accessible devices per unit area	To support a connection density of up to 10^6/km^2, for example in massive machine-type communication scenarios
Energy efficiency (a) Network side	Quantity of information bits transmitted to/received from users, per unit of energy consumption of the radio access network (RAN) (in bit/Joule)	Target is at least 10x on network energy efficiency The 5G network must not consume more energy, while providing enhanced features
(b) Device side	Quantity of information bits per unit of energy consumption of the communication module (in bit/Joule)	
Spectrum efficiency	Average data throughput per unit of spectrum resource and per cell (bit/s/Hz)	3–5× increase in spectrum efficiency
Area traffic capacity	Total traffic throughput served per geographic area	10 Mbit/s/m^2 in hotspot scenarios

architecture, where the macro layer provides coverage to users moving at high speeds or for secure control channels, while a lower layer comprising network nodes with smaller capabilities provides high data rates and connectivity to other RATS (say, WiFi or new mmWave RATs). Moreover, a 5G device may have simultaneous active connections to more than one network node, with the same or different RATs, each connection serving a specific purpose, for example one connection to a given node for data and a second connection to another node for control. In addition, the use of remote radio heads connected to central processing nodes with the aid of ultra-high-speed backhaul is expected to be extended to more areas. Fast and high-capacity backhaul will enable tighter coordination between network nodes in a larger area. All of these changes will require a high level of integration of different nodes in the network and of technologies located even within the same node. In short, the 5G system will

need to provide a flexible technological framework in which networks, devices, and applications can be co-optimized to meet the great diversity of requirements anticipated by 2020.

As the 5G usage models and networks evolve, 5G device architectures will also be more complex than in 4G. Devices will be capable of operating in multiple spectrum bands, ranging from RF to mmWave, while being compatible with existing technologies such as 3G and 4G. The need to support several RATs with multiple RF-chains will impose tremendous challenges for 5G device chipset and front-end module suppliers, as well as system and platform integrators. Another key feature of 5G devices will be their advanced interference suppression capabilities. The dense deployment of network nodes and increasing sources of interference will require that the devices deployed autonomously detect, characterize, and suppress interference from any source: intra-cell, inter-cell, or D2D. The task of interference cancellation will be exacerbated by the existence of strong self-interference in the case of simultaneous transmission and reception. In addition, devices will be required to actively manage all the available network connections, including D2D links, as well as to share contextual information with network layers so that network resources can be efficiently utilized. All of these enhanced features will need to be implemented in such a way that energy consumption is optimized for a small wireless device platform.

1.4   Outline of the Book

In this book we bring together a group of visionaries and technical experts from academia and industry to discuss the applications and technologies that will comprise the 5G system. It is expected that some of the new technologies comprising 5G will be evolutionary, covering gaps and enhancements from 4G systems, while some of the technologies will be disruptive, covering fundamentally new waveforms, duplexing methods, and new spectrum. These technologies will encompass the end-to-end wireless system: from wireless network infrastructure to spectrum availability to device innovations.

The book is organized into three parts. Part I has four chapters. In Part I, we provide an overview of 5G, address trends in applications and services, and summarize 5G requirements that will be need to be addressed in next-generation technologies and system architectures. We also provide an overview of some 5G research programs around the world: Horizon 2020 in Europe and Intel's 5G University Research Program in USA.

Part II has nine chapters. In Part II, we address evolutionary technologies that will be needed to meet 5G requirements, including:

• co-operative radio access architectures to enable greater energy efficiency and network performance
• small-cell networks with in-built caching
• multiple RAT integration, which is inevitable to provide a seamless user experience
• distributed resource allocation
• advances in device-to-device communications
• energy-efficient network design
• multi-antenna processing and interference co-ordination techniques
• design for M2M communications
• design for ultra-low latency.

These technologies are already being developed in 3GPP Release 11 and beyond as part of the evolution of 4G systems.

Part III has five chapters. In Part III, we discuss "revolutionary" candidate technologies: those that are essentially disruptive and different from 4G. These include:

- new physical layer waveforms that offer enhanced flexibility and performance
- massive MIMO technologies that enable large numbers of simultaneous users
- mmWave technologies to harness new spectrum for access and backhaul
- simultaneous transmit and receive on the same time/frequency resource
- software defined networking and network function virtualization to enable software-based flexible infrastructures.

References

[1] ITU-R, Recommendation M.1645: "Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000", June 2003.

[2] ITU-R, Document 5D/TEMP/625-E: "IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond", 17 June 2015.

[3] Cisco, "Cisco Visual Network Index: Global mobile traffic forecast update", 2013.

[4] Ericsson, "Traffic and market data report," 2011.

[5] Ericsson, White paper "More Than 50 billion connected devices", 2011. URL: http://www.ericsson.com/res/docs/whitepapers/wp-50-billions.pdf.