

1

Introduction

Anthony C.K. Soong¹ and Rath Vannithamby²

¹ Futurewei Technologies, Plano, TX, USA

² Intel Corporation, Portland, OR, USA

Abstract

This chapter looks at the features in the first phase of fifth generation (5G) that supports the vertical industry. Moving onto future phases of 5G, Third Generation Partnership Project and the cellular industry has invited and offered to work with all vertical industries to define additional requirements for future 5G releases. One interesting vertical industry, for example, is professional audio production which requires strict synchronization of devices to function. The chapter discusses the significance of network slicing for the support of the vertical industry services. It examines a number of other issues that impact the vertical services such as virtual network function placement, multi-access edge computing, and artificial intelligence. The chapter further presents an overview of the key concepts discussed in the subsequent chapters of this book. The book examines the impact of 5G cellular communications on the various vertical industries.

Keywords 5G cellular communications; network issues; network slicing; radio access; Third Generation Partnership Project; ultra-reliable and low-latency communications standardization; vertical industries

1.1 Introduction

It is without a doubt that commercial wireless cellular communication has changed how we, as human beings, interact with each other. As young graduate students over 30 years ago, we could not have imagined how enhancing our

5G Verticals: Customizing Applications, Technologies and Deployment Techniques, First Edition.

Edited by Rath Vannithamby and Anthony C.K. Soong.

© 2020 John Wiley & Sons Ltd. Published 2020 by John Wiley & Sons Ltd.

communications would have altered the way we interact in the world. At that time, our dream was to bring ubiquitous wireless telephony to the world. The attraction of that dream is obvious, to give users the freedom of movement while being able to continue a telephone call. There was no discussion about the killer application because voice was the only application. Commercial wireless systems were in their infancy, analog modulation was king, and only the very elite would have a telephone in their car. If you wanted to take the phone with you, it was carried in its own briefcase because the size of the phone was that large.

Decades ago, the concept of cellular was considered new. It necessitated the development of base stations serving a small area, or cell, because of the requirement of providing high capacity mobile telephony that did not require a very large number of channels. Arguably, cellular communication was born on 4 January 1979, when the Federal Communications Commission authorized Illinois Bell Telephone Co. to conduct a trial of a developmental cellular service in the Chicago area. Around the same time, American Radio Telephone Service Inc. was authorized to operate a cellular service in the Washington–Baltimore area. The feasibility and affordability of cellular services where the same channel may then be re-used within a relatively small distance was then demonstrated. More importantly, from a communication system point of view, the concept of capacity increase by densification¹ was established and full commercial service first began in Chicago in October 1983.

From these humble beginnings, cellular communication has grown into a common and necessary part of everyday human interaction. The transformation of the cellular phone from a telephony device to a pillar of human social interaction can be laid at the development of the so-called “smart phone”. Humanity changed in 2008 and is now dependent upon the applications on the smart phone as a harbinger of information, as well as to enable socialization with others in an individualized way. Indeed it is ubiquitous customized socialization that has transformed our society. In some sense, it has not only brought us closer to each other but also closer to our humanity.

As we move into the era of 5G cellular communication, humanity will move from the age of human social communication to a world where communications is fusing together the physical, digital, and biological worlds; the so-called “fourth industrial revolution” [3]. This will represent an unprecedented opportunity to transform our industry and simultaneously drive profitability and sustainability. Innovations will enhance the production cycle and connect manufacturers with their supply chains and consumers to enable closed feedback loops to improve products, production processes, and time-to-markets. As an example of

¹ It should be noted that almost all of the gain in the capacity of the commercial wireless communication system in history has come from densification of the network. A discussion of fifth generation (5G) densification can be found in [1, 2].

the potential economic impact, the World Economic Forum studied the impact of this on the state of Michigan [4]; the birthplace of modern automotive mass manufacturing. Today, the manufacturing environment of the automotive industry is undergoing unprecedented change spurred on by the changing expectation of digital consumers, the emergence of the smart factory, and the rise of connected vehicles.² This change is coinciding with the transition from mass manufacturing to hyper-customization demanded by the customer. No longer are consumers looking to just buy a product; they are looking for a complete customized user-centric experience.

Sustainability has now also become a prime concern for consumers and manufacturers alike. Embracing sustainability and green principles is not just a mere marketing tool anymore. Consumers and regulators are no longer satisfied when economic growth happens at the expense of the environment and are setting higher and higher sustainability requirements. The fourth industrial revolution presents an opportunity to decouple this relationship by providing both economic growth while enhancing simultaneously the environment [4]. As an example, the auto manufacturers are, more and more, making bold commitments to sustainability. The vision of General Motors (GM) of zero crashes, zero emissions, and zero congestion demonstrates the auto industry's strategy of coupling growth with sustainability. As part of this, GM is committed to using 100% renewable energy by 2050. Both GM [5] and Ford [6] have committed to the United Nations 2030 sustainability goals [7].

For the state of Michigan alone, the fourth industrial revolution is projected to add \$7 billion to the automotive industry by 2022 [4]. From a global perspective, the main societal benefits are the time and cost savings and efficiency gains from specific hyper-customized services and applications while enhancing sustainability. Paramount within this transformation is that human-to-human communication will only form one pillar. The other pillars will be the revolutionary changes to other vertical industries enabled by ubiquitous communication. This book will examine the impact of 5G cellular communications on the various vertical industries.

1.2 5G and the Vertical Industries

As the fourth generation (4G) cellular system, embodied by Long-Term Evolution (LTE), is now reaching maturity, the cellular industry has developed the first standards for 5G: Third Generation Partnership Project (3GPP) Release 15 [8, 9]. 5G, however, arrives at a challenging time for the industry because the industry

² See Chapters 9 and 10.

has reached 8.8 billion global mobile connections from 5.1 billion unique subscribers [10]. This implies that almost every person who wants a mobile connection already has one. The economic impact is such that increasing revenue for mobile carriers from enhanced mobile broadband (eMBB) service will be difficult. The first wave of 5G users will not be new users but mostly users that are upgrading their services from older generations to the 5G eMBB service. It can then be argued that the success of 5G, for the industry, will depend upon more than the success of the 5G eMBB service. In other words, one must look beyond the eMBB service in order to expand the 5G footprint.

There is, however, a pot of gold at the end of the rainbow. It is expected that in the future, the number of connected things will far exceed the number of humans on Earth (7.6 billion). One research report predicted that by 2020 the number of connected devices will reach 20 billion [11]; roughly three times the number of humans on Earth. Furthermore, the growth of this number is expected to be exponential and unbounded. Consequently, connected devices (the so-called Internet of Things [IoT]) provides a growth path for the cellular industry.

This growth path, however, is not without its challenges. While the eMBB service is a monolithic service, the IoT potential includes a large number of vertical industries; each with its own service requirements. Indeed the service is so diverse that it resemble the long tailed services discussed in [12] with eMBB as the head service and the vertical services representing the long tail (Figure 1.1). The main characteristic of the long tail is that each vertical service in the long tail does not generate significant revenue but because the tail is long, there is a large number of different vertical services, the total revenue is significant; often more than the so-called “head service”. This is the classic so-called “selling less to more” scenario.

The size of the long tail has been examined by a number of studies. For example, [13] estimated that the economic impact would be a massive 2016 \$12 trillion per year by 2035 (Figure 1.2). A similar projection of \$14 trillion by 2030 was reported by [14]. To give a sense of the scale, the current global revenue of the cellular industry is about \$1 trillion per year in 2018 [10]. The total size of the tail has the potential to be 12 times that of the head in 2018. Even capturing a part of this would represent significant growth for the cellular industry.

The European Commission has an in-depth study on the growth in vertical business for the mobile 5G service providers [15]. This report focused on just a few key verticals but the trend is very clear. It showed that the revenue mix in 5G, is perhaps, the biggest difference between 5G and 4G (Figure 1.3). Currently the revenue from IoT services is a very small part and most of the revenue is from the eMBB service. By 2025, however, it is anticipated that the revenue from the vertical services would surpass that of eMBB. This would be just as transformative a milestone in the history of the commercial wireless industry as the first time that

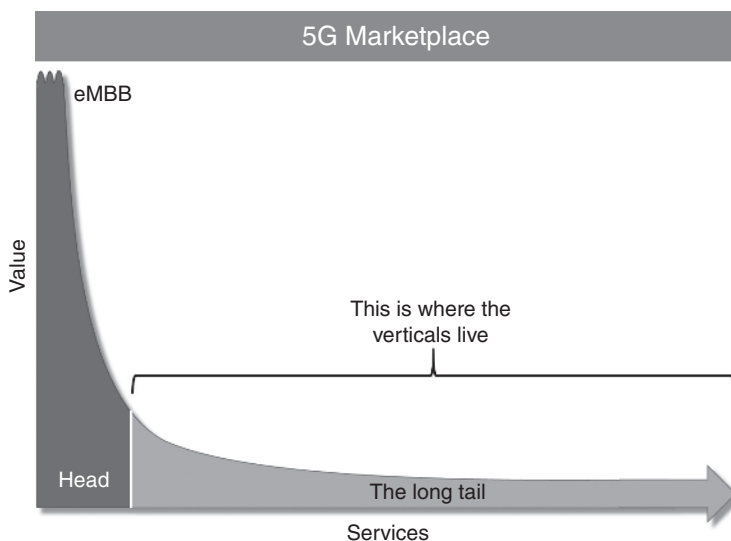


Figure 1.1 A schematic representation of the vertical services as long tailed services. The figure is modified from that in [12] to reflect the wireless industry.

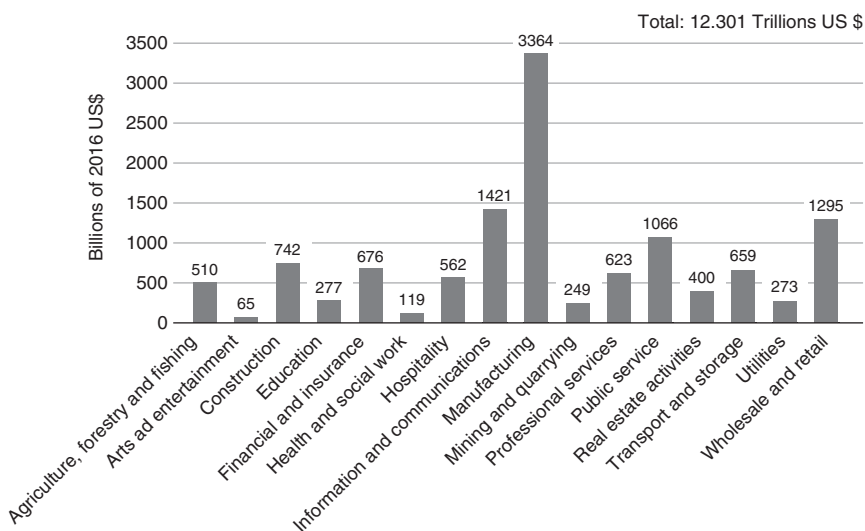


Figure 1.2 An estimate of the economic impact of different industries by 2035. Source: Data from [13].

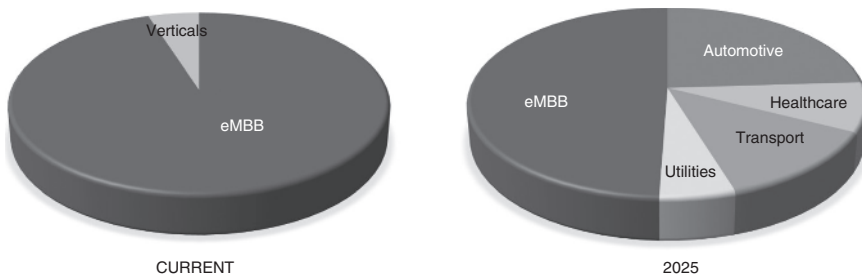


Figure 1.3 The projected operator revenue from difference services. *Source:* Projection data from [15].

the revenue from data exceeded that from voice. 5G has the capability for such a transformation because it offers more to the vertical industry than just connectivity. It offers an innovative communication platform to not only do things more efficiently but to do more and different kind of things; ranging from smart factories to smart cars to smart unmanned aerial vehicles.

From service providers' point of view, the road to 5G is clear; they have to begin to restructure their businesses around the vertical industry opportunities that 5G provides. Several tier one operators have already begun to restructure their organization to take advantage of these new opportunities. In the past, the operator might have been divided into a mobile, an enterprise, and maybe an IoT business unit. Today, we are starting to see them structure by vertical industries. This allows them to provide the needed focus on the new business initiatives built around specific vertical industries.

While carriers are at the beginning of their journey to support the vertical industry, some of the ecosystem players are further ahead [14]. Intel, Qualcomm, and Huawei are well on their way to building interoperable integrated solutions across diverse vertical sectors. Apple and Samsung are providing horizontal interoperability between devices with their integrated home IoT suites. The commercial mobile industry will provide the key backbone that supplies essential connectivity between billions of cloud applications and sensors with the biggest opportunities in providing consumer and enterprise applications and services to the verticals making the IoT a reality.

Mobile carriers are not the only ones that can see this new opportunity. Over the top (OTT) players, such as Amazon that started the whole heavy tail transformation, are also extremely excited about the potential new 5G business opportunities. They are also transforming themselves to provide new customized services for the vertical market. The carriers, if they are to stop the trend, namely OTT provides the service and the carriers just provide the pipe for the communications that is needed by the services, that 4G started, will need to find ways to out innovate

the OTT players. All is, however, not bleak for the carriers. As the rest this book will make evident, 5G is ushering in the age of ultra-customization. This means that the competitive advantage for the carriers is laid in the unique network intelligence that allows the carriers to customize their 5G service. For example, they can use machine learning techniques to customize the network slice for the application that will provide a better service experience.

Socioeconomically 5G support of the vertical industry has the potential to offer the most promise for eliminating the disadvantages associated with the digital divide [16]. The access to mobile broadband, from 4G, has significantly closed the digital divide by providing broadband access to those without fixed broadband access. The socioeconomic disadvantage, however, persisted even though access to broadband improved because the digital divide and physical divide are often in conjunction. The physical isolation of these vulnerable populations lock them out of opportunities and services even if they are provided with a broadband link. For example, they are unable to maintain regular contact with service providers which hampers their ability to monitor chronic diseases, connect with job opportunities real time, or connect with assistance for homework and research papers to aid their academic pursuits. The support of the different verticals with 5G can potentially solve all these issues by bringing the services to the disadvantaged. It can untether expert medical care from physical hospitals, expert career coaching from physical offices, and expert teachers from their academic institution, to offer customized services to their clients. The report [16] studied a number of vertical use cases and concluded that 5G vertical support, along with the hyper-customized technologies and applications that they will support, can be the great socioeconomic equalizer by providing emerging pathways for economic and social opportunities.

The concept of hyper-customization, selling less to more, has a profound impact on the requirement of the cellular systems. For one thing, it is no longer economical to build a customized system for every service. That does not mean that each vertical service does not need customization but just that such customization can no longer be at the hardware level. Fortunately, at about the same time that 5G was being developed, the phenomenon of movement of data to the cloud so that it can be easily accessed from anywhere with any device was taking shape [1]. The end points and the time frame for which network services are providing are thus fundamentally redefined. The resulting network needs to be much more nimble, flexible, and scalable. The two enabling technologies are: network function virtualization (NFV) and software defined networking (SDN). Network functions that have been traditionally tied to customized hardware appliance via NFV can now be run on a cloud computing infrastructure in the data center. It should be understood that the virtualized network function here does not necessarily have a one to one correlation with the traditional network functions but rather can be

new functions created from, e.g. the functional decomposition of traditional functions. This gives the network designer extreme flexibility on how to design the virtualized functions and interconnection. SDN provides a framework for creating intelligent programmable networks from the virtualized network functions that are, possibly, interconnected on all levels [17]. Together these technologies endow the 5G network with significant nimbleness through the creation of virtual network slices that support new types of vertical services [18] that may leverage information centric network features [19]. A more precise and detailed discussion of network slicing will be given later, for now we will just use the concept that a network slice contains all the system resources that are needed to offer a particular services so that each vertical industry can be supported in a slice. Consequently, 5G solves the economic issue by affording the network designer to create a single network that can be customized (slice) and scaled [20] for different vertical markets economically via software.

1.3 5G Requirements in Support of Vertical Industries

There are now a significant number of papers in the literature that discuss 5G systems support for vertical industries. The main characteristics of these systems are rapidly becoming open ecosystems built on top of common infrastructures [21]. In essence, they are becoming holistic environments for technical and business innovation that integrates network, computer and storage resources into one unified software programmable infrastructure. Moreover, the strict latency requirements of verticals, such as factory 2.0, is forcing the network designer to put significant network resources at the edge of the network for distributed computing. This is the so-called multi-access edge computing (MEC) phenomenon.

The 5G service and operational requirements have been detailed in [22]. Detailed requirements from the automotive, eHealth, Energy, Media, and Entertainment, and Factory of the Future were analyzed. The needs of these industries can be condensed into five use cases: dense urban information society (UC1), virtual reality office (UC2), broadband access everywhere (UC3), massive distribution of sensors (UC4), and actuators and connected car (UC5). These can then map to the following radio network requirements:

- 1) 300 Mbps per user in very dense outdoor and indoor deployments.
- 2) Broadband access (50 Mbps per user) everywhere else.
- 3) 1 Gbps per user for indoor virtual reality office.
- 4) Less than 10 ms exchange to exchange (E2E) latency everywhere.
- 5) One million devices per square kilometer.

- 6) Battery life of 10 years.
- 7) Ultra reliable communication with 100 Mbps per user with 99.999% reliability.
- 8) Ultra low latency communication with 5 ms E2E latency.

These requirements were developed into the standards [9] as eMBB, massive machine type communication (mMTC) and ultra-reliable and low-latency communications (URLLC) services. Since the use of unlicensed carriers in coordinated, on demand service-orientated fashion can offer high performance system gains [23] for certain vertical industry services, the standards also developed the support for Licensed Assisted Access.

The role of the network operator can clearly be seen as to provide a tailored communication system to its customers (end users, enterprises, and vertical industries). This requires the 5G system to have the ability to flexibly integrate the cellular communication system for various business scales ranging from multi-nationals to local micro-businesses [24]. Having to adopt to a variety of requirements, some of which may not be known during initial system design, has led to the concept of network slicing. This network flexibility will, thus, become a first key design principle in the 5G control plane. Other key design principles which allows the customer to customize, manage and even control the communication slice include: the openness of the control plane for service creation, connectivity via a multitude of access technologies (e.g. licensed and unlicensed access) and context awareness by design. The technology enablers are:³ network slicing, smart connectivity, modular architecture, resource awareness, context awareness, and control without ownership. Some of these concepts, such as control without ownership, are still under discussion in 3GPP and will not be finalized in the first release of the standards.

The aforementioned requirements are the first step in supporting the vertical industries. More detailed requirements for supporting the different verticals will be discussed in later chapters. The remainder of this chapter will look at the features in the first phase of 5G that supports the vertical industry. Moving onto future phases of 5G, 3GPP and the cellular industry has invited and offered to work with all vertical industries to define additional requirements for future 5G releases. One interesting vertical industry, for example, is professional audio production which requires strict synchronization of devices to function [25]. The requirement for low latency and synchronization goes beyond what is being developed in the current URLLC standardization. 3GPP will consider these requirements and the suite of options for vertical support in 5G in future releases of the standard will, thus, be enhanced as needed.

³ A more detailed discussions of the 5G network technologies can be found in Chapters 6 and 7.

1.4 Radio Access

The 3GPP has been working on specifying the 5G radio interface, also referred to as New Radio (NR) [26].⁴ There is always a need for mobile broadband with higher system capacity, better coverage, and higher data rates. In 5G, the needs are for more than mobile broadband. One such need is for URLLC, providing data delivery with unprecedented reliability in combination with extremely low latency, for example targeting industrial applications in factory setting.

Another is for mMTC, providing connectivity for a very large number of devices with extreme coverage, very low device cost and energy consumption. NR is being designed to support eMBB, URLLC, and mMTC. Moreover, it is possible that other use cases that are not yet known may emerge during the lifetime of NR. Therefore, NR is designed to support forward compatibility, enabling smooth introduction of future use cases within the same framework.

At present, 3GPP plans to develop the technologies and features to support eMBB, URLLC, and mMTC over multiple releases, e.g. 3GPP Releases 15, 16, and 17. The Release 15 specification covers technologies and features that are needed to support eMBB and URLLC. 3GPP has developed technologies, e.g. enhanced machine type communication and narrowband IoT that can handle mMTC use cases and can satisfy mMTC 5G requirements, and further technologies are expected to be developed to support other mMTC use cases in the later releases.

The eMBB service is to support a range of use cases including the ones identified in [27], namely (i) broadband access in dense areas, (ii) broadband access everywhere, and (iii) higher user mobility. Broadband access must be available in densely populated areas, both indoors and outdoors, such as city centers and office buildings, or public venues such as stadiums or conference centers. As the population density indoors is expected to be higher than outdoors, this needs correspondingly higher capacity. Enhanced connectivity is also needed to provide broadband access everywhere with consistent user experience. Higher user mobility capability will enable mobile broadband services in moving vehicles including cars, buses, and trains.

The URLLC is a service category designed to meet delay-sensitivity services such as industrial automation, intelligent transportation, and remote health [28]. Since the human reaction time is in the order of a millisecond (e.g. around 1 ms

⁴ A more detailed discussion of how the industry developed the 5G standards, its technology and evaluated the performance can be found in [8].

for hand touch and 10ms for visual reaction), packets for the mission-critical applications should be delivered in the order of a microsecond [28].

The NR is designed in such a way that eMBB, URLLC, and mMTC use cases are supported over a unified, flexible, and scalable frame structure. The NR interface provides a flexible framework that can be used to support different use cases. This is accomplished through a scalable Orthogonal Frequency Division Multiplexing (OFDM) based numerology and flexible frame structure. When scheduling URLLC traffic together with eMBB traffic within the same frame, since the URLLC traffic needs to be immediately transmitted due to its hard latency requirements, the URLLC transmission may overlap onto previously allocated eMBB transmissions. Various techniques to efficiently provide the resource allocation for eMBB and URLLC are given in [29, 30].

The service requirements of 5G also cause significant changes in the concept of a cellular system. For example, the concept of a cell is no longer relevant. It has evolved, in 5G, to a concept known as multi-connectivity. In principle, multi-connectivity refers to a device sharing resource of more than one base station. This concept is not really new and it has its roots in Release 12 dual connectivity. For 5G FDD/TDD dual connectivity, this allows for the transmission of multiple streams to a single UE that is semi-statically configured. The higher layer parallel transmission here is the key. It makes dual connectivity applicable to those deployment scenarios without requiring ideal, almost zero latency backhauls. 5G support LTE-NR dual connectivity which allows one of the carriers to be LTE while the other is NR.

Also within the multi-connectivity umbrella, it allows for uplink (UL)/downlink (DL) decoupling; having different cells associated with the UL and DL. The basic configuration is for one cell to be configured with two ULs and one DL. One of the UL carriers is a normal TDD or FDD UL carrier while the other is a supplementary uplink⁵ (SUL) band (Figure 1.4). This configuration allows for the dynamic scheduling and carrier switching between the normal UL and the SUL. The UE is configured with two ULs for one DL of the same cell, and uplink transmissions on those two ULs are controlled by the network. One major advantage of the SUL is for it to be in a lower frequency while the paired UL and DL carrier is in a higher frequency. This is extremely important because the UL is power restricted for health reasons. For high data rates, configuring the SUL in this configuration can give extra range to the UL to balance the UL and DL coverage. This capability is key to certain vertical use cases that are more UL intensive.

5 In 5G, an UL-only carrier frequency is referred to as the SUL frequency from a NR perspective. See [8].

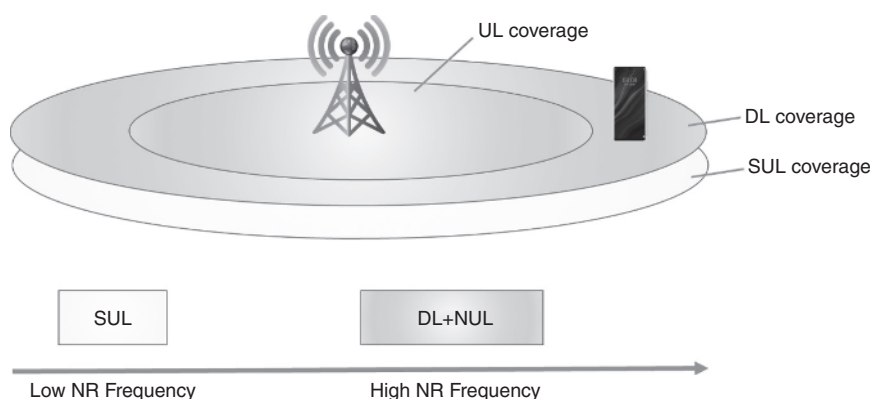


Figure 1.4 Single cell with two uplink carriers and one downlink carrier. NUL, normal uplink.

1.5 Network Slicing

It should now become clear to the reader that we cannot understate the importance of network slicing in supporting the vertical industries. On account of the recent advances in the NFV and SDN technologies [31], mobile network operators (MNOs) or even mobile virtual network operators⁶ (MVNOs) can use network slicing to create services that are customized for the vertical industries [32].

The impact of slicing in a 5G system was studied in [33] for the specific case of a vehicle network. It confirmed that slice isolation is a key requirement in three different aspects: (i) slice A should not be influenced by other slices even when the other slices are running out of resources, (ii) to prevent eavesdropping, direct communication between the slices should not be allowed, and (iii) a mechanism to prevent “through the wall hijacking” should be in place. Security isolation was the topic of investigation in [34]. Their conclusion is that a prerequisite for running a highly sensitive service in a network slice is full isolation of the slice from all other users. This needs a trusted relation between the vertical industry and the mobile operator from which the vertical is renting the resources. The operator needs to ensure to the vertical that it will guarantee the needed isolation and security for their traffic and devices. Currently in 3GPP there is a study item to look at the isolation issue. One proposal that is favored by some of the industrial 2.0 companies is for the vertical industry to take full control over the network slice. This

⁶ The difference between a MNO and a MVNO is that the MNO owns the underlying network resources. A MVNO offers mobile services like a MNO but does not own the network resources. It rents the resources from one or more MNOs.

allows for the vertical industry to dictate the isolation and security. It should be noted that this point is still begin discussed in 3GPP and the solution is currently not yet settled.

The feasibility of using NFV and SDN for slicing has been recently studied. In [35] a network slicer which allows vertical industries to define vertical services based upon a set of service blueprints and arbitrating, in the case of resource shortages, among the various vertical industries was shown to be effective for network slicing implementation. In a similar concept, a slice optimizer which communicates with an SDN controller to receive information regarding the network slice and adapts the slices according to the network state was proposed and analyzed in [36].

In parallel to the NFV and SDN advances, semantic interoperability development allows for the exchange of data between applications, as well as an increase in the level of interoperability, analytics, and intelligence [37]. This technology is now overcoming the limitation of static data models and bridges the gap between the different vertical industries' network slices.

An overview of the solutions proposed in the literature [38, 39], and the current network slicing status in 3GPP are detailed in [40]. A network slice is defined, in the standards, as a logical network that provides specific network capabilities and network characteristics. From a user's point of view, it behaves like a customized wireless system, which may be dynamically reconfigured, dedicated for its use. The standards have taken a strong step toward a cloud native approach to network slicing [41]. This is achieved through virtualization and modularization of the network function while providing a notion of network programmability through a system service-based architecture. The intension is for network function services to be flexibly used by other authorized network functions by exhibiting their functionality via a standardized service-based interface in the control plane. The question as to whether the slice extends and how far into the radio access network (RAN) [42] was discussed extensively in 3GPP. RAN slicing is particularly challenging because of the inherently shared nature of the radio channel, the desire for multi-user diversity gain, and the impact of the transmitter on the receiver [43]. The conclusion is that for the first phase of 5G, it was agreed that the RAN will only be slice-aware so as to treat the sliced traffic accordingly. The RAN will support inter-slice resource management and intra-slice quality of service differentiation. The network designer has full flexibility of how to achieve this in the system.

It is now clear that NFV and SDN allow network resources to be isolated into a programmable set of slices in order to guarantee the E2E performance of the network. Note that this guarantee is regardless of what is happening elsewhere in the network. This is known as "slice isolation". We will see later on that slice isolation is one of the reasons that makes the design of the 5G network so challenging

because one cannot just blindly implement the IT cloud concepts in the mobile network and hope to obtain an efficient network.

Efficient resource management [38] and orchestration [44], with and without mobility [45], to effectively manage the network slice to exploit its inherent flexibility will be paramount to servicing the vertical industries economically. The first challenge is for the graceful life cycle management of the virtual network functions (VNFs) and scaling of the system resources in a dynamic fashion. This is complicated by the fact that the VNFs are just pieces of software that can be instantiated anywhere in the network topology. The consequence of this is that the traditional protocol stack was designed without this flexibility of the location of the VNFs in mind. The interrelationships between the protocol stack and any instantiation has not been optimized in Release 15 which put certain constraints on the locational relationship between the VNFs. 3GPP is currently studying the issues and the optimization of the protocol stack for anywhere instantiation may be a feature in future releases. For example, there may need to be a design of a slice aware mobility management protocol to optimize the mobility challenges in network slicing [46].

The second challenge is to manage and orchestrate the resources such that the multiple network slices implemented on a common infrastructure can maintain isolation. The network would need to have resource elasticity to optimize the network sizing and resource consumption by exploiting statistical multiplexing gains [44] while maintaining isolation. Here, resource elasticity is defined as the ability of the wireless communication system to allocate and deallocate resources for each slice autonomously to scale the current available resources with the service demand gracefully.

The standards do not specify how the management and orchestration is done for network slicing. Such design is left up to the system designer. Part IV of this book will give some insights on how to achieve this in support of the vertical industry. Clearly, in order to customize the network for the service, joint resource allocation strategy that takes into account the significance of the resource to a particular service would be advantageous in supporting the diverse vertical applications on a common infrastructure [47]. Furthermore, to guarantee isolation, there is a desire to augment each virtual resource at instantiation with backup resources. This ensures that when failures occur, sufficient resources are available to maintain the service. In so doing, it becomes clear that there is a utilization tradeoff between reliability, isolation and efficient utilization of the resource. One method to circumvent this is to use pooled and shared backup resources [48]. The analysis showed that significant utilization improvement can be obtained.

One of the main challenges to slicing is an efficient method of exposing the capabilities of the network using network abstraction to the orchestration layer. This is compounded by the fact that the network is distributed and the service

may span across several domains. There are many options for the system designer to implement the service orchestrator; ranging from simple distribution to federation. In [49] an orchestrator component was proposed that monitors and allocates virtual resources to the network slices and makes use of federation with other administration domains to take decisions on the end-to-end virtual service.

The distribution of the service over many domains has a profound impact on the security of the service. In the extreme case, the distributed infrastructure that the service runs on may come from different MNOs. The MVNO leases these resources in a dynamic fashion to fulfill its service requirement. 3GPP, in the development of the standards, states that non-management prerequisites “such as trust relationships between operators, legal and business related, to create such a slice are assumed to exist” [50] between the MNOs and MVNOs for multi-operator slice creation. Consequently, the open environment which enables the trading of resources, perhaps in the form of slices, is required to facilitate the isolation and scalability of the vertical services implemented over a shared infrastructure. In [51], using Factory of the Future as the exemplary use case, it was suggested that a 5G network slice broker in a blockchain can reduce the service creation time while autonomously handle network service request. The sharing of crosshaul capabilities in a 5G network via a multi-domain exchange was analyzed in [52] for the architectures proposed in the EU H2020 project. Another proposal, [53], is to use fair weighted affinity-based scheduling heuristic to solve the scheduling of micro services across multiple clouds.

Several reports demonstrating slice management in a 5G network have been reported in the literature. An overview of the different experimental SDN/NFV control and orchestration for the ADRENALINE test bed was presented [54]. The test bed can then be used for the development and testing of end-to-end 5G vertical services. In [55], the authors proposed an end-to-end holistic operational model following a top-down approach. They planned to realize the service operational framework within the MATILDA EU H2020 project.

1.6 Other Network Issues

The previous section discussed the significance of network slicing for the support of the vertical industry services. This section will examine a number of other issues that impact the vertical services such as VNF placement, MEC, and artificial intelligence (AI).

In a 5G network, where network function are virtualized and running on distributed hardware, the network topology to instantiate the VNF can have significant impact on the efficiency of the service. This problem is further compounded by the non-uniform nature of the service demand and the irregular nature of the network

topology. This problem is not addressed in the standards because it is considered an implementation issue but is an extremely active area in the research literature. One solution [56] is to map the non-uniform distribution of signaling messages in the physical domain to a new uniform environment, a canonical domain, and then use the Schwartz–Christoffel conformal mappings to place the core functions. The analysis showed that the solution enhanced the end-to-end delay and reduced the total number of activated virtual machines. An affinity-based approach was shown in [57] to solve the function placement problem better than the greedy approach using the first-fit decreasing method. A multi-objective placement algorithm [58] only performs 5% less than the one obtained with the optimal solution for the majority of considered scenarios, with a speedup factor of up to 2000 times. The foregoing is not meant as an exhaustive review of the work on VNF placement but rather to show the vibrant research that can be used for supporting vertical services.

Another area of significant research activity is in distributed edge computing. The distribution of functions, possibly, over multiple cloud infrastructures, and the control and data traffic through these functions is known as service function chaining (SFC) [59]. An architecture for providing cost effective MEC and other services for the vertical industry was detailed and validated in [60]. An interesting problem, if there is now significant computation done at the edge, is that mobility would imply that there needs to be some mechanism to enable service migration at the edge. Network virtualization and distribution for data services over distributed enhanced packet core was discussed in [61]. Reference [62] developed the concept of a Follow Me Edge-Cloud leveraging the MEC architecture to sustain requirements of the 5G automotive systems. In principle the exact algorithm of linking the service to the cloud is up to the system designer, however, the standards provides the support to link the mobile service to the edge computing. Data plane distribution also has profound impact on the control plane side. To satisfy the low latency requirement, like the data plane, the control plane needs to be hierarchical in nature [63]. This then implies [64] that for an efficient mobile wireless system, certain data in the control plane need to be synchronized in a distributed architecture. This is particularly important because of the statefulness of the commercial wireless system.

One area of research that is getting a resurgence is using AI for network management. 3GPP has studied the usage of AI and is currently standardizing a network Data analytic function and its interface in the second phase of 5G (Release 16). The Internet Engineering Task Force (IETF) Autonomic Networking Integrated Model and Approach (ANIMA) working group is working on self-managing characteristics (configuration, protection, healing, and optimization) of distributed network elements, and adapting to unpredictable changes. The support of AI for vertical services is on its way. It should be noted that the standards will not standardize the AI methodology but rather will only standardize the interface to support AI functions. The interested reader can find details on the challenges and opportunities in [65].

1.7 Book Outline

It should now be clear that 5G is a network that was endowed, from the very beginning, with capabilities that allow for economical customization of the wireless network for different vertical industries; indeed, maybe to such a fine scale as each application. This chapter is the only chapter in Part I of this book. The rest of the book will elucidate what enables 5G to have such attributes and how it will revolutionize the communications within the vertical industries. It is organized as shown in the following.

Part II of this book is on deployments and business model opportunities and challenges. There is one chapter (Chapter 2) in this part. This chapter describes how the 5G network is designed to support a variety of vertical services from the operator point of view. The chapter introduces a variety of 5G services and their requirements. This chapter also gives a detailed presentation on the 5G network deployment architecture. Furthermore, this chapter also discusses a service-aware SON and some practical use cases. Performance benefits are also analyzed.

Part III of this book is on radio access technologies for 5G verticals. The discussions of the 5G standards here are necessarily laconic. The interested readers are encouraged to consult the more pedagogical discussion in [8]. There are three chapters (Chapters 3, 4, and 5) in this part.

Chapter 3 is on NR radio interface for 5G verticals. This chapter provides an overview of the 3GPP NR radio interface design and explains how it can be tailored to meet the requirements of different verticals. The chapter also covers advanced technologies for NR such as scalable OFDM numerology, flexible slot structure, massive multiple-input multiple-output (MIMO), beamforming, advanced channel coding, millimeter-wave deployment, spectrum aggregation, and dual connectivity and how these features enable various verticals.

Chapter 4 is dealing with one of the most important issues in millimeter-wave communications, i.e. the effects of dynamic blockage in multi-connectivity millimeter-wave radio access. This chapter provides a tutorial on modeling the dynamic blockage processes for 5G NR connectivity, and then shows how to improve the session connectivity in the presence of dynamic blockage.

Chapter 5 is on radio resource management techniques for 5G verticals. This chapter discusses radio access network resources, network slicing and its challenges for achieving efficient resource management; then it exposes the reader to the resource management approaches that can support and build efficient network slices for 5G verticals. Furthermore, resource allocation techniques and performance analysis are shown for virtual reality use case.

Part IV of this book is on network infrastructure technologies for 5G verticals. There are two chapters (Chapters 6 and 7) in this part.

Chapter 6 is on advanced NFV, SDN and mobile edge technologies for URLLC verticals. This chapter gives an overview of several URLLC vertical scenarios, their requirements and different deployment scenarios; it then discusses SDN, NFV and

5G core network functions to provide high precision networking to match the bandwidth, latency, and reliability targets of different URLLC applications.

Chapter 7 is on edge clouds complementing 5G networks for real time applications. This chapter describes the basics of edge cloud with respect to 5G networking and deployment, and SDN, NFV, and the need for disaggregation of control from data planes to provide the best practices at the edge where control from the cloud is collapsing with the central RAN control at the edge. The chapter also provides the state-of-the-art for edge cloud deployment options.

Part V of this book is on 5G key vertical applications. There are three chapters (Chapters 8, 9, and 10) in this part.

Chapter 8 is on 5G connected aerial vehicles. This chapter summarizes the needs and challenges to support aerial vehicles over current and future cellular networks, and provides a review of the research and development of drone communications in general from both academia and industry. This chapter also gives a discussion on the 5G challenges for aerial vehicle solutions and further work needed in this area.

Chapter 9 is on 5G connected automobiles. This chapter focuses on the value 5G brings to connected vehicles, specifically on high data rates, low latency, and edge computing features. This chapter concentrates on a couple of main use cases, vehicle platooning, and high definition maps, and how 5G and edge computing can assist to provide the best experience to the end-user.

Chapter 10 is on 5G for the industrial application. It will study the capabilities of the current release of 5G and its applicability to the smart factory use cases. The gaps to the requirement will be identified and how the industry is working to find solutions to close the gaps will be elucidated. The chapter concludes with a discussion on the spectrum situation of industrial use as well as some early trials and demonstrations.

References

- 1 Andrews, J.G., Buzzi, S., Choi, W. et al. (2014). What will 5G be? *IEEE Journal on Selected Areas in Communication* 32 (6): 1065–1081.
- 2 Liu, J., Xiao, W., and Soong, A.C.K. (2016). Dense network of small cells. In: *Design and Deployment of Small Cell Networks* (eds. A. Anpalagan, M. Bennis, and R. Vannithamby), 96–120. Cambridge: Cambridge University Press.
- 3 Schwab, K. (2016). *The Fourth Industrial Revolution*. Geneva: World Economic Forum.
- 4 World Economic Forum (2019). A New Era of Manufacturing in The Fourth Industrial Revolution: \$7 Billion of Possibilities Uncovered in Michigan. <https://www.weforum.org/whitepapers/a-new-era-of-manufacturing-in-the-fourth-industrial-revolution-7-billion-of-possibilities-uncovered-in-michigan>.
- 5 General Motors (2019). U.N. Sustainable Development Goals. <https://www.gmsustainability.com/unsdg.html> (accessed 29 April 2019).

- 6 Ford (2018). Contributing to the UN SDGs. <https://corporate.ford.com/microsites/sustainability-report-2017-18/strategy-governance/sdg.html> (accessed 29 April 2019).
- 7 United Nations (2015). Transforming our World: The 2030 Agenda for Sustainable Development. <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication>.
- 8 Lei, W., Soong, A.C.K., Jianghua, L. et al. (2020). *5G System Design: An End to End Perspective*. Cham: Springer International Publishing.
- 9 3GPP (2018). 3GPP Specification Series 38. <http://www.3gpp.org/DynaReport/38-series.htm>.
- 10 Global Data (n.d.). GSMA Intelligence. <https://www.gsmainelligence.com> (accessed 20 August 2018).
- 11 Hung, M. (2017). *Leading the IoT*. Gartner Research.
- 12 Anderson, C. (2006). *The Long Tail*. Westport: Hyperion.
- 13 Campbell, K., Diffley, J., Flanagan, B. et al. (2017). The 5G economy: How 5G Technology will contribute to the global economy. <https://cdn.ihs.com/www/pdf/IHS-Technology-5G-Economic-Impact-Study.pdf>.
- 14 World Economic Forum (2017). Digital Transformation Initiative: Telecommunication Industry. <http://reports.weforum.org/digital-transformation/wp-content/blogs.dir/94/mp/files/pages/files/dti-telecommunications-industry-white-paper.pdf>.
- 15 Directorate-General for Communications Networks, Content and Technology (European Commission) (2017). Identification and quantification of key socio-economic data to support strategic planning for the introduction of 5G in Europe. <https://publications.europa.eu/en/publication-detail/-/publication/ee832bba-ed02-11e6-ad7c-01aa75ed71a1>.
- 16 Lee, N.T. (2019). Enabling opportunities: 5G, the internet of things, and communities of color. <https://www.brookings.edu/research/enabling-opportunities-5g-the-internet-of-things-and-communities-of-color>.
- 17 Tudzarov, A. and Gelev, S. (2018). 5G and software network paradigm. 23rd International Scientific-Professional Conference on Information Technology (IT).
- 18 Chiosi, M., Clarke, D., Willis, P. et al. (2012). Network Functions Virtualisation – An Introduction, Benefits, Enablers, Challenges and Call for Action. http://portal.etsi.org/NFV/NFV_White_Paper.pdf.
- 19 Ravindran, R., Chakraborti, A., Amin, S.O. et al. (2017). 5G-ICN: delivering ICN services over 5G using network slicing. *IEEE Communications Magazine* 55 (5): 101–107.
- 20 Rost, P., Mannweiler, C., Michalopoulos, D.S. et al. (2017). Network slicing to enable scalability and flexibility in 5G mobile networks. *IEEE Communications Magazine* 55 (5): 72–79.
- 21 Gavras, A., Denazis, S., Tranoris, C. et al. (2017). Requirements and design of 5G experimental environments for vertical industry innovation. 2017 Global Wireless Summit (GWS).

- 22 Elayoubi, S.E., Fallgren, M., Spapis, P. et al. (2016). 5G service requirements and operational use cases: Analysis and METIS II vision. 2016 European Conference on Networks and Communications (EuCNC).
- 23 Pateromichelakis, E., Bulakci, O., Peng, C. et al. (2018). LAA as a key enabler in slice-aware 5G RAN: challenges and opportunities. *IEEE Communications Standards Magazine* 2 (1): 29–35.
- 24 Mahmood, K., Mahmoodi, T., Trivisonno, R. et al. (2017). On the integration of verticals through 5G control plane. 2017 European Conference on Networks and Communications (EuCNC).
- 25 Pilz, J., Holfeld, B., Schmidt, A., and Septinus, K. (2018). Professional live audio production: a highly synchronized use case for 5G URLLC systems. *IEEE Network* 32 (2): 85–91.
- 26 3GPP (2018). TR 38.913 Study on scenarios and requirements for next generation access technologies v15.0.0. <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2996>.
- 27 NGMN (2015). NGMN 5G White Paper. <https://www.ngmn.org/5g-white-paper/5g-white-paper.html>.
- 28 Ji, H., Park, S., Yeo, J. et al. (2018). Ultra-reliable and low-latency communications in 5G downlink: physical layer aspects. *IEEE Wireless Communications* 25 (3): 124–130.
- 29 Simsek, M., Aijaz, A., Dohler, M. et al. (2016). 5G-enabled tactile internet. *IEEE Journal on Selected Areas in Communications* 34 (3): 460–473.
- 30 Alsenwi, M., Tran, N.H., Bennis, M. et al. (2019). eMBB-URLLC resource slicing: a risk-sensitive approach. *IEEE Communications Letters* 23 (4): 740–743.
- 31 Ordonez-Lucena, J., Ameigeiras, P., Lopez, D. et al. (2017). Network slicing for 5G with SDN/NFV: concepts, architectures and challenges. *IEEE Communications Magazine* 55 (5): 80–87.
- 32 Kalyoncu, F., Zeydan, E., and Yigit, I.O. (2018). A data analysis methodology for obtaining network slices towards 5G cellular networks. IEEE 87th Vehicular Technology Conference (VTC Spring).
- 33 Soenen, T., Banerjee, R., Tavernier, W. et al. (2017). Demystifying network slicing: From theory to practice. 2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM).
- 34 Schneider, P., Mannweiler, C., and Kerboeuf, S. (2018). Providing strong 5G mobile network slice isolation for highly sensitive third-party services. IEEE Wireless Communications and Networking Conference (WCNC).
- 35 Casetti, C., Chiasserini, C.F., Deiß, T. et al. (2018). Network slices for vertical industries. IEEE Wireless Communications and Networking Conference Workshops (WCNCW).
- 36 Rezende, P.H.A. and Madeira, E.R.M. (2018). An adaptive network slicing for LTE radio access networks. 2018 Wireless Days (WD).
- 37 Osseiran, A., Elloumi, O., Song, J., and Monserrat, J.F. (2018). Internet-of-Things (IoT). *IEEE Communications Standards Magazine* 2 (2): 70.

- 38 Richart, M., Baliosian, J., Serrat, J., and Gorricho, J.-L. (2016). Resource slicing in virtual wireless networks: a survey. *IEEE Transactions on Network and Service Management* 13 (3): 462–476.
- 39 Foukas, X., Patounas, G., Elmokashfi, A., and Marina, M.K. (2017). Network slicing in 5G systems. *IEEE Communications Magazine* 55 (5): 94–100.
- 40 Kaloxylos, A. (2018). A survey and an analysis of network slicing in 5G networks. *IEEE Communications Standards Magazine* 2 (1): 60–65.
- 41 Sharma, S., Miller, R., and Francini, A. (2017). A cloud-native approach to 5G network slicing. *IEEE Communications Magazine* 55 (8): 120–127.
- 42 Ksentini, A. and Nikaein, N. (2017). Toward enforcing network slicing on RAN: flexibility and resource abstraction. *IEEE Communications Magazine* 55 (6): 102–108.
- 43 Sallent, O., Perez-Romero, J., Ferrus, R., and Agusti, R. (2017). On radio access network slicing from a radio resource management perspective. *IEEE Wireless Communications* 24 (5): 166–174.
- 44 Gutierrez-Estevez, D.M., Gramaglia, M., De Domenico, A. et al. (2018). The path towards resource elasticity for 5G network architecture. *IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*.
- 45 Zhang, H., Liu, N., Chu, X. et al. (2017). Network slicing based 5G and future Mobile networks: mobility, resource management, and challenges. *IEEE Communications Magazine* 55 (8): 138–145.
- 46 Li, X., Samaka, M., Chan, H.A. et al. (2017). Network slicing for 5G: challenges and opportunities. *IEEE Internet Computing* 21 (5): 20–27.
- 47 Narmanlioglu, O., Zeydan, E., and Arslan, S.S. (2018). Service-aware multi-resource allocation in software-defined next generation cellular networks. *IEEE Access* 6: 20348–20363.
- 48 Yeow, W.-L., Westphal, C., and Kozat, U.C. (2011). Designing and embedding reliable virtual infrastructures. *ACM SIGCOMM Computer Communication Review* 41 (2): 57–64.
- 49 Li, X., Mangues-Bafalluy, J., Pascual, I. et al. (2018). Service orchestration and federation for verticals. *IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*.
- 50 3GPP (2018). TR 28.801 v15.1.0: Study on management and orchestration of network slicing for next generation networks. http://www.3gpp.org/ftp//Specs/archive/28_series/28.801.
- 51 Backman, J., Yrjölä, S., Valtanen, K., and Mämmelä, O. (2018). Blockchain network slice broker in 5G: Slice leasing in factory of the future use case. *Internet of Things Business Models, Users, and Networks*, Copenhagen, Denmark (23–24 November 2017), IEEE.
- 52 Contreras, L.M., Bernardos, C.J., Oliva, A.D.L., and Costa-Perez, X. (2017). Sharing of crosshaul networks via a multi-domain exchange environment for 5G services. *IEEE Conference on Network Softwarization (NetSoft)*.

- 53 Bhamare, D., Samaka, M., Erbad, A. et al. (2017). Multi-objective scheduling of micro-services for optimal service function chains. *IEEE International Conference on Communications (ICC)*.
- 54 Muñoz, R., Vilalta, R., Casellas, R. et al. (2017). Integrating optical transport network testbeds and cloud platforms to enable end-to-end 5G and IoT services. *19th International Conference on Transparent Optical Networks (ICTON)*.
- 55 Gouvas, P., Zafeiropoulos, A., Vassilakis, C. et al. (2017). Design, development and orchestration of 5G-ready applications over sliced programmable infrastructure. *29th International Teletraffic Congress (ITC 29)*.
- 56 Laghrissi, A., Taleb, T., and Bagaa, M. (2018). Conformal mapping for optimal network slice planning based on canonical domains. *IEEE Journal on Selected Areas in Communications* 36 (3): 519–528.
- 57 Bhamare, D., Samaka, M., Erbad, A. et al. (2017). Optimal virtual network function placement in multi-cloud service function chaining architecture. *Computer Communications* 102: 1–16.
- 58 Arouk, O., Nikaein, N., and Turletti, T. (2017). Multi-objective placement of virtual network function chains in 5G. *IEEE 6th International Conference on Cloud Networking (CloudNet)*.
- 59 Halpern, J. and Pignataro, C. (eds.) (2015). RFC 7665: Service Function Chaining (SFC) Architecture. <https://www.rfc-editor.org/info/rfc7665>.
- 60 Vilalta, R., Lopez, V., Giorgetti, A. et al. (2017). TelcoFog: a unified flexible fog and cloud computing architecture for 5G networks. *IEEE Communications Magazine* 55 (8): 36–43.
- 61 Kaippallimalil, J. and Chan, H.A. (2014). Network virtualization and direct Ethernet transport for packet data network connections in 5G wireless. *IEEE Global Communications Conference*.
- 62 Aissioui, A., Ksentini, A., Gueroui, A.M., and Taleb, T. (2018). On enabling 5G automotive systems using follow me edge-cloud concept. *IEEE Transactions on Vehicular Technology* 67 (6): 5302–5316.
- 63 Sama, M.R., Contreras, L.M., Kaippallimalil, J. et al. (2015). Software-defined control of the virtualized mobile packet core. *IEEE Communications Magazine* 53 (2): 107–115.
- 64 Kaippallimalil, J., Mademann, F., Shiyong, T. et al. (2015). Data distribution and synchronization in next generation mobile core network. *IEEE Conference on Standards for Communications and Networking (CSCN)*.
- 65 Mwanje, S., Decarreau, G., Mannweiler, C. et al. (2016). Network management automation in 5G: Challenges and opportunities. *IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*.