A Survey on the 5th Generation of Mobile Communications: Scope, Technologies and Challenges

M. Sousa\textsuperscript{a,cd}, P. Vieira\textsuperscript{bd}, M. P. Queluz\textsuperscript{ad}, A. Rodrigues\textsuperscript{a,cd}

\textsuperscript{a}Departamento de Engenharia Electrotécnica e de Computadores, Instituto Superior Técnico, Portugal
\textsuperscript{b}Departamento de Engenharia Electrónica e Telecomunicações e de Computadores, Instituto Superior de Engenharia de Lisboa, Portugal
\textsuperscript{c}CELFINET, Consultoria em Telecomunicações, Lda.
\textsuperscript{d}Instituto de Telecomunicações, Lisboa, Portugal

marco.sousa@celfinet.com pvieira@deetc.isel.pt [paula.queluz, ar]@lx.it.pt

Abstract—The 5th Generation (5G) of mobile communications will impact the customers Quality of Experience (QoE) by addressing the current mobile networks usage trends and providing the technological foundation for new and emerging services. Additionally, 5G may provide a unified mobile communication platform, with multiple purposes, leveraging industries, services and economic sectors. In this paper, a 5G tutorial is presented, including the 5G drivers, main use cases, vertical markets and a current status of the standardization process. Furthermore, several 5G key enabling technologies are presented, concerning the Radio Access Network (RAN) and Core Network (CN) perspectives. Finally, a brief outline over the Internet of Things (IoT) concept and current research topics is presented.

Keywords: 5G, eMBB, mMTC, URLLC, SDN, NFV, mMIMO, mWaves, MEC.

I. INTRODUCTION

Mobile wireless communication networks have been experiencing enormous advances throughout its successive generations. Starting with 1st Generation (1G), which was an analog technology only for voice calls in the 1980s, industry moved on to 2nd Generation (2G), that entered into the digital domain in the 1990s; besides voice, it also supported Short Message Services (SMSs). The breakthrough of the 3rd Generation (3G) was the global access to mobile data services, video streaming, web browsing, e-mail, etc, in the 2000s. The introduction of the 4th Generation (4G) eliminated the circuit switching domain to embrace an all Internet Protocol (IP) network, enhancing data services, in the 2010s. Finally, the 5th Generation (5G) will be disruptive in the sense that while previous generations had the purpose of connecting people, 5G will connect not only people but also the physical world (things). Also, the 5G will provide the technological foundation for a wide range of new applications and services, headed for an increasingly connected world.

In the economy field, the concept of General Purpose Technologies (GPT), identifies technologies whose adoption introduce changes that redefine work processes as well as the rules of competitive economic advantages [1]; the printed press, the Internet, and the computer, are a few winning examples. According to [1] from IHS Markit, the 5G has the potential to enter the exclusive group of GPTs, as a technological breakthrough.

In this paper, a state of the art about the 5G is presented. As mentioned, 5G is significantly different from previous generations and will have a significant impact in society and economy. In that sense, besides the 5G technological foundations, also the new services/applications and economic sectors, which are likely to adopt or benefit from the 5G, will be considered.

This paper is organized as follows. In Section II, the development scope of the 5G is presented; it includes the technology drivers, the main 5G use cases and the regulatory perspective of 5G. In Section III, an overview of possible technologies to facilitate the 5G networks, is presented; it includes key technologies for both Radio Access Network (RAN) and Core Network (CN). Section IV, is dedicated to the Internet of Things (IoT), in light of being one key difference between 5G and legacy networks. Finally, in Section V conclusions are draw.

II. 5G SCOPE

This section provides a wide scope analysis of 5G networks, starting with the main drivers for its development and foreseen use cases. Since the Mobile Network Operators (MNOs) business approach will be also transformed along with 5G deployment, new vertical players are anticipated, such as the automotive industry, which has shown a strong interest in taking advantage of the upcoming networks. The efforts with standardization and regulation, from entities such as the 3rd Generation Partnership Project (3GPP) and International Telecommunication Union (ITU), are crucial for the first 5G commercial deployments, and will be also overviewed in this section.

A. 5G DRIVERS

The 5G networks drivers can be loosely classified as customer or industry associated. With respect to mobile network costumers, there are several key trends. Firstly, the number of mobile subscriptions is growing at 6% annual rate [2], reaching 7.8 billion in the third quarter of 2017 (Q3), and the respective grow rate for mobile broadband subscriptions is around 20% year-on-year [2]. Besides the mobile subscriptions up rise trend, the mobile data traffic generated by each subscriber is also increasing. Latest statistics account 65% annual increase
in mobile traffic between the 3rd quarter of 2016 and 2017 [2], and the momentum is expected to continue and reach a 8-fold increase in 2023. Currently, more than 50% of the mobile traffic is video streaming [2], being the leading application in traffic generation.

The 4G networks are not fully developed yet, for instance Self-Organizing Networks (SON) features, as covered in [3] [4] can improve network performance, also MNOs are starting to evolve their Long Term Evolution (LTE) networks to Long Term Evolution - Advanced (LTE-A), enhancing the 4G network performance. Nonetheless, even considering such solutions, 4G networks will not be able to cope with the subscribers increasing demand. Furthermore, considering services such as video streaming, where the uprising trend in video resolution will generate more content in 4K or even 8K, or considering the appearance of new services, like virtual reality and augmented reality, its feasibility with the current mobile networks is limited. Overall, such services require high throughputs, and constitute the first main 5G use case, the Enhanced Mobile Broadband (eMBB) [5].

On another perspective, the fourth industrial revolution known as “Industry 4.0” is starting. According to [6], Industry 4.0 is defined as the digitalization process of the manufacturing sector; it consists on embedding sensors for monitoring all products and equipment, using cyber physical systems, and applying cognitive analysis to all collected data. From the technological point of view, 5G will be an important facilitator in the full realization of the Industry 4.0 concept. Accordingly, the second main use case, defined by ITU, is the Massive Machine Type Communications (mMTC) [5]. In essence, it is characterized by a large number of connected devices, collecting and sending non-delay-sensitive data [5]; it can also be seen as one of the applications of IoT [7].

Lastly, ITU also defined the Ultra-Reliable Low Latency Communications (URLLC) [5] requirements. This will establish the technological foundations for the self-driving [8] endeavor. There is a strong economic interest from the automotive industry in developing not only the autonomous vehicles, but also a connected and intelligent infrastructure. It requires both Vehicle-to-Vehicle (V2V) communications and Vehicle-to-Infrastructure (V2I). Along this way, the transmitted data is as time sensitive as intolerant to failures and errors. Satisfying the above mentioned requisites entails in the development of wireless communication systems with high availability, ultra reliable and with very low latency. Similarly to the automotive industry, other sectors may also develop new services based on URLLC, such as Wireless Tele Surgery (WTS) [9] in the health care sector.

B. Use Cases

Figure 1 presents the main 5G applications in a three dimensional space, where each dimension corresponds to one of the main 5G requirements (eMBB, mMTC and URLLC).

These requirements are deeply analyzed in the following subsections.

1) Enhanced Mobile Broadband: The eMBB essentially deals with the human centric use cases, aiming to deliver multimedia content and data services. It depends on the ability to provide high throughputs, especially when considering high quality video streaming or 3D video services. In a recent survey [10], the authors identified that network throughput and security are the factors that mostly condition the subscribers expectations about the upcoming 5G networks which is aligned with the eMBB requirements. In [11], the authors classified eMBB as the first 5G “killer app”, based on a costumer survey where the experienced bottlenecks, of smartphone users, were identified as one of the factors that most impact the Quality of Experience (QoE).

ITU has defined minimum requirements for the main network capabilities [12]. Nonetheless, depending on the usage scenario, some capabilities are more relevant than others. This relative importance is depicted in Figure 2.

Considering the eMBB, the key capabilities are: network energy efficiency, area traffic capacity, peak data rate, user experienced data rate, spectrum efficiency and mobility. In [11], and considering the RAN, the authors identified three pillars to address the requirements of eMBB, densify networks, deliver higher spectral efficiency and the usage of new spectrum bands. In this matter, the development and standardization of the 5G New Radio (NR), that is an ongoing process, will provide the technology to deliver the performance require-
ments for eMBB. The main key aspects of the 5G NR are the shared access scheme, Orthogonal Frequency Division Multiplexing (OFDM) based or Non-Orthogonal Multiple Access (NOMA), new system architectures enabling Cloud Radio Access Network (C-RAN) and network slicing, new spatial-domain processing techniques including Massive Multiple-Input Multiple-Output (mMIMO) and 3D beamforming, and also the use of Millimeter Wave (mmWave) bands.

2) Massive Machine Type Communications: The services belonging to the mMTC are such that machine-type devices are used for monitoring, sensing and metering. On a 5G perspective, the network is expected to support a large number of these machine-type devices, typically requiring low throughput and sparse communications. Also, it is required to support a high connection density of these devices [13]. Basically, 5G mMTC encompasses the IoT use cases which are non delay-sensitive, using a standardized technological solution. The 3GPP group has already standardized the Narrowband Internet of Things (NB-IOT), in LTE Release 13 and LTE Release 14, to provide wide-area mMTC connectivity for IoT [14]. Besides this licensed band standard, other proprietary solutions, such as Sigfox and Lora, were developed in unlicensed bands [15]. Nonetheless, these solutions are struggling in a scenario where the number of devices significantly overpass the resources. As these solutions rely on orthogonal transmission principles, in recent years non-orthogonal strategies have been also proposed to accommodate more users than the more classical orthogonal approaches [13].

The 5G eMBB and the mMTC require two completely different communication system designs. Not only the eMBB service category is heavily focused on the downlink communication, while the mMTC service is focused on the uplink, but also the eMBB requires high packet sizes and throughputs, whereas the mMTC requires low values of these parameters [16]. Consequently, the mMTC advocates a set of technologies that are quite different from those of the eMBB use case. The first main difference is, as mentioned before, the medium access scheme, in order to support the highest number of connected devices. While the orthogonal medium access tightly sets the available resources according to the number of supported users, a non-orthogonal medium access allows some degree of resources overloading [16]. Moreover, the grant-based access, used in LTE, requires a good prediction of uplink requests and additional signaling, which is not ideal for the mMTC scenario. In that sense, a grant-free solution is expected to enhance the requirements feasibility of mMTC [16]. Both solutions imply a complexity increase from the base station in order to simplify the devices complexity and achieve the stipulated requirements. Technologies such as Sparse Code Multiple Access (SCMA), Compressed Sensing based Multi User Detection (CS-MUD) and Continuous Phase Modulation (CPM) can be strong candidates to enable massive access.

3) Ultra-reliable and Low Latency Communications: The URLLC supports the applications that are latency sensitive as remote control and autonomous driving [17]. Another emerging application is tactile internet [18]; it allows humans to control real and virtual objects wirelessly, and hence, equivalent to human touch, visual and auditory perceptions would be transmitted seamlessly via data networks. This application requires 1 ms end-to-end latency [19]. Other services for 5G URLLC require latencies in the range of 1-10 ms. Currently, LTE networks are characterized by latencies in the range of 30 to 100 ms [13]. This end-to-end latency values are due to the best-effort policies, typically used in the backbone network. In order to comply with the latency requirements of the 5G URLLC, changes have to be conducted not only in the backbone network but also in the RAN. The use of Software Defined Networks (SDNs), Network Function Virtualization (NFV), and the concept of network slicing, can establish dedicated connections for URLLC services [17]. Likewise, the use of Multi-Access Edge Computing (MEC) can reduce the latency even more [20]. On the RAN, given that a large portion of the latency is introduced by the control signaling [13], the communication overhead for URLLC has to be reduced. To accomplish this, the packet and frame structure, and the scheduling schemes, have to be revisited [17]. The use of polar codes for large-sized packet and of Sparse Vector Coding (SVC) for small-sized ones are technological enablers to achieve the latency requirements.

C. New Players

Even though the number of mobile subscribers has been rising, the MNOs revenues have been flutting out in the past years [21]. With the anticipated capabilities of 5G, MNOs can achieve a much larger growth, addressing key challenges in digitalization of manufacturing, automotive and other industries, also known as vertical markets. In this ecosystem, MNOs can become, besides network developers, service enablers or even service creators [21]. In this new reality, the role of MNOs, might evolve from Business-to-Consumer (B2C) to Business-to-Business (B2B) providers. In the following text, an overview of main vertical markets, in a 5G perspective, is presented.

1) Automotive: The automotive industries have been adopting several connectivity technologies pursuing the long term goal of autonomous driving. In that sense, the automotive industry is clearly interested in the possibilities that 5G might enable. The 5G can improve the Vehicle-to-Everything (V2X) communications [22] and also the in car “infotainment” [23]. Other new use cases such as Platooning [24] or teleoperated vehicles [25] might be developed with the support of 5G networks.

2) Manufacturing: Within the manufacturing industry, and as pursued by the Industry 4.0, 5G networks will be able to provide the underlying unified communication platform to fulfill the digital transformation [26]. The 5G can be a key enabler for remote assistance and robot control, logistics tracking or process automation within factories [23]. Throughout, efficiency improvements and automation, all over the supply chain and product’s life cycle, are expected.

3) Agriculture: Even more traditional sectors, such as agriculture, are undertaking profound technological innovations. With IoT, agriculture can move to smart farming and precision agriculture [23]. The baseline is to monitor crop yields, moisture levels and/or terrain, providing more data to support
assertive farming decisions. The 5G can provide a robust network for IoT and also for remote control of the farming machinery [23].

4) Energy and Utilities: Regarding the energy and utilities environment, the energy industry is facing some challenges; the increased electricity consumption associated with some degree of uncertainty in modeling the demand, and considering several sources of energy generation (renewable sources), contribute with additional threats. In this complex scenario, emerging concepts as smart metering and smart grids [27], have been investigated. The smart metering can be considered part of the mMTC applications whereas the smart grid concept might be more challenging. The smart grid concept requires communication between sensors, control systems, energy generation and storage to monitor and to optimize the grid in real-time. The 5G can support this use case, specially where fiber based access is not cost effective [28].

D. Regulation

The regularization and the full standardization of 5G is a lengthy process, involving several entities and organizations, private and public ones. Around 2014, the International Telecommunication Union Radiocommunication Sector (ITU-R) started working in the definition towards the 5G technology performance requirements, as outline in Figure 3.

Fig. 3: Timeline for IMT-2020 in ITU-R [29].

Within ITU-R, the 5G is known as International Mobile Telecommunications - 2020 (IMT-2020). In 2016, ITU-R produced the first draft of the IMT-2020 performance requirements. The submitted proposals, for the new radio interfaces, to be included in IMT-2020 specification, will be evaluated by independent external groups. The whole process is expected to be due by 2020, with the approval of the specifications [29].

Meanwhile, the 3GPP is developing its proposal for the IMT-2020 call from ITU-R. The final specifications for the IMT-2020 are planned to be completed with Release 16, by the end of 2019, as shown in Figure 4.

The 3GPP group has released an early drop of Release 15 containing the Non-Standalone (NSA) 5G radio specifications. The full Release 15 will include the standalone version of the 5G NR by mid 2018. While the Release 15 focus on the eMBB, Release 16 is expected to evolve more in depth the mMTC and the URLLC specifications.

Fig. 4: 3GPP Timeline towards 5G [30].

III. TECHNOLOGIES

In the previous section, the 5G environment and the three major use cases were considered. Several technologies were identified as key enablers to meet the 5G requirements. In this section, these technologies are presented considering RAN and CN domains, respectively. Moreover, the different network architectures that are being developed within the 3GPP group, are also considered.

A. Architecture

Regarding the upcoming 5G network architecture, there are two distinct architecture approaches: the NSA 5G architecture and the standalone 5G architecture. The 3GPP group, in a early drop version of Release 15, focused their effort mainly towards the specification of the NSA 5G network. The complete Release 15 should include also the standalone 5G architecture specification. The left side of Figure 5 presents the NSA architecture, that should be used on the first commercial deployments of 5G; the right side of the Figure 5 displays the standalone option for 5G.

Fig. 5: Proposed 5G architectures in 3GPP Release 15 [31].

According to the specifications, the NSA 5G architecture, should have a 5G Base Station (BS) anchored in a LTE network. A User Equipment (UE) is expected to be connected to the LTE BS, using this link for user and control plane, and while connected with a 5G BS for user plane only. Also, the 5G BS is connected with the LTE Evolved Packet Core (EPC). Considering the standalone option, the 5G RAN is connected...
only with the 5G Core Network (5G-CN) providing both user and control planes, forming a full 5G network.

B. Radio Access Network

Taking into account the 5G RAN, there are several expected evolutions compared with current networks, that are described next.

1) Massive MIMO: Traditional Multiple-Input Multiple-Output (MIMO) systems, improve the Signal-to-Noise Ratio (SNR) in uplink due to the diversity gain, mitigating the fading effects and increasing capacity due to the achieved spatial multiplexing gain [32]. There are two major MIMO types: Single-User MIMO (SU-MIMO) and Multi-User MIMO (MU-MIMO). Whereas SU-MIMO allocates the time-frequency resources to a single user, MU-MIMO simultaneously serves several users in the same time-frequency resource using beamforming [33]. Finally, mMIMO is a form of MU-MIMO [33], where the number of BS antennas is much larger [34] than in MU-MIMO. The MU-MIMO requires a highly reliable Channel State Information (CSI), which is unlikely and complex in practical scenarios [32]. Additionally, in [35] the authors showed that when considering a large number of antennas at the BS MU-MIMO allows the simplest sort of precoding on the forward link and processing on the reverse link. Thus, mMIMO deliver high throughputs (high spectral efficiency) reliably, both in the uplink and downlink channels, and in fast-changing propagation environments, without the effects of uncorrelated noise and fast fading. It only remains the inter-cellular interference due to pilot contamination [35]. Even though, work is being developed to mitigate the pilot contamination using: protocol based methods [36], blind methods [37], precoding [38], Genetic Algorithm (GA) based allocation scheme [39] or the use of dual pilot sequences [40]. Hence, mMIMO will be a key enabler for the 5G use cases. Moreover, combining the high antenna gains (mMIMO) with large available bandwidths (mmWave), 5G network throughput barriers can be pushed even further.

2) Millimeter Waves: The limited available spectrum, and its cost, have always been a concern for MNOs. In that sense, efforts such as spectrum refarming, through the re-purposing of the terrestrial TV spectrum [34], or by refarming the 2G bands to other technologies, have been implemented. Also, spectrum sharing techniques [41] [42] have been proposed. Nonetheless, some open issues remain to be solved [34], as how to increase the available bandwidth at microwave frequencies. Alternatively, there is a great available bandwidth at the mmWave, ranging from 3 to 300 MHz [34]. Due to the associated high pathloss for higher frequencies, the use of mmWave towards 5G networks was initially considered for short-range indoor locations [43]. More recently, it has been tested in outdoor Non-Line-of-Sight (NLoS) environments, with satisfactory results [44] [45]. Moreover, combining mmWave with other techniques, such as beamforming, can enhance the outdoor mmWave performance [46].

Besides some regulatory constraints, key technical constraints prevent the use of mmWaves. The high pathloss, the assumption of needed Line-of-Sight (LoS), the infeasible nature of some hardware components due to technical limitations and the high Doppler shift, remain as key challenges [47]. More recently, all these challenges have been subject of important developments.

Regarding the high pathloss, it can be mitigated with the use of smart antennas, with high gains [45] [48]. The LoS requirement can be solved by exploiting the multipath reflections [45]. Considering the hardware limitations, new antenna array designs have been proposed [49], but there are still some challenges, as the linear relationship between power consumption and the sampling rate, in the analog to digital converters, which is even more critical at millimeter bands [49]. Finally, the Doppler shift is also mitigated when considering directional antennas [48].

Aside from the mmWave applications for the RAN, backhaul and relay solutions [50] are also available.

3) Others: Some other 5G RAN technologies, not included in the previous items, can be considered. The RAN random access scheme proposed in the 3GPP Release 15, is OFDM based. Nonetheless, new waveform candidates are expected in future releases. For instance, non-orthogonal waveform enables higher spectral-efficiency compared with the orthogonal ones [51]. Some of the non-orthogonal access schemes are the NOMA [52], the SCMA [53], the Multi-user Shared Access (MUSA) [54], the Pattern Division Multiple Access (PDMA) and Successive Interference Cancellation Aenable Multiple Access (SAMMA) [55]. All these schemes achieve higher throughputs, or spectral-efficiencies, than Orthogonal Frequency-Division Multiple Access (OFDMA) based schemes.

Other key technology towards 5G is Full-Duplex (FD) wireless systems, which enables a radio transceiver to receive and transmit simultaneously in the same frequency and time frame [56]. It can potentially double the system capacity and increase the spectrum utilization efficiency. The main challenge of these systems is Self-Interference (SI) mitigation [57] [58]; in fact, although feasible, some practical imperfections still limit its performance [56].

Aiming to target the spectrum shortage problem and the problem of spectrum under-utilization (spatial and/or temporal), the Dynamic Spectrum Sharing (DSS) concept has been investigated [59].

Another, key technology towards 5G is network slicing [60]. The concept was initially proposed for the 5G-CN [61] and was extended for End-to-End (E2E) network slicing. In [61], the authors evaluate the impact of network slicing in the RAN, namely traffic differentiation, efficient management mechanics to setup and operate new slices, among others.

C. Core Network

Moving on to the 5G-CN, there are also new technologies and concepts being introduced.

1) Cloud Radio Access Network: The current RAN architecture in mobile networks contributes to the network resources underutilization [62], in the sense that the network load at different locations and time instances varies due to the user movements; hence, some BSs can be overloaded while
others are idle (e.g., office locations vs. residential areas during the day). Consequently, the BS processing power is being underutilized in some BSs and used at the maximum in other BSs [63]. To overcome this challenge, the concept of C-RAN was proposed [64], which consists in physically separate the Baseband Units (BBUs) from the Remote Radio Units (RRUs), by centralizing the BBUs into a shared BBU Pool.

Under a C-RAN architecture, depicted in Figure 6, all computational resources are concentrated in a cloud platform which includes the BBU Pool; the BS contains only the RRU to receive and transmit the radio signals to the cloud platform [62].

The BBU pool will serve a particular area providing all the processing to the macro and small cells, allowing BSs to share computational resources. The distance between the cloud platform and the BSs can be up to 40 km where the distance limitation comes from the propagation delay [64]. The fronthaul communications are analyzed in dept in [66].

Another advantage of the C-RAN architecture is the network energy efficiency [64], as a result of a more efficient use of the computational resources. Additionally, as the BSs became less complex, containing only the RRU, the deployment and maintenance cost of new BS is expected to drop [64]. The C-RAN concept will take a key role towards the 5G networks.

2) Multi-Access Edge Computing: Another important concept towards 5G is the MEC. Recent years have seen a paradigm shift towards decentralized computing and cloud platforms, with its realization in MEC, in mobile communications. This aims to push computing, storage and control resources to the edges of the network, thus near the final users. In the MEC architecture, presented in Figure 7, the latency experienced by the users is reduced, while enabling a more efficient usage of the network backhaul and of the core network [67].

MEC results from a strong synergy between Information Technology (IT) and mobile networks domains, which is a current trend, in the sector. Thus, the MEC, can leverage the development of a wide range of new services and applications, especially if information such as contextual information and location awareness, is used to deliver, in real-time, a customized mobile broadband experience to users [69]. Also, as MEC offers an open radio network edge platform, MNOs can monetize it, by allowing third-parties to access the storage and processing capabilities; this may facilitate service enhancements, or even new services, towards not only mobile subscribers but also to enterprises and vertical segments. This should be a secure cloud platform that, through Application Programming Interfaces (APIs), shares access to third-parties [69].

Similar to MEC, there are several related concepts, such as Mobile Cloud Computing (MCC), Local Cloud, Cloudlet and Fog Computing.

MCC takes advantage of mobile computing and cloud computing to enable virtualized computing and storage resources [68], for mobile end-users. It provides all resource-intensive computing executed in clouds without a device with a powerful configuration. Also, it extends the battery life and storage of the end-users devices.

Local Cloud is managed by internal or external sources and are intended to provide services exclusively to a group or institution [68]. In [70], the authors proposed a cooperative scheduling algorithm to manage the local cloud resources together with Internet cloud resources.

Cloudlet is an emerging paradigm, where a small-box data center is deployed at one wireless hop away from mobile devices [68]. It is the middle entity between the mobile device and the cloud, with data routing and security functionality [71]. Thus, cloudlet aims latency and resource sensitive mobile applications. It can be deployed in public places such as hospitals, shopping centers, etc.

Fog Computing is also known as edge computing, supporting ubiquitous connected devices [68]. In this concept, processing is carried out in the local area network or in the IoT gateway. In fog computing, one could retrieve data from several sensors and act accordingly, without having to access a remote cloud platform [68].

As a concluding remark, MEC is highly complementary with C-RAN [72]. Not only the collocation of both technologies allows a more cost effective investment for MNOs but also the MEC APIs can provide access to RAN information.
which, otherwise, would not be trivial, enabling other services.

3) Software Defined Networks/Network Function Virtualization: MNOs have high expectations on these two technologies, SDN and NFV, as they promise to reduce network costs, improve the network scalability and flexibility and provide the base ground for a more dynamic and efficient network.

Firstly, SDN is a centralized networking paradigm where there is a key separation between control and user data [73] [74], as seen in Figure 8.

![SDN Architecture Overview](https://example.com/sdn-diagram.png)

**Figure 8: SDN architecture overview [73].**

Moreover, as the network control functions are centralized in one or more SDN controllers, it enables to simplify the data forwarding of applications and network services [73]. The communication between control and data planes is managed through a communication protocol, being OpenFlow the most used [75]. This is an open source protocol which is controlled by Open Networking Foundation (ONF) [76], and allows to access the flow tables [75] that control the traffic routing in switches and routers. Nonetheless, other protocols exist, such as the Internet Engineering Task Force (IETF) ForCES [77] protocol. As already stated, a key element in a SDN architecture is the SDN controller. It should be not only a platform for deploying SDN applications but also a SDN application development environment [78]. There are several controller implementations, such as the OpenDayLight [78] or the ONOS [79]. Overall, SDN applications interact with the SDN controllers using, as interfaces, Representational State Transfer (REST) and JavaScript Object Notation (JSON); the SDN controller uses communication protocols such as OpenFlow [75].

Thus, SDN presents itself as an alternative to standardized networking protocols, by providing the same role as a centralized software application.

Regarding the NFV [80], it is basically the process of replacing dedicated hardware with software instances which run on cloud environments or general purpose servers [73]. Within this concept, each conventional Network Function (NF) runs in a Virtual Machine (VM) or even in multiple VMs. Each implementation of a NF, using VMs, is called Virtual Network Function (VNF). These instances are deployed and executed in a Network Function Virtualization Infrastructure (NFVI). This is composed by physical resources (computing, storage and networking) which are used through a virtualization layer by the VNFs. Ultimately, and using as example the LTE EPC, each core entity (e.g., Mobility Management Entity (MME), Serving Gateway (S-GW), etc.) could be virtualized as VNFs, where each type of VNF, by forming a common pool, can be scaled independently and according with the network requirements and resources [73]. Towards 5G, not only NFV but also SDN will be key technologies. Many components of a 5G network can be turned into VNFs, allowing accelerated service deployment, when compared to the traditional hardware deployment, network flexibility, scalability and capacity. Above all, VNFs enable network slicing, where the same physical infrastructure can be used to provide multiple and independent logical networks, where the users experience similar conditions as having a dedicated physical infrastructure [81].

Additionally, even though SDN and NFV are independent technologies, both can be enhanced when used together due to their complementary traits. Explicitly, NFV can serve an SDN architecture by virtualizing components such as the SDN controllers or the forwarding data entities. Also, SDN can serve NFV by allowing programmable and dynamic network connectivity between VNFs [73]. This synergy between SDN and NFV is called Software Defined Network Virtualization (SDNV) [82], and is an emerging research area. Both SDN and NFV have the potential to redefine the evolution of network architectures and the potential to be key enablers for the 5G networks.

### IV. INTERNET OF THINGS

Although IoT [7] has been an active research topic in the past years, the concept of a network of smart devices has been around since the mid 80s. Concretely, Mark Weiser’s paper about ubiquitous computing [83], in 1991, set some cornerstones of the actual IoT concept.

The IoT concept of a network of connected physical objects, cyber-systems and sensors everywhere, combined with recent technological advances, opens up the range of IoT applications in different environments, which are being proposed and developed. Besides, the applications mentioned in Section II-C, the concept of smart home [84] and the smart cities are other examples of IoT applications. Extensive work has been developed in creating applications towards enhancing the user lifestyle, especially through the retrieval and analysis of relevant city information [85]. Also, the concepts of Intelligent Transportation System (ITS) and smart healthcare have been receiving important contributions by the IoT community [86].

#### A. Design Requirements

To fully implement the proposed IoT use cases, including the 5G mMTC and URLLC scenarios, there are some main key design principles. A low device (e.g., a IoT sensor) cost will be a key enabler for mass-market IoT applications within the 5G mMTC use case. The IoT network deployment cost, including Services (OpEx) must be minimized in order to provide massive IoT through a feasible economic perspective [86]. In this scenario,
the usage of cellular networks, compared to dedicated Low-Power Wide-Area (LPWA) networks, might be preferred. Also, a key point is energy efficiency, mainly on the IoT device side. Considering that these devices are battery powered and are expected to have a time span of years, without human intervention, energy efficiency is crucial. Thus, the development of lightweight protocols and scheduling optimization are important research areas [86] [87]. Other design principle, specially towards smart metering, is the coverage requirement. Some smart meter locations are in deep indoor scenarios, which require enhanced coverage [88]. Another key topic in the IoT design is the security and privacy of personal data [89]. In [89], the author describes extensively the main concerns in providing security in a IoT environment.

B. Architecture

The IoT generic architecture is organized in three layers: the application layer, the transport layer and the sensing layer [90]. The sensing layer is composed by all cyber-physical objects and sensors. All these devices may collect any kind of data, which is then delivered to the application layer through the transport layer. In the application layer, the data is aggregated and analyzed using intelligent computing, in order to extract valuable information or to trigger actions towards the sensing layer devices [90]. A general IoT architecture is presented in Figure 9.

![General Architecture of IoT](image)

Fig. 9: General Architecture of IoT [90].

Nonetheless, since it is expected that IoT networks will contain millions of devices, large scale IoT (scalable and flexible architecture) is required. Thus, in [91] the authors proposed a self-configuration peer-to-peer architecture. This architecture type, provides automatic discovery mechanisms, enabling the absence of human intervention in the configuration phase.

Other architectures have been proposed, such as the SDN [92] and the cloud computing based [93].

C. Communication Technologies

IoT networks have been widely investigated on the last years, resulting in several IoT communication technologies, which solve the transport layer on the IoT general architecture. There are three types of technologies: long-range, short-range and cellular networks [86]. Long-range networks, or LPWA technologies, are among the most popular IoT approaches. One of them is LoRa [94], which is a physical layer protocol targeting low-cost, low-power and long-range communications [86]. LoRa architecture is a star shaped network where each device has a direct connection with a LoRa gateway.

Also in the long-range IoT, Sigfox [95] technology offers end-to-end IoT connectivity. Sigfox relies on Ultra-Narrowband (UNB) communication technologies, well adapted to a wide range of conventional IoT use cases, that rely on sparse and low throughput communications requirements [96]. Others, as DASH7 [97] or Weightless [98] are also promising long-range solutions.

On the short-range IoT, one of the most used is the Bluetooth [99]. Even though Bluetooth was standardized by the Institute of Electrical and Electronics Engineers (IEEE) as a communication technology for replacing wires in mobile devices, it has evolved to other applications, especially in IoT smart home. The main drawback of Bluetooth is the restriction of only one-to-one communication. In that sense, the Bluetooth Smart Mesh working group was proposed to standardize a new mesh architecture aimed at IoT use cases.

ZigBee, which is developed on top of the physical and data link layers defined in IEEE 802.15.4, is also a short-range communication base for IoT [86]. The main difference of ZigBee compared with Bluetooth, is the range. Whereas Bluetooth operates around a 50 meters range, ZigBee can provide service through hundreds of meters.

Also in the short-range communication domain the Wireless Local-Area Network (WLAN) technology is widely used. Initially, it was designed to support high bandwidth communications between devices. As it does not verifies the modern IoT network requirements, IEEE proposed a low power WLAN, IEEE 802.11ah [100], as an amendment to the legacy standard. WLAN based IoT applications include parking metering, autonomous lightning, smart security, smart home thermostats, etc., [101].

While IoT essentially aims to interconnect a great number of devices and extract value from all collected data, a new paradigm has recently emerged to enhance even more IoT, called Cognitive Internet of Things (CIOT) [102]. The CIOT arises from the application of cognitive abilities to the IoT network. Within the scope of CIOT, heterogeneous smart objects inter-operate through a cognitive centralized entity. This adds an intelligent decision making layer to conventional IoTs networks [102]. Compared with IoT, CIOT is a new and hot research topic.

V. CONCLUSION

The next generation of mobile communications, 5G, is around the corner. Besides MNOs, several economic sectors glimpse at 5G as a key promoter for socio-economic development. Clearly, the mobile network costumers are enthusiastic concerning the upcoming 5G networks and associated services.
The evolution of the current mobile networks to 5G is expected to be gradual, minimizing the CapEx for MNOs. The first 5G commercial deployments are expected by the end of 2018, in a NSA solution, where 5G BSs are anchored in the LTE EPC.

In this paper, apart from an overview of the 5G ecosystem, several key technologies were pinpointed. Concerning the 5G RAN, mMIMO and mmWave will be crucial to accomplish the eMBB requirements. Nonetheless, these two topics are still open research areas. In mMIMO systems, the channel estimation techniques, pilot contamination and the rich scattering environments dependence are still being investigated. Also, the joint operation of mMIMO and mmWave is being explored.

The 5G-CN will constitute an advance of mobile core networks, being empowered by several IT driven concepts. CRAN, MEC, SDN and NFV are candidate solutions heading for the 5G-CN, where work is in motion to solve associated open issues. Overall, these technologies are associated with practical constraints, regarding system inter-operation, orchestration, resource management, flexibility and scalability. The association of multiple technologies, from the ones mentioned, is again a research area of importance.

Also, E2E network slicing is a promising solution towards the development of the future 5G networks. As open challenges, dynamic slicing, end-to-end slice provisioning, slice isolation, are front runners.

As a concluding remark, from the technological point of view, 5G can be considered an evolution of LTE, specially in the early 5G phase. From a socio-economic point of view, it certainly can be seen as a revolution.

VI. ACKNOWLEDGEMENTS

The authors would like to thank FCT for the support by the project UID/EEA/50008/2013. Moreover, our acknowledgment concerning project MESMOQoE (n° 023110 - 16/SI/2016) supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF).


