From 1G to 5G, What Next?

Opeoluwa Tosin Eluwole, Nsima Udoh, Mike Ojo, Chibuzo Okoro and Akintayo Johnson Akinyoade

Abstract—**From the era of electrical telegraphy to the modern internet protocol (IP)-based networks, the evolution from the first generation (1G) to the nascent, fifth generation (5G) networks has been largely gradual, sometimes meteoric. Today, there is the 'big data' buzz, the internet is nearly ubiquitous and everything appears to have 'smart' capabilities given fascinating technologies such as cloud computing, machine-to-machine (M2M) and the internet of things (IoT). Through these eras, developments have witnessed spells of very low data rates (about 10 kilobits per second (kbps)) to very high rates in a staggering magnitude of 1 terabit per second (Tbps), which is hypothetically attainable in 5G technology. This survey paper presents a recapitulation of the evolution from 1G to 5G. It further elaborates the key capabilities and performance targets of 5G; with emphases on intriguing characteristics (ultra-low latency, ultra-reliability, ultra-responsiveness, ultra-fast data rate, ultra-connectivity and ultra-densification), and enabling technologies (software defined networking (SDN), network functions virtualisation (NFV), network slicing (NS), mobile edge computing (MEC) millimetre wave (mmWave), and massive MIMO). The current technological trends due to the evolution are also presented, with highlights regarding the key issues faced by professionals in the field and technology-savvy users in a constantly developing world. Some potential areas of research and development (R&D) are also presented.**

Index Terms—**5G, IoT, MEC, NFV, NS, R&D, SDN**

I. Introduction

SINCE the days of advanced mobile phone service (AMPS), technological advances in cellular systems and telecommunications in general have not ceased. Decade after decade, we have witnessed periods of purely analog systems with no data capabilities (1G), digital circuit-switched systems optimised for full-duplex communication and superb voice telephony (2G), broadband and multimedia systems (3G), all-IP network revolution, and the burgeoning era of unified IP, massive and seamless end-to-end connectivity and mobility (5G). Notably, a critical examination of the timeline from 1G through 5G depicts a palpable trend of the birth of a new generation of cellular wireless technology occurring circa every decade.

Without any doubt, the evolution from 1G to 5G has been all-encompassing, involving significant changes at an architecture level (core network (CN) and radio access network (RAN)) from one generation to the next. These have covered amongst other things, numerous changes at physical, application, transport, session and security layers within the overarching architecture. Even at an end user level, mobile equipment (MEs), mobile units (MUs), mobile stations (MSs), or user equipment (UEs) have changed progressively in line with the technological evolutions attained on the road from 1G to 5G. There have been periods of bulky UEs using macro subscriber identity modules (SIMs) and universal SIMs (USIMs) in comparison to the modern era of smart, portable devices and tablets using micro and nano SIMs.

While 2G will arguably remain the breakthrough moment in the chronology of cellular wireless networking and telecommunications, it may be safe to consider that the nascent 5G technology could be the game changer for the future, especially for companies who can leverage the capabilities it promises to offer, albeit there might be several ups and downs as the technology evolves. There is a potential of generating revenues in excess of trillions of United States Dollars (USDs) across the technology, media and telecommunications (TMT) sector as a whole given the feasibility of the ability to seamlessly combine high-speed multimedia with mobile portability and IoT capabilities as 5G technology continues to develop. In fact, several giants in the networking and telecommunications world (including Qualcomm, Cisco, Samsung etc) have undertaken several studies and provided some forecasts. Some of the study results have revealed that there would be an excess of 20 million 5G-related jobs by 2035 [1]. By 2020, monthly global mobile data traffic would be around 30 Exabytes (EB) (in terms of scale) with approximately 75% of these attributed to mobile video [2].

If we consider all the happenings from 1G till date, one thing is clear; from a technological standpoint, we are heading towards a truly "networked" society, where everyone and everything is connected to everyone and everything respectively, whilst being able to seamlessly access or share information everywhere and every time. From the provider and consumer viewpoints, this will generate numerous technical, commercial and management challenges on the long term realisation of the so-called networked society [4], [6].

The remainder of the paper is organised as follows: Section II provides a concise, but insightful background on the evolution from 1G to 4G. Section III presents the nascent 5G technology, describing what it is and how its architecture will pan out. The section further presents an architectural

Manuscript received January 20, 2018; revised May 24, 2018.

O. T. Eluwole was with Underwriters Laboratories, Basingstoke, UK and EY LLP, London, UK. He is now with BDO LLP, Southampton, UK (e-mail: ope.eluwole@bdo.co.uk).

N. Udoh was with Underwriters Laboratories, Basingstoke, UK. He is now with EY LLP, London, UK. (e-mail: Nsima.Udoh@uk.ey.com).

M. Ojo was with the Politecnico di Milano, Milano, Italy. He is now with the University of Pisa, Pisa, Italy (e-mail: mike.ojo@ing.unipi.it).

C. Okoro was with John Lewis Partnership and BAE Systems, UK. He is now with HSBC, London, UK (e-mail: chibuzo@gmail.com).

A. J. Akinyoade is with the University of Rochester, New York, USA. He is now with Visa Inc., USA. (e-mail: akakinyo@visa.com).

comparison of 5G and prior wireless technology generations. Section IV presents the main features, capabilities and performance targets of 5G, underlining six fascinating characteristics and six enabling technologies. Sections V and VI provide a useful insight into the main current technological trends due to the evolution from 1G to 5G, highlighting some of the key challenges and concerns associated with 5G and these technological trends. Section VII concludes the paper.

II. Recapping 1G – 4G

A. 1G

1G represents an analog transmission technology designed to provide basic voice service. AMPS provided the first commercial cellular phone concept. The radio signals used are analog in nature with no data capabilities, though digital signaling is used to connect the radio towers to the rest of the telephone system by modulating (frequency modulation (FM)) the voice calls to higher frequency of about 150MHz. In other words, frequency division multiplexing (FDM) is used to divide the bandwidth into specific frequencies that are assigned to individual calls. The cell size for a typical 1G network is about 2-20km [7]–[9].

Figure 1 shows a typical 1G AMPS cellular architecture. The fact that 1G networks are based on analog signals/protocol technology (FM) means that one common problem would be susceptibility to interference, which reduces call quality. In addition, there is fundamentally lack of security, as analog signals do not allow the implementation of advanced encryption methods. 1G communication technology is mired by limited capacity (limited number of subscribers), large phone size, poor voice quality, battery life and handover reliability (frequent call dropping) [7]–[9].

In order to summarise the wide-ranging facts regarding the various cellular standards, release versions as well as the pros and drawbacks of each generation from a technical and/or commercial standpoint, tables I and II have been put together to provide a succinct and relevant summary of each generation (i.e. 1G to 5G) in terms of nomenclature, key features and performance targets such as peak data rate, available standards, bandwidth, latency etc.

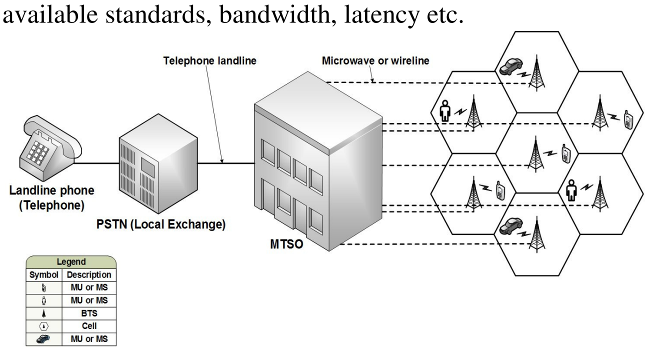

Fig. 1. 1G AMPS architecture; Redrawn version based on [9], [15].

B. 2G

2G technology brought digitisation to cellular networking as it provided the first digital systems as overlays or parallels to analog-based systems. 2G was able to provide a significantly improved voice quality and gave birth to the first data service offering (albeit limited) in the evolution of the cellular networks. Largely, through the use of a more efficient bandwidth/spectrum allocation by way of multiple access schemes such as frequency division multiple access (FDMA), time division multiple access (TDMA) or code division multiple access (CDMA), 2G communication technology was very successful and excellent for voice applications [10]. The latter could not only be digitally encrypted, but also able to provide secure short message service (SMS) and multimedia messaging service (MMS) services to overcome some of the limitations of 1G. In addition, 2G was able to provide a semi-global roaming system to foster connectivity all over the world, a feat that was not achieved by 1G. In practice, the 2G global system for mobile communications (GSM) specification supports cell sizes of up to 35km using macro, micro, pico or femto cells [7]–[9].

Figure 2 represents a typical 2G GSM architecture. After its inception, the 2G era drastically evolved from GSM to general packet radio service (GPRS) and enhanced data rates for GSM evolution (EDGE) (also called pre-3G systems) in 1999, partly due to the insatiable nature of the users who always wanted more in terms of data services, quality of service (QoS) and throughput speeds [10]–[11]. Nevertheless, a 2G network, in particular GSM, has its fair share of drawbacks such as interference issues (including co-channel interference (CCI) or adjacent channel interference (ACI)) due to frequency reuse [10], the pulse nature of TDMA and angular decay curve under unfavourable terrain, topographic or electromagnetic conditions, which could cause intermittent call dropouts or total failure.

At a security level, authentication, encryption and anonymity form the key aspects of security provision in GSM. Although authentication and encryption are provided through A3, A8 (both implemented in the SIM) and A5 algorithmic mechanisms, a number of drawbacks such as crypto flaws, eavesdropping attack (due to invalid security assumptions), SIM attack, false or fake base station (BS), absence of replay protection and denial of service (DOS) are known to be shortcomings of the security arrangements in GSM [12]–[14].

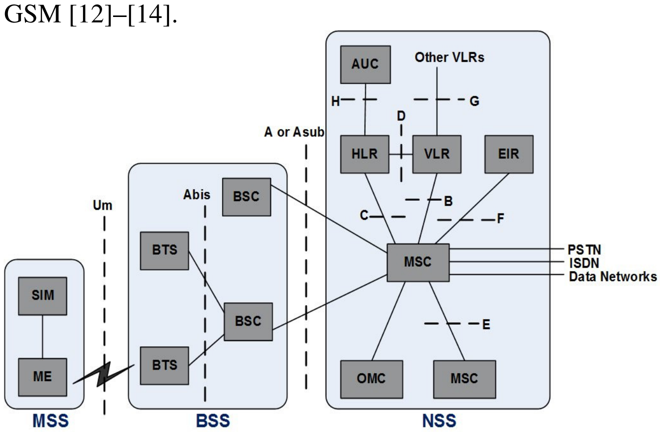

Fig. 2. 2G GSM architecture, Redrawn version based on [7]–[14]

C. 3G

3G provides dedicated digital networks used to deliver broadband/multimedia services. Figure 3 shows the 3G universal mobile telecommunications service (UMTS) architecture. Driven partly by the advancement in internet and IP network technology, 3G architecture provides support for an enhanced data rate (throughput speed) and QoS. Services such as global roaming and enhanced voice quality are important feats of the 3G technology. A slight drawback of the technology is in the area of energy efficiency, 3G UEs consume significantly more power compared to most 2G models and it is less economical to set-up, operate and maintain a 3G network in contrast to the prior generation of networks [3], [11], [17]–[18].

In addition, 3G UMTS is backward compatible with prior generations of cellular wireless technologies through its ability to exist in heterogeneity with the legacy GSM or AMPS technology.

The evolution from UMTS through high speed packet access (HSPA) and evolved HSPA (HSPA+) further provided significantly enhanced end-to-end network performance and eventually led to the development of the next generation of networks i.e. 4G.

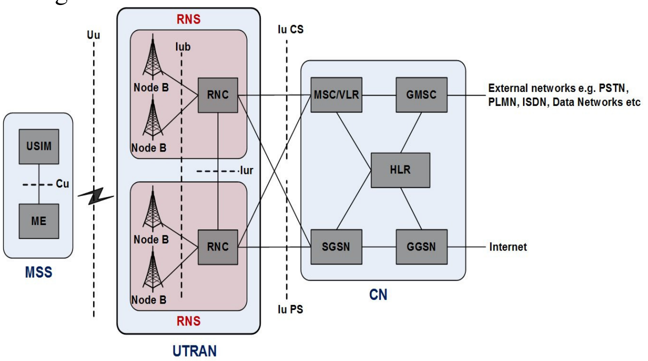

Fig. 3. 3G UMTS architecture; Redrawn version based on [17]

D. 4G

4G represents the generation of mobile cellular communication technology anticipated to productively deliver the demands for broadband data transmission and broadcasting, in addition to very high-volume voice users [3].

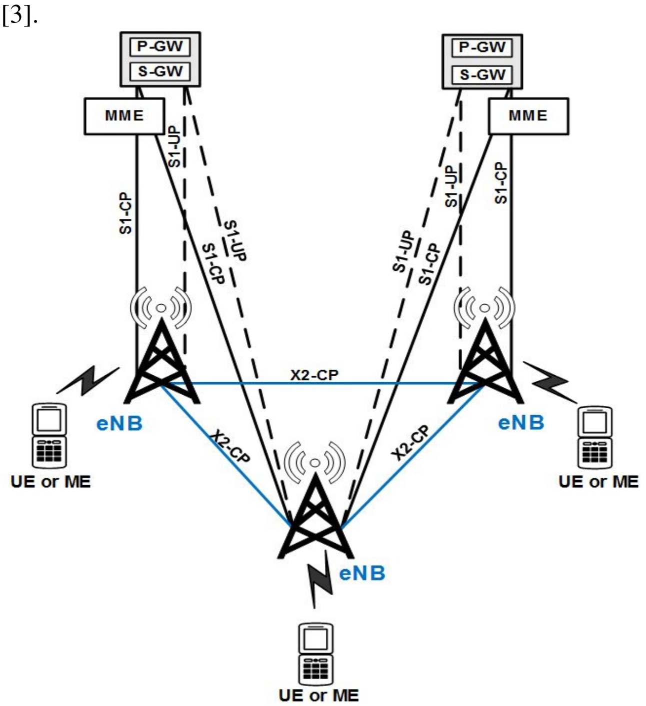

Fig. 4. 4G LTE architecture; Redrawn version based on [19], [20], [26], [27]

From an architecture standpoint, 4G long term evolution (LTE) network is designed with the aim of providing support for packet-switched traffic with seamless mobility, QoS and minimal latency. This approach allows for the support of all services (data, voice, multimedia) through packet connections [20]. Using only two types of nodes namely the enhanced node B (eNB or eNodeB) and the mobility management entity (MME)/system architecture evolution gateway (SAE GW), a highly streamlined architecture (4G-RAN) can be defined for LTE as depicted in Figure 4 [19]–[21].

In Figure 4, the UE or mobile phone is connected wirelessly to the eNB or 4G BS. All radio protocols, mobility management, header compression, ciphering, reliable delivery of packets and all packet retransmissions are orchestrated by the eNB, as the radio network controller (RNC) is incorporated into the latter. On the control side, eNB incorporates functions such as admission control and radio resource management (RRM). The CN is streamlined by separating the user and control planes (UP and CP). The eNBs can communicate with each other using an X2 interface while the eNBs can communicate with the MME in the control plane and/or SAE in the user plane using an S1 interface. The MME/SAE is called the evolved packet core (EPC), while for the whole system the term evolved packet system (EPS) can also be used [19]–[20].

The capability to achieve minimal latency, advancements in multiple input multiple output (MIMO) techniques through various radio access technologies (RATs) such as orthogonal frequency division multiple access (OFDMA) and single carrier frequency division multiple access (SC-FDMA) are part of the key requirements and feats of 4G technology. Numerous innovative concepts such as carrier aggregation, relaying and coordinated multipoint (CoMP) transmission and/or reception (explained in more detail in the next section) ultimately designed and implemented to provide significantly improved peak data rates, support for heterogeneous network deployment and spectrum flexibility amongst other capabilities are wonderful accomplishments provided by 4G technology [19], [22].

Other major enhancements brought by the 4G era are in the areas of multicasting, and interference mitigation. 3GPP Release 12/13 defines several key capabilities and requirements for green computing (energy efficiency), LTE for public safety, emergency and location services communication (machine-type communication (MTC), M2M and IoT); in addition to multi-broadcasting and multicasting services i.e. evolved multimedia broadcast multicast service (eMBMS) [23]–[25].

CCI has been a major bottleneck to achieving higher capacity in cellular networks. However, the evolution from 1G through 4G has produced various interference coordination schemes and interference-aware receivers aimed at mitigating CCI. These have produced promising performance improvements compared to receivers viewing CCI as Additive White Gaussian Noise (AWGN). In particular, 3GPP Release 12 implements a technique, known as network assisted interference cancellation and suppression (NAICS), through considerable improvements to intra- and inter-cell interference mitigation at the receiver side. This is achieved by using advanced receivers to increase the degree of awareness about interfering transmissions (broadcasts) with potential assistance in the network [23]–[25].

It is worth mentioning that other wireless technologies that evolved around the 2G, 3G and 4G era include Bluetooth, Wi-Fi, worldwide interoperability for microwave access (WiMAX) and ZigBee amongst others. Since 2G and 3GPP family of standards (releases), interoperability with previous generations of mobile, cellular wireless technologies has been a fundamental principle in design, development and deployment. This is reflected by the fact that 4G is backward compatible and integrates with various wireless communications technologies ranging from 2G GSM to Wi-Fi and ZigBee.

III. What is 5G and how will its Architecture pan out

A. 5G defined

4G was once tagged the next generation network while 5G has also recently been referred to as network of the future by some scholars in the field [28]. According to [29], in 1G the foundation of mobile telephony was established while in 2G mobile telephony became available for everyone. Fast-track to 3G, the foundation of mobile broadband was realised and the evolution of the latter became the order of the day in 4G.

While it can be argued that 3G and 4G technologies connected people and partly things (objects or artefacts) in the case of 4G long term evolution (LTE) / LTE-Advanced (LTE-A), the developmental efforts to date towards the realisation of 5G makes it look increasingly indisputable that 5G will be able to connect everything, providing a seamless, coalescing connectivity fabric for at least the next decade and possibly beyond [30]–[35]. In other words, it suffices to assert that the advent of 5G will provide limitless access anywhere, at anytime, for anyone, and for anything [6]. This is partly because this anticipated generation of technology, if successful, would create a unified air interface in establishing end-to-end connectivity between mundane things such as smartphones, fridge, freezer, boilers, cars, wearables, utility meters and many more [30]–[35].

To put it tersely in technical terms, 5G brings a world of an appreciably enhanced mobile data broadband, ultra-responsiveness, ultra-reliability, ultra-low latency, ultra-fast data rate and enormous MTC/M2M or IoT capabilities. In 5G, the essential ingredients of radio resource allocation (a key component of RRM) including latency, throughput, reliability, QoS and QoE are expected to be significantly optimised to entirely new, unprecedented levels.

Atop several key concepts, techniques and schemes such as CoMP, SC-FDMA, OFDMA, frequency division duplex-time division duplex carrier aggregation (FDD-TDD CA) etc, 5G will be epitomised by filter bank multicarrier (FBMC) [103], non-orthogonal multiple access (NOMA) [104], beam division multiple access (BDMA) [105], Universal Filtered Multi-Carrier (UFMC) [106] and multi-RAT amongst others. Some of these will be discussed in subsequent sections of the paper. One other key aspect of 5G is the evolution from cell centricity into device centricity, which exploits and harnesses intelligence at the device side (human or machine) such as via device-to-device (D2D) communication of UE assisted mobility [36]. As it implies, D2D is such that there is direct communication between devices and the exchange of UP traffic, essentially avoiding going through a network infrastructure [6]. D2D is further described in section V.

B. 5G architecture – Next Generation RAN (NG-RAN), Non-Standalone (NSA) and Stand Alone (SA) modes

From an architectural standpoint, 5G network design (governed by 3GPP Releases 14, 15 and 16, and projected to be finalised around June 2018), is expected to reuse the 4G LTE CN i.e. EPC, being deployed either in an NSA or SA mode. There is a notable CN in 5G, known as the next generation core or nextgen core (NGC). In 4G, the LTE eNB (4G BS) enables connectivity to the EPC; however, in 5G, the eLTE eNB will enable connectivity to both the EPC and the NGC. The NGC establishes connectivity to the new radio nextgen node B (NR gNB) (5G BS) via the NG interface [37]–[45].

From a RAN perspective, the overall system architecture for 5G i.e. NG-RAN is depicted in Figure 5. The NG-RAN comprises gNBs which provide the NG-RA UP and CP protocol terminations towards the UE. The gNBs are interconnected with each other and connected to the NGC via the Xn interface and the NG interface respectively. The main functions of the gNB include all RRM tasks (radio bearer control (RBC), radio admission control (RAC), connection mobility control (CMC), scheduling), IP header compression and encryption of user data stream, routing of UP data towards user plane functions (UPFs), scheduling and transmission of paging messages and system broadcast information which originate from the access and mobility management function (AMF), measurement and measurement reporting configuration for mobility and scheduling [41].

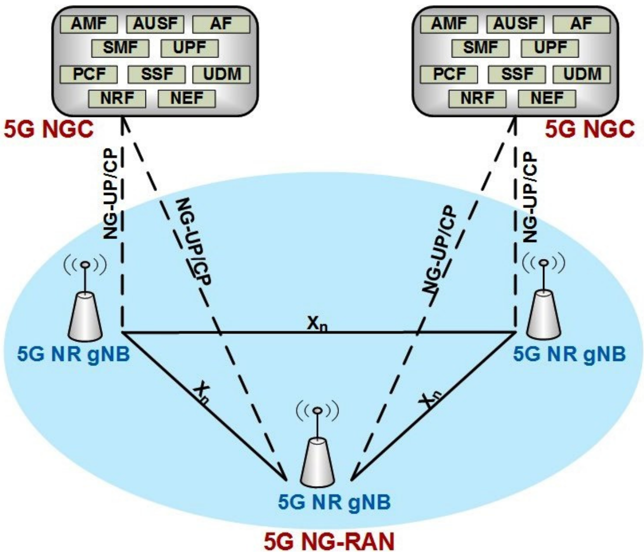

Fig. 5. 5G NG-RAN; Redrawn version based on [41]

In a typical 5G NSA mode shown in Figure 6, the architecture does not really require the use of the NGC as the gNB connects directly with the EPC via the UP interface. In this way, the gNB acts as a secondary serving cell to enhance throughput and capacity, whilst the eNB also connects to the EPC to provide CP functions e.g. paging, tracking, session and mobility management. From a CN viewpoint, not depending on the NGC whilst deploying the gNB, in addition to minimal or no modification required to the EPC makes the NSA architecture very attractive to service providers and commercial mobile network operators (MNOs) [38]–[39], [41].

The 5G SA model on the other hand, shown in Figure 7, is such that, there is a direct connectivity between the NGC and the gNB to facilitate the provision of both UP and CP functions. In this way, the NGC acts as the primary CN for establishing 4G and 5G access via connectivity with the eLTE eNB and the gNB respectively. This is conceptually analogous to the how the EPC establishes 3G and 4G access. The SA mode which uses the NGC as the common CN for all access types is the long term goal for 5G deployment as envisioned by the stakeholders and standardisation bodies driving 5G development and implementation. One of the reasons for this as described by [39] is that this allows network operators to migrate from the EPC to NGC, whilst also providing the ability to offer high-mobility, low-latency, access-agnostic, ultra-reliable and "follow-the-user" services in a flexible manner.

As depicted in Figures 5, 6 and 7, the NGC has several modules, including (but not limited to) AMF, authentication server function (AUSF), application function (AF), session management function (SMF), UPF, policy control function (PCF), slice selection function (SSF), unified data management/user data management (UDM), network function repository function (NRF) and network exposure function (NEF) [40]–[41], [49]. The modularised or layered design of the NGC makes it possible to support various services in a flexible manner [40]. The main functions of the highlighted modules are presented in table III.

While there are ongoing research works, trials, debates and considerations by various institutions and bodies geared towards standardising and ensuring the success of 5G technology, [50] provides some of the notable upcoming events in the run up to the realisation of 5G technology by 2020. These include (but not limited to):

- 2018: Launch of large-scale trials of 5G systems and technology
- 2018: 5G technology showcasing at PyeongChang 2018 Winter Olympics
- 2019: World Radio Conference (WRC) – International agreement on radio spectrum for 5G
- 2020: Early commercial deployment of 5G systems for selected advanced uses
- 2020: Gigabit connectivity based on 5G technology makes debut at Tokyo 2020 Summer Olympics
- Beyond 2020: Full commercial 5G infrastructure deployment.

C. How 1G, 2G, 3G, 4G and 5G architectures compare

Using figures 1–7, table IV provides some key comparisons/contrasts between the main elements of 1G, 2G, 3G, 4G and 5G architectures.

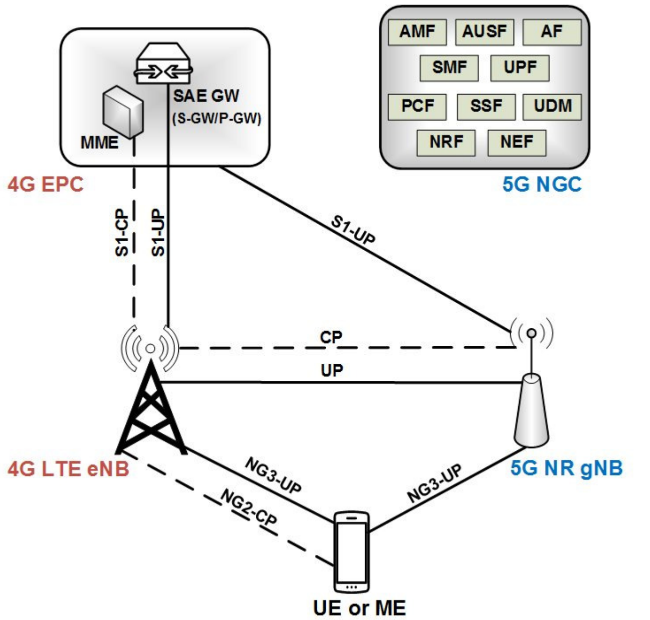

Fig. 6. 5G NSA architecture mode; Redrawn version based on [37]–[45]

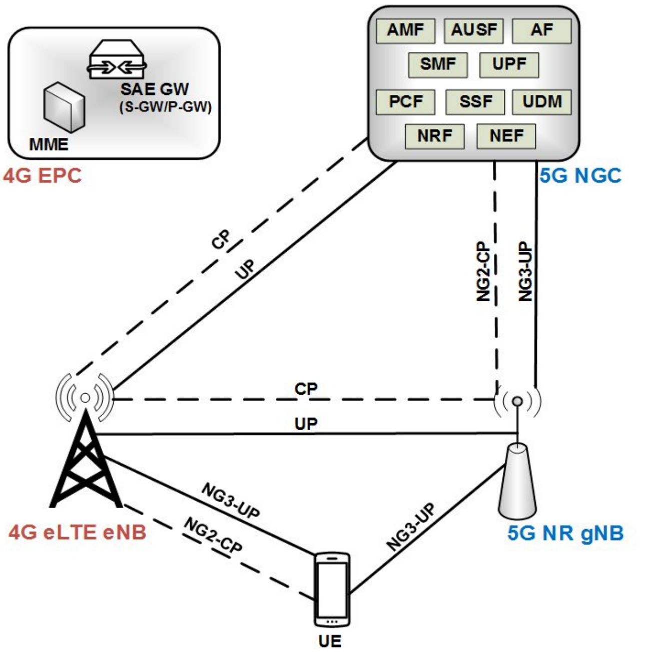

Fig. 7. 5G SA architecture mode; Redrawn version based on [37]–[45]

TABLE I
1G TO 5G – TIMELINE, STANDARDS, RELEASES, RATS, ETC

Generation	Development period	Standard/Technology	Protocol family/Release	Modulation scheme/ MIMO techniques	Protocol/RAT	Forward Error Correction (FEC)
1G	1980s	AMPS NMT TACS/JTACS PSTN	IS-95 TIA-EIA 95 CDMAOne	FM FSK	FDMA	-
2G	1992 – 1997	GSM	CDMAOne	GMSK	FDMA TDMA CDMA	-

2.5G	1998	GPRS	3GPP Release 97	GMSK	TDMA CDMA	-
2.75G (2.9G)	1999	EDGE EGPRS ECSD	3GPP Release 98	8PSK	TDMA CDMA	-
3G	2000 – 2001	UMTS	CDMA2000 IMT-2000 3GPP Release 99 (R99) 3GPP Release 4	QPSK	TDMA CDMA WCDMA EV-DO Rev. A EV-DO Rev. B TD-SCDMA	Turbo codes
3.5G	2002 – 2007	HSPA: HSDPA HSUPA	3GPP Release 5 3GPP Release 6	DL: 16QAM or QPSK (HSDPA)	EV-DO	Turbo codes
3.75G	2007 – 2008	HSPA+	3GPP Release 7	DL: 64QAM or 2x2 MIMO stream UL:16QAM	EV-DO	Concatenated codes
3.9G	2009 – 2010	LTE	3GPP Release 8 3GPP Release 9	DL: 4x4 MIMO UL: 64QAM SISO	DL: OFDMA UL: SC-FDMA	Concatenated codes
4G	2011 – 2015	LTE-A	3GPP Release 10 IMT-Advanced 3GPP Release 11	DL: MU-MIMO (8x8) UL: SU-MIMO (4x4) TDD FDD	DL: OFDMA UL: SC-FDMA	Turbo codes
4.5G (pre- 5G)	2015 – 2016	LTE-A Evolution LTE-M LTE-U comprising LTE-A Pro (LTE- LAA); LTE-LWA and LTE-MulteFire	3GPP Release 12 3GPP Release 13 eIMT-A	eMIMO FD-MIMO Elevation beam forming eCoMP UL: SU/MU MIMO DL: 3D-MIMO FDD-TDD CA 256QAM	DL: OFDMA UL: SC-FDMA	Turbo codes
5G	2016 – Date 2020*	NR?	3GPP release 14 3GPP release 15 3GPP release 16 ITU/IMT-2020	FQAM FBMC Massive MIMO Advanced MIMO	BDMA Multi-RAT NOMA	Low-density parity check codes

*Anticipated release date

TABLE II

1G TO 5G – KEY PERFORMANCE TARGETS

Generation	Data rate/Throughput*	Spectrum flexibility/Channel bandwidth	Spectral efficiency	Latency
1G	9.6kbps	Analog Up to 30KHz (radio channel)	UL: Analog frequency channel (BS) DL: Analog frequency channel (MS)	>1000ms
2G	64kbps	200kHz 1.25MHz (CDMA)	-	300–1000ms**
2.5G	384kbps	200kHz	-	600–750ms**
2.75G (2.9G)	2Mbps	200kHz	-	600–750ms**
3G	2.4Mbps	5MHz 1.25MHz (CDMA)	-	100–500ms** Typical 120ms
3.5G	DL: 14.4Mbps UL: 5.76Mbps	5MHz 1.25MHz (EV-DO)	-	150–400ms**
3.75G	DL: 28.8–168Mbps UL: 11.5–22Mbps	1.4MHz – 20MHz	-	100–200ms** <100ms
3.9G	DL: 100–300Mbps UL: 50–75Mbps	Up to 20MHz (LTE)	DL: 15bps/Hz UL: 3.75bps/Hz	UP: ~10ms (<10ms) CP: <100ms 40-50ms**
4G, 4.5G (pre-5G)	DL: 1–3Gbps UL: 0.5–1.5Gbps	Up to 100MHz (LTE-A)	DL: ~30bps/Hz UL: ~15bps/Hz	CP: <100ms, typically 45ms (actual) UP: ~5ms (sub 10ms) 40-50ms**
5G	1Tbps (over 100m) DL: ≥20Gbps UL: ≥10Gbps	Up to 100GHz	DL: 30bps/Hz UL: 15bps/Hz	≤1ms

*Values denote peak or theoretical maximum potentially achievable unless stated otherwise

**Typical/actual values for deployed networks

TABLE III
FUNCTIONS OF 5G NGC MODULES

Module	Function
AMF	UE-based authentication, authorisation and accounting (AAA) Access control and mobility management Orchestration of network slices selection Controls the SMF and is independent of the access technologies Non-access stratum (NAS) signalling termination and security Access-stratum (AS) security control Inter CN node signalling for mobility between 3GPP access networks Tracking area list management for UE in idle and active mode
AUSF	Data storage for UE authentication
AF	Provision of packet flow information to PCF
SMF	Network policy-based session management (setting up, maintaining and tearing down) Allocation of IP addresses to UEs Selection and control of UPF for data transport DL data notification Traffic steering configuration at UPF to route traffic to proper destination
UPF	Configuration and location management based on service type Packet routing and forwarding Packet inspection of UP traffic Traffic utilisation reporting Anchor point for intra-/inter-RAT mobility when applicable DL packet buffering and data notification triggering
PCF	Provision of policy framework for roaming, network slicing, mobility management, session management and QoS support
SSF	Selection of (network) slice (instance) to which a service must connect
UDM	Storage of subscriber data and profiles of UE
NRF	Provision of registration and discovery functionality i.e. assisting node in discovering network services
NEF	Exposure and publishing of network data

TABLE IV
COMPARING 1G, 2G, 3G, 4G AND 5G ARCHITECTURES

Element/Network	1G AMPS	2G GSM	3G UMTS	4G LTE	5G NR
Key element for mobility management	MTSO	MSC	RNC	EPC	NGC
Base station	BTS	BS or BTS	NodeB	eNB (eNodeB)	gNB (gNodeB)
Switching	Circuit	Circuit	Circuit / Packet	Packet	Packet
Interfaces		Um, Abis, A, B, C, D, E, F, G	Iur, Iub, Iu, Cu, and Uu	X2-CP, S1-UP, S1-CP, S2	NR – NG-CP, NG-UP, S1-UP, S1-CP, NX, Xn, other N-interfaces e.g. N2, N3 etc
Connectivity to external network (PSTN/internet)	Via the MTSO	Via the MSC	Via the GMSC and GGSN	Via EPC?	Via EPC or NGC as per an NSA/SA mode of operation
Applications	Voice only	Voice, data	Voice, data, video calling	Voice, data, video calling, online gaming, HD TV or video streaming	Voice, data, video calling, online gaming, HD TV or video streaming, UHD video, VR/AR/SR, 8K UHD video streaming

IV. KEY FEATURES, CAPABILITIES AND PERFORMANCE TARGETS OF 5G

With commercial deployments set for 2020 and beyond, 5G is expected to be able to [30]–[36]:

- Provide minimal or no latency, gigabit, immersive multimedia experience (i.e. AR and VR) and tactile internet. The latter is discussed in more detail under 5G's intriguing characteristics.
- Meet extremely high capacity and performance demands
- Provide service continuity (high availability) in trains, sparse and dense areas
- Support connectivity for user devices and IoT/M2M devices in excess of 20 million and one trillion respectively, at or very near, 100% reliability.

Some of the key features, requirements, performance targets and enabling technologies for 5G are discussed below and further in subsequent sections of the paper.

A. *Spectrum flexibility*

One or the key characteristic features of 5G includes scalable and flexible bandwidths ranging from 1GHz to circa 100GHz. A typical 5G wireless spectrum will be in the order or 5GHz. For 5G to be considered successful, one of the key criteria will be the efficiency in providing massive system capacity. This apparently calls for network densification. There is no doubt that 5G networks will be much denser than its immediate predecessor (4G) in order to deliver significantly higher cell capacities and per user data rates i.e. ultra-fast data rate for low-mobility users. In making this become a reality, 5G would allow for highly flexible and dynamic allocation of TDD (for higher frequency bands i.e.>10GHz in highly dense environments) transmission resources without any restrictions on DL and UL configurations. In the same vein, flexibility would need to be permitted in a dynamic manner for FDD in orchestrating lower frequency bands [30]–[36].

Given the astronomical rate at which devices are being connected to each other in the modern day via the deployment of wireless sensors or actuators for MTC/IoT, 5G technology will need to be able to handle the network demands to support new models in respect of device and connectivity management. Currently, about a billion of wireless sensors have been deployed worldwide and at this rate [53], virtually everything we can ever imagine will become connected sooner or later once 5G becomes fully operational.

B. MIMO/modulation techniques

Conceptually, MIMO uses multiple receiving and transmitting antennas and actually exploits the effects of multipath as opposed to compensating or eradicating them. The valuable consequences of using MIMO are increased throughput and greater range of operation [54].

Beyond the technological advancements achieved by 4G, 5G design and implementation would incorporate several innovative techniques such as massive MIMO, enhanced CoMP (multi-BS cooperation) and network densification through high density small cells [31], [34].

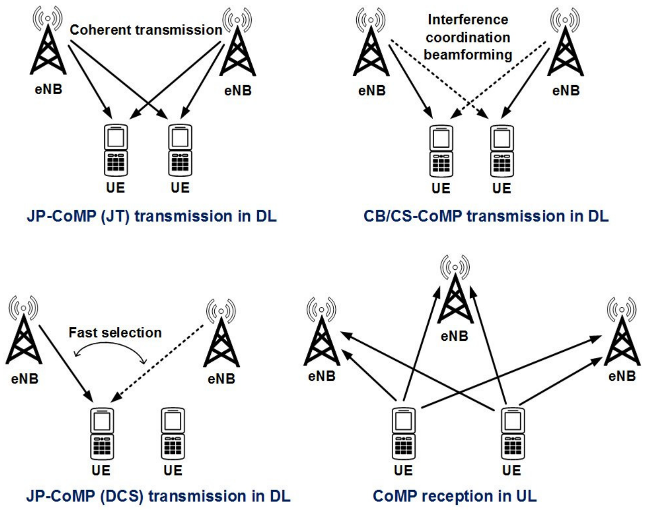

Fig. 8. CoMP technologies; Redrawn version based on [22], [57]

Massive MIMO (also called enhanced multi-user (MU) MIMO, full-dimension (FD) MIMO, large-scale antenna systems, very large MIMO or hyper MIMO) is a significantly enhanced form of MIMO technology which uses a collection of antennas, orchestrated to concurrently serve multiple tens of UEs using a one-time frequently slot i.e. same time-frequency resource. In this way, the benefits of MU-MIMO are expanded to a larger scale [55]–[56]. Massive MIMO is further described under 5G's key enabling technologies (subsection F).

CoMP technologies (shown in Figure 8) on the other hand, aid collaborative processing, transmission and reception of user data in a cellular network and can be very vital in achieving capacity, mobility and efficiency improvement, ranging from spectral through cell-edge throughput to energy efficiency. On the basis of data availability at multipoint, possible types of CoMP include joint processing (JP) CoMP and beamforming-CoMP (CB/CS). JP-CoMP may take forms such as joint transmission (JT) or dynamic cell selection (DCS) while beamforming-CoMP can be coordinated beam switching (CBS) or coordinated beamforming (CBF) (e.g. elevation beamforming, azimuth beamforming etc). In terms of coordinated points, classes of CoMP include Intra-eNB and Inter-eNB CoMP [22], [57].

Massive MIMO in conjunction with multiple access schemes, adaptive coding and modulation such as BDMA, NOMA, FBMC and frequency shift keying and quadrature amplitude modulation (FQAM) will help deliver significantly enhanced spectral efficiency (peak, average user and cell-edge), whilst also seamlessly providing data rates in the orders of Gbps (or more) anywhere. FBMC technique has been envisioned for 5G technology in order to overcome some of the shortcomings of the OFDM-based 4G technology, which includes out-of-band emission and side lobes in the OFDM waveform. The FBMC technique provides a more complex method to drastically diminish side lobes within the spectrum [6], [58].

C. Peak spectral efficiency (PSE)

According to [59], PSE (in bps/Hz), refers to the maximum data rate under ideal i.e. error-free conditions normalised by channel bandwidth (BW). In 5G networks, PSE is critical in evaluating enhanced mobile broadband (eMBB) usage situation. Assuming an antenna configuration that supports 8 spatial layers (8 multiple streams) in DL and 4 streams in UL, the minimum requirements for DL and UL PSEs are 30bps/Hz and 15bps/Hz respectively. In 4G networks, SE requirements are set between 1 – 3 bps/Hz; this is increased to 10bps/Hz in 5G and this takes the PSE achievable in 5G up to 22 times that of a 4G LTE/LTE-A technology [30]–[35].

D. Cost, energy consumption and efficiency

Through the support of multiple RATs providing improved network energy performance, 5G devices are expected to be available at very low cost, with a battery life spanning an excess of 10 years without recharging when used to support MTC/IoT applications–critical machine-type communication (cMTC), ultra-reliable low-latency communication (URLLC) and eMBB (discussed shortly). Until recent years, the evolution of cellular mobile network technologies from 1G through 4G has been such that emphases have been placed largely on peak data rates. 5G is expected to achieve significantly better energy efficiency compared to 3G or 4G or any previous generation of cellular mobile network technology [22], [30]–[35].

E. Intriguing characteristics of 5G

1. Ultra-fast data rate

Peak data rate refers the maximum achievable data rate (in bps) under ideal conditions i.e. absolutely error-free conditions on received data bits transferrable to a single MS under maximum radio resource utilisation (barring those radio resources used for reference signals, guard bands, guard times and physical layer synchronisation). Mathematically, the user peak data rate is defined as the product of the channel BW and the PSE [59].

The peak data rate that can be achieved in 5G is theoretically up to 1Tbps, albeit this is anticipated to be realised around 2030 [34], [60]. Practically, peak data rates in excess of 10Gbps may be achieved in specific scenarios such as indoor and dense outdoor environments. A range of 10-50Gbps can be achieved for low mobility users, with ≥ 100Mbps cell-edge data rate guaranteed for 95% of the users. In urban and sub-urban environments, rates in excess of 100Mbps is attainable while about 10Mbps is possible for IoT applications almost everywhere, including sparsely populated rural areas [30]–[35], [59].

It is worth noting that the IMT-2020 has set minimum requirements for peak data rates in a workable 5G network to be 20Gbps and 10Gbps in the DL and UL respectively [59]. Also, [6] provides a detailed compilation/summary of 5G-related activities being undertaken by several universities, technology groups and other research institutions across the globe.

2. Ultra-low latency

By definition, latency refers to the time (delay), apparently measured in seconds, between the generation alongside transmission of data from one device (such as a sensor) and the error-free reception of the data by another device (such as an actuator) [29]. While throughput tells us how fast data is sent through a network, latency describes how long it takes data sent from a particular source to successfully reach the assigned destination (target). Broadly, latency has four associated components namely: transmission delay, queuing delay, propagation delay and processing delay [16]. Extremely low latency is vital in order to support 5G's essential services or major uses cases, which include a triad of massive machine type communication (mMTC), URLLC and eMBB [30]–[35].

The first class of service, mMTC (also called mIoT) is designed to help achieve the ultra-connectivity (discussed shortly) requirement i.e. wide area coverage for up to several thousands of devices per square kilometre of coverage using very cost-effective software and hardware, whilst achieving highly energy-efficient operations. The letter "m" in mMTC refers to massive number of devices (typically sensors and actuators) which are generally low cost and tend to consume minimal amount of energy for prolonged battery life. Typical mMTC applications include smart agriculture, logistics, fleet or vehicular management and so on [33]. Smart, MTC and IoT technologies amongst others are discussed further in section IV.

URLLC (also called cMTC or cIoT) is essential in order to achieve the ultra-reliability (discussed shortly) requirement. The letter "c" in cMTC denotes critical applications which require very high availability and maximum reliability, incurring minimal or no latency in the process. These applications include automated energy distribution in smart grids, sensor networking and industrial process control or any type of application where monitoring and control occur in real-time [33].

The third category of service, eMBB essentially relates to the provision of ultra-high data rate and ultra-connectivity through an extended support of conventional MBB.

Four main types of latencies can be incurred by a user initiating a new request in a mobile cellular network namely: control plane, user plane, core network and internet routing latency [16], [61]. These latencies are often due to associated components of delay as mentioned previously. In designing cellular network architectures, two of these latencies represent some of the key technical considerations that need to be made in the overarching network design, these are the UP and CP latency.

Assuming unloaded service conditions and an active state of MS, UP latency is the additional time taken in delivering a packet (application layer message) as it traverses from a protocol layer 2/3 service data unit (SDU) ingress point to the protocol layer 2/3 egress point of the radio interface in either DL or UL. In contrast, CP latency is the transition time from an idle state (optimal battery efficiency) to an active state (start of continuous packet transfer). Both latencies are critical in assessing eMBB and URLLC usage scenarios [59].

In terms of metrics, the expected end-to-end latency for 5G is envisioned to be < 5ms, which is a tenth of the typical end-to-end latency achievable by 4G. In fact, 5G is anticipated to incur a UP latency of < 1ms over-the-air–this is nearly imperceptible or real-time, to the extent that it will be able handle practically challenging low-latency services and applications such as self-driving cars, augmented reality (AR), ultra-high definition (UHD) multimedia streaming etc [30]–[35]. To summarise, at a minimum, the following latency requirements [59] have been set for 5G:

- 4ms for eMBB (UP)
- 1ms for URLLC (UP)
- 20ms (CP)

3. Ultra-reliability

According to [62], reliability can generally be viewed as the availability or provisioning of a certain level of service approximately 100% of the time. In the context of cellular wireless networking, [59] defines reliability as the ability to transmit a specified amount of traffic within a pre-determined period of time with high probability of success. This is akin to the definition provided in [29], which defines reliability as being capable of guaranteeing a successful message transmission/delivery within a defined latency budget. Ultra-reliability is vital in assessing the URLLC usage scenario in 5G [59].

Irrespective of the possible types of latencies that may be encountered in a cellular mobile network and the numerous factors which could cause impairment to reliability [63] e.g. irrepressible interference, equipment failure, signal power instability etc, 5G mobile technology is expected to provide an ultra-high reliability and availability, guaranteeing successful packet delivery within 1ms or less with a probability of nearly 100% (99.9999%).

4. Ultra-connectivity

A wide gamut of new applications, scenarios and use cases such as wearable devices, smart cities, homes, stores, offices or cars, critical infrastructure, telemedicine and industrial processes amongst others will enjoy seamless, ubiquitous wireless connectivity with the advent of 5G. To date, 4G LTE/LTE-A has been evolving in a manner that identifies the need to provide excellent coverage for mobile uses; new air interface technologies in 5G networks will be designed to be backward compatible and interoperable with 4G's OFDM access technology in order to serve extreme MBB demands as well as other high-bandwidth and high-traffic utilisation scenarios [30]–[35].

Through the use of advanced D2D and mmWave technology (explained in the next subsection), 5G will foster the attainment of extremely low-latency, whilst also providing support for enhanced connection density in excess of 10^6 concurrent connections per km^2. This will be very vital especially for big data, MTC/M2M and IoT applications [30]–[35].

5G will also support mobility at speeds up to 500km/h (envisioned for high-speed trains) and terminal localisation within 1m [30]–[35]. Mobility (measured in km/h) refers to the maximum MS speed at which a defined QoS can be achieved under a mobility interruption time of 0ms. Four categories of mobility are defined for 5G, including stationary (0km/h), pedestrian (0 – 10km/h), vehicular (10 – 120km/h) and high-speed vehicular (120 – 500km/h) [59].

5. Ultra-responsiveness

Responsiveness refers to a time-based measure of the ability of a component, system, or an entire functional unit to complete a specified task. In order to bring an entirely new dimension to human-to-machine (H2M) communication in 5G cellular networking i.e. the ability to transmit touch and actuation in real-time, the need for an ultra-responsive connectivity becomes vital [36].

From fixed internet, through mobile and things internet (IoT), the evolution of the internet has reached a new form referred to as the 'tactile internet' (TI), a term coined by Professor G. Fettweis some few years ago, which is now widely accepted by the community [64]–[65], [68].

Aside from fundamentally providing ultra-fast data rate, ultra-low latency, ultra-reliability and ultra-connectivity capabilities through 5G, it is vital that all the afore-mentioned are accomplished in an ultra-responsiveness manner in order to be able to efficiently deliver haptic experiences remotely. This is because TI centres on the remote delivery of a physical haptic experience using an ultra-fast, ultra-reliable, and more importantly, ultra-responsive network connectivity, so that real-time interactive systems which are able to steer and control both real and virtual objects can be built [62].

TI enables haptic communications by providing the medium for transporting touch and actuation in real-time [36]. In an interview in January 2017 [66], it was asserted that, "Conceivably, a person or a machine could be in one place, yet apply their physical skills in another place through the tactile internet. This possibility will enable a burst of innovation in so many aspects of our lives that we simply cannot imagine all the applications at this point". [67] upheld this by stating that "the tactile internet will be an enabler for remote skillset delivery and thereby democratize labour and wealth globally".

Being an essential ingredient for 5G, the design goals and technical requirements of TI apparently include: 1ms round trip latency (ultrafast reaction times), high availability (carrier-grade), reliability, robustness, coexistence of human-to-human (H2H), H2M and M2M, and security. If these requirements are adequately met, TI will be vital in multifarious remote applications such as robot steering, monitoring and medical surgery, education and training, servicing and decommissioning etc. Other areas where TI will prove very useful include smart grid (supplier synchronisation), telepresence, drones, self-driving cars, public safety communications systems, AR and virtual reality (VR) amongst a host of others [36], [64]–[65].

6. Ultra-densification

Network densification is a promising cellular technique that leverages spatial reuse to enhance coverage and throughput for 5G cellular network. Network densification has to do with the addition of more BSs and access points (APs) and exploiting spatial reuse of the spectrum, thereby improving network capacity [107]. It is advocated that ultra-dense networks (UDNs) will be the main technology enabler for achieving the 5G requirement of 1000 times increase in mobile network data throughput compared to LTE.

Network densification significantly reduces transmission distance and enables proximity communication, which makes the signal propagation to transit from long to short range propagation [108]. As network densification has a significant impact on network capacity, it also makes interference more difficult to handle. We introduce in the next few paragraphs different ways of combatting interference and further enhancing the system performance in UDNs.

Network capacity can be enhanced via the application of techniques such as interference management, NOMA and mmWave communications. Interference cancelation and interference coordination are the two predominant interference management techniques that have the potential to combat interference and improve network capacity [109]–[111]. Interference cancelation has to deal with the rebuilding of interfering signals, decoding, whilst also removing such from the aliasing signal until the desired signal is retrieved. Interference coordination involves the use of various techniques such as beamforming, power control, user scheduling, advanced receiver techniques etc, such that the desired signal and the interfering signals are forced to be spatially orthogonal at the receiver.

NOMA serves as a promising method to improve user connectivity and network capacity by fully multiplexing available spectrum via non-orthogonal spectrum sharing. NOMA allows multiple users to share time and frequency resources in the same spatial layer via power domain or code domain multiplexing. Compared to conventional orthogonal multiple access technologies, NOMA can accommodate much more users via non-orthogonal resource allocation, and also allows controllable interferences to realize overloading at the cost of a tolerable increase of receiver complexity. Therefore, the demands of spectral efficiency and massive connectivity for 5G can be partially fulfilled by NOMA [104] [112].

mmWave communications technologies over 30-100GHz would serve as a promising complementary to sub 6GHz technologies in ultra-dense networks [113]. Specifically, higher data rates and larger network capacity can readily be guaranteed under mmWave bands. Subsection F below includes addition information on mmWave.

F. Key enabling technologies of 5G

In discussing the key enabling technologies for 5G, it is important to have a quick recapitulation of the cloud concept as this relates to a fair share of 5G's key technology components.

A cloud can be referred to as a huge pool of highly scalable, dynamically reconfigurable, on-demand virtualised computer resources such as storage, processing power, input/output devices memory. These resources are usually provided by a harmony of hardware and software platforms, orchestrated in a manner that an optimum resource utilisation can be achieved at scale. The main characteristics of a cloud or cloud computing system include on-demand self-service (autonomous provisioning), broadband network access (availability for heterogeneity), rapid elasticity (scalability), resource pooling (dynamic resource allocation), and measured service level agreements (SLAs) [69]–[71].

Nearly every organisation or business nowadays including content delivery and social platform giants such as Netflix, Facebook, Twitter, Instagram etc, has one form of cloud-based, technology-enabled service or the other including servers, storage and applications from a cloud service provider delivered to its information technology (IT) landscape via the internet as needed and paid for on a pay-as-you-use basis. Microsoft Azure is a popular cloud platform which provides integrated cloud services. Through a provisioned cloud infrastructure at a data centre and following a public, private, community or hybrid deployment, cloud services can be delivered through three basic service models, including software-as-a-service (SaaS) e.g. Salesforce.com, platform-as-a-service (PaaS) e.g. Google App engine and infrastructure-as-a-service (IaaS) e.g. Amazon web services [69]–[71].

From the manifold research undertaken to date, it is clear that the design and implementation of 5G will encompass emerging, complementary and supporting technologies such as SDN, NFV, NS and MEC amongst others, in order to provide virtualisation and programmability of network, services and control functions. A common theme, called "softwarization" has been coined in recent years, to depict the systemic evolution and introduction of concepts and technologies such as SDN, NFV and MEC to the wider, traditional telecommunication systems landscape [72], [74]. This is possibly rightly so because these technologies involve tactical use of software to perform myriads of functions within the end-to-end network domain across all network layers ranging from the services to the application layer. Perhaps, it suffices to envision that MEC, SDN, NFV (as will be discussed shortly) and related concepts would serve as the platform where core IT and traditional telecommunications converge [73].

From a converged network viewpoint, the key performance targets of 5G such as ultra- low latency, reliability and connectivity will be achieved through network densification, virtualisation and optimisation. Virtualisation enables the simplification of system management by providing the ability to abstract and democratise computer resources while softwarization via programmability involves the use of software to orchestrate applications and services within the network, whilst also simplifying the scaling and management of network infrastructure.

1. NFV

NFV leverages central processing unit (CPU) virtualisation and other cloud computing technologies to establish network functions migration from dedicated hardware to virtual machines (VMs), effectively reducing hardware footprint [73]. Implemented through the virtual network function (VNF) (a software function), NFV enables the separation of network functions from hardware infrastructure so that they can be managed as a software module deployed in a cloud computing infrastructure, effectively providing greater degree of abstraction and increasing overall network flexibility. The VNF software undertakes the task of handling specific network functions that run on the VMs on top of the hardware networking infrastructure such as routers, switches or gateways [72], [74]–[75].

2. SDN

SDN is a networking paradigm which is typically implemented in the CN. SDN establishes a decoupling of control and data planes (CP and DP), enabling direct programmability of network control via software-based controllers. This changes the limitations of current network infrastructures by decoupling the forwarding plane (DP) and the network's control logic (control plane) traditionally coupled with one another. The control plane is implemented in a logically centralised controller (or network operating system), simplifying policy enforcement and network configuration and evolution, and moreover reducing hardware footprint. User data and system control separation in SDN helps to deliver high degree of device-centric optimisation of the active radio links in the network [73], [75]. SDN can also be extended to the RAN in the form of a self-organising network (SON) [62].

NFV can serve SDN by virtualising the SDN controller to be rendered in the cloud, thus allowing dynamic migration of the controllers to the optimal locations while SDN can serve NFV by providing programmable network connectivity between VNFs to achieve optimised traffic engineering. Both technologies are related and paired to each other, but are based upon differing standards with no combined standardised architecture. There are a number of technical aspects which impact several network elements concerning SDN and NFV. Several key areas need to be addressed for standardisation in order to accelerate the adoption of networks evolving with SDN and NFV. Several standardisation bodies like ETSI, IETF, ONF, 3GPP, and IEEE itself are involved in standardising different technical aspects [5].

Furthermore, both technologies can be adopted in the mobile packet core (MPC) network architectures such as the EPC, which is the most recent MPC network that represents the core of the LTE system. SDN and NFV can be adopted in the main functional entities of the EPC, such as the serving gateway (S-GW), packet data network gateway (P-GW), the MME, the home subscriber server (HSS), and the policy control and charging rules function (PCRF) [5].

3. NS

By definition, a network slice can be viewed as an autonomous logical network, created by the interconnectivity of a subset of required building blocks (CP and DP elements) that can be autonomously instantiated and operated over a physical or virtual infrastructure [76]. In other words, with the NS concept, multiple different logical network architectures (i.e. slices) can be defined on top of the same physical (IP services) infrastructure so that resources may be dedicated wholly to a single slice or shared among different slices. Practically, the network slice would be defined across the entire communication system (i.e. end-to-end), including both AN and CN functions alongside their corresponding nodes and end-systems such as MEs [62].

NFV and SDN in conjunction with enhanced analytic tools for RRM will provide autonomous optimisation capabilities to MNOs by following well-defined policy controls [35]. With a much greater degree of abstraction through the unification of SDN, NFV and SON, 5G will be able to offer a NS capability i.e. provision of network-on-demand functions through the delivery of connectivity services based on various custom SDN-established functions that control availability, geographical coverage area, robustness, capacity and security in a dynamic and flexible manner [62], [75].

4. MEC

MEC has been identified as a key technology and architectural model for enabling the transformation to 5G, as well as providing an environment for innovation and value creation [77]. MEC ensures compute, storage and network resources (e.g. MEC application server) are seamlessly integrated with the BS to orchestrate computationally-intensive and latency-sensitive applications such as AR and VR [28]. MEC is about providing IT and cloud computing capabilities within the RAN in close contiguity to mobile subscribers. With MEC, mobile BSs are transformed into smart or intelligent service hubs via the exploitation of context, contiguity, agility and speed. As a result, a wide range of services such as unified mobile communications, RAN-aware context optimisation, IoT, AR, video analytics, distributed content and domain name system (DNS) caching etc, can be provided with minimal latency, high rate of data processing, streamlined network performance and enhanced QoE [36], [77].

Through the provision of a highly distributed computing environment, MEC helps to bring the cloud's concept (dynamic resource management capability) closer to the edge of the network (RAN), simplifying processing and storage of content (information) and bringing the latter in close propinquity to mobile users. This effectively means that users' growing reliance on mobile devices in undertaking compute- and storage-intensive applications are able to be offloaded to happen in the cloud (e.g. the MEC server platform would host software for real-time analytics and machine learning (ML) applications), achieving better flexibility and performance, whilst also extending battery life. MEC helps to add intelligent capability to the traditional RAN typically used for voice calls and data in mobile cellular works through the overlay of distributed edge cloud computing onto it, essentially turning it into a cloud RAN (C-RAN) to provide edge intelligence [28], [36], [75], [77].

The C-RAN concept involves transferring the baseband processing units from cell sites to a centralised location to serve a wide area via fronthaul in order to reduce the hardware footprint at the cell site, whilst more importantly, minimising latency [35], [78]. In conjunction with SDN and NFV, MEC is an essential ingredient in achieving extremely low latency, very high bandwidth and speed in 5G technology, which aims to achieve seamless interconnectivity of trillions of devices [28].

5. mmWave

A 5G small cell deployment in up to 100GHz band using a 2GHz carrier BW" is considered to represent an mmWave [34]. In 4G, spectrum flexibility (channel BW) is scalable up to 20MHz; this will be significantly increased to about 100GHz in 5G technology with the advent of the mmWave technology. MmWave carrier frequencies enable larger BW allocations, which directly translate to higher data transfer rates. A resultant effect will be an increase in data capacity alongside a reduction in the latency for digital traffic as the RF channel BW is increased [80].

One of the key recommendations for 5G deployment centres on establishing multi-connectivity between LTE-A, centimetre wave (cmWave) and mmWave, which can appreciably improve cell-edge performance (with deployment inter-site distance of about 100m), whilst reducing the required density for small cell deployment. An mmWave radio can provide significantly enhanced capacity in the region of several Tbps/km^2 due to additional carrier BW and multiple sectorized antennas orchestrated for the mmWave APs. As frequency is inversely proportional to wavelength (in the "Golden Rule" for waves); having a small wavelength in an mmWave frequency system facilitates the design and orchestration of massive arrays of antennas (e.g. via CoMP-based MIMO), providing high beamforming gains (via adaptive beamforming, relaying and inter-cell interference mitigation) necessary to combat propagation loss in the mmWave band [31], [34], [80]–[82].

6. Massive MIMO

Massive MIMO has been recognised as a promising technology to meet the demand for higher data capacity for 5G networks by 2020 and beyond. As alluded to earlier in the paper, massive MIMO is a communication system where a BS having multiple hundreds (or possibly thousands) of antenna arrays concurrently serve many tens of user terminals, each having a single antenna, in the same time-frequency resource [55], [127]. The BS with multiple antennas sends independent data streams to multiple terminals in the same time-frequency resource.

Massive MIMO relies on spatial multiplexing, which in turn relies on the BS having good enough channel knowledge of both the UL and the DL. On the UL, this is easy to accomplish by having the terminals send pilots, based on which the BS estimates the channel responses to each of the terminals. However, it is more challenging to implement massive MIMO on the DL.

The basic advantages offered by massive MIMO are multiplexing gain [55], [127], energy efficiency [128], spectral efficiency [128], [129], increased robustness and reliability [130], [131], simple linear processing [130], [132] and cost reduction in RF power components [131].

V. Some current technological trends due to the evolution to 5G

A. *Smart, M2M and IoT*

Smart

In English grammar, the word smart has numerous synonyms such as clever, quick-witted, wise, intelligent etc. Technically, smart basically means harmonising or bringing together (i.e. seamlessly integrating) a miscellany of components such as transducers (sensors and/or actuators), monitoring systems, automated controls, modelling systems, decision-support applications etc more 'intelligently' in a bid to ultimately accomplish massive improvements in performance, flexibility and scalability. Therefore, the so-called smart apparently derives input from somewhere (a unit, application or system), applies some wits and then takes actions accordingly [82]–[83].

Interestingly, [83] asserts that, "If a machine/artefact does something that we think an intelligent person can do, then we consider the machine to smart". In other words, being smart entails having an inherent capability to carry out sensing, capturing, monitoring, adapting, inferring, relaying, learning, retrieving, anticipating, translating, self-creating, self-organising and self-sustaining tasks, since these actions require some sort of intelligence or brainpower, albeit at varying degrees. In our world today, we have smart devices (e.g. phones and tablets), cars, systems, homes, environments and even cities.

Smart homes, smart cities and smart grids increase dense and diverse connectivity. Practically, these translate into very high data usage levels, which is extremely daunting in 4G LTE cellular systems. 5G is envisioned be able to provide greater spectrum availability at untapped mmWave spectrum. In addition, 5G networks are expected to provide cumulative capacity for multiple simultaneous users in both unlicensed and licensed spectrum [80], providing required capabilities to address enormous data rates and connectivity in smart cities, homes and grids [133].

M2M

M2M, sometimes called MTC, is a way of establishing direct communication between devices or objects using any one of, or both, wired or wireless communication channels [84]. It refers to communication among devices (e.g. a sensor, actuator or meter) with minimal or no human intervention in the process of capturing an event or status (e.g. humidity or temperature related data), which is then transmitted via a connection-oriented and/or connectionless network to an application (software program) which is able to process and translate the captured event into meaningful information [85].

The fundamental components of an M2M system therefore includes sensors, radio frequency identification (RFID), some form of wired and/or wireless communication link (e.g. Ethernet, Wi-Fi, cellular link) and an autonomous computing software set to assist a networked device manipulate and make meaningful decisions of captured events. Using these components, an M2M system can be built explicitly to undertake a specific task and/or utilise a particular device [84]–[86].

The ETSI M2M architecture describes two main domains in a typical M2M service construct: the M2M device domain (includes the area network) and the M2M network domain (includes the core network) which are connected via gateways. In providing service capabilities, the latter manage the M2M area networks of the M2M devices following an established communication with the M2M core network applications. Communication infrastructure, service functions and application logic are essential to building an M2M solution that is able to provide service capabilities such as remote entity management and transaction management amongst others [85].

At the media access control/physical layer, 2G/3G/LTE cellular communication technology protocols are consequential for M2M development and deployment. They form the fundamental building block upon which the network, transport and service layers could be built for a complete M2M system. LTE-A 3GPP Release10 specifications includes the MTC architecture, features, and standards [85].

The state-of-the-art architecture, recent advances and challenges in communication technologies for M2M are described in [114]. The authors in [114] also describe the architectural enhancements and novel techniques expected in 5G networks related to M2M. Joint use of carrier aggregation and relay station in OFDMA-based 5G wireless is proposed in [115]. Latest advances and developments in architecture, protocols, standards and security for M2M evolution from 4G to 5G wireless are discussed in [116]. Moreover, the authors in [117] present an overview of several features introduced in 3GPP Releases 13 and 14 to address the requirements of M2M communications together with the new air interface characteristics associated with the new UE categories. We expect the cognitive radio to emerge and assist in developing novel cognitive M2M architecture for sensing and using the available frequency bands [118].

IoT

Historically, the IoT concept was first alluded to by researchers at Carnegie-Mellon University in 1982 having successfully connected a vending machine (https://thefutureofsmart.wordpress.com/2014/10/04/the-first-iot-device-a-coke-machine/) to the Internet. Further IoT-related development work on 'ubiquitous computing' was undertaken in the early 1990's [87] prior to the eventual official naming and formulation of the term/acronym IoT by Kevin Ashton (Executive Director of Auto-ID Labs at MIT) whilst making a presentation for Procter & Gamble in 1999 [88]–[90].

The IoT represents a network of connected devices which are able to communicate over the Internet and doing so autonomously without the need for human intervention. IoT is simply about everything being interconnected and interrelated. The IoT depicts a world of universally connected objects, things and processes which collectively creates a digital characteristic of the people and their surroundings in real-time, ultimately providing the best living experience. Simply put, IoT is any IT design, build and implementation having an overarching goal of providing increased connectivity of people and things. The IoT allows the connectivity of smart objects to the internet with a distinguishing capability to enable an exchange of data never previously available, whilst bringing information to users in a much more efficient manner [53], [91].

With a quintet of technology components, specifically sensors (RFID), connectivity, infrastructure platform (for storing /processing data), analytics platform (for making informed decision of data) and user interface (for presenting data), an IoT product can be designed, built and implemented [53], [89]. By using various types of wireless technologies such as 3GPP LTE and LTE-A, Wi-Fi, ZigBee and Bluetooth, seamless connectivity between IoT user-end devices and other IoT gateways or servers can be established [86], [102].

As of 2015, a typical smart phone had about a dozen sensors for orchestrating various IoT functionalities such as temperature, inclination, humidity etc. As of today, fewer than 1 billion sensors have been deployed globally for establishing IoT applications [53]. However, research undertaken by multiple institutions and research organisations have predicted that there will be up to 50 billion connected devices by 2020 and an excess of 30EB of mobile data traffic every calendar month across the globe.

Examining Smart, M2M and the IoT in today's world – benefits, impacts and concerns

Although there are obvious discrepancies, some of the real-life applications and connectivity nature of smart, M2M and IoT technologies appear akin to each other in a manner that is sometimes nearly inseparable, and therefore we may collectively refer to them as a 'triplet' concept. For example, while a smart technology has an intrinsic capability to automatically adapt its behaviour to fit an environment, it practically uses sensors, databases, wireless access etc to collaboratively sense, adapt and provide information to users within the surroundings akin to the IoT technology [82].

A common theme apparently exists between the IoT and M2M systems– minimal or no need for human intervention. In fact, it can be rightly or wrongly argued that M2M forms the basis and serves as an enabler for the IoT technology. It can also be argued that a product or device with in-built M2M capabilities can be considered as being smart. According to [86], IoT networks are also known as M2M area networks. There is often, if not always, a constant link amongst the triad of technologies either in terms of connectivity, remote access/control, or real-world applications and, therefore they may be used interchangeably.

Smart, M2M and IoT technologies have certainly revolutionised the way we live today. On a commercial note, they have brought newer, intriguing and fantastic opportunities to individuals and organisations across the world, whilst also providing solutions to myriads of problems which cut across multiple, if not, all industries and sectors, ranging from utility to healthcare sector. Applications of M2M technology can be seen in remote monitoring e.g. vending machine and telemetry applications, warehouse management, robotics, logistics and supply chain management etc [84], [91]. Also, in the manufacturing, logistic or supply chain sector, having a well-established IoT system provides the ability to facilitate document and fleet maintenance as and when needed. This helps to simultaneously achieve reduced costs, increased efficiency and reliability [92].

From smart devices such as wearable technology products, phones or tablets, to smart cars (adaptive cruise control systems, self-parking etc), smart homes (lighting, controlled heating, door and window solutions such as Nest and Hive) and smart cities (intelligent road networks and parking bays etc.), the IoT and smart technologies are indeed turning the world into a 'global village'. As alluded to earlier in the paper, 5G is expected to achieve the 'real' networked society, where seamless connectivity is attainable everywhere and every time for everyone and everything.

It is almost impossible to defy the good tidings brought to humanity by the advent of the internet since the late 1960s; in fact, we can easily come up with a seemingly endless list of the benefits and logical deliberations in favour of the internet. Fast forward to 1999, the world wide web was invented, and by all accounts, the term IoT soon became a formality.

There is a potential capability to integrate sensor data from numerous distributed artefacts through the use of smart, M2M and IoT technologies, which in turn provide the ability to thwart crime and asymmetric warfare. In addition, an IoT-enabled pervasive positioning technology can help trace lost and stolen goods from a superstore. Popular demand together with advances in smart and M2M technologies could drive extensive dispersion of the IoT that could, akin to the contemporary internet, yield invaluable contributions to the global economic landscape. However, to the extent that commonplace objects turn out to be IT security risks, the IoT appears to have the propensity to distribute such risks much more rapidly and extensively than the internet has to date [92].

According to [93]: "Certainly, the internet has moved the world forward; also, as a result of using the Internet, a lot has gone wrong, quite a lot is going wrong, and a lot more could go wrong. Therefore, it is crucial to develop a balance on how to manage the deployment and efficient use of the internet so that it can be fit for purpose". Unsurprisingly, the IoT has actually been named as one of the top six disruptive civil technologies out of a bucket of 102 potentially disruptive technologies identified at the outset of the study undertaken in [92]. In the report, it was explicitly stated that: "By 2025 internet nodes may reside in everyday things–food packages, furniture, paper documents, and more. Today's developments point to future opportunities and risks that will arise when people can remotely control, locate, and monitor even the most mundane devices and articles".

As with every new or emerging technology, challenges, threats and real risks, particularly to smart and IoT advancements do exist. In a world of connected objects at an unprecedented scale, there is bound to be concerns and challenges. From a technical standpoint, interoperability, interusability, latency, reliability, energy efficiency (most IoT devices are battery-powered) and synchronisation issues have been found to affect the design and deployment of IoT systems [53]. Above all, there is a palpable sense of jeopardy to data privacy and confidentiality, which is understandable because the torrent of data generated via the interconnectivity of smart and IoT-enabled devices in the modern society has an apparent tendency to reveal so much about people than is currently being envisaged by the advocates and developers of these technologies.

B. D2D

D2D communication is a promising technique to improve resource utilisation in 5G cellular networks by offloading the traffic from backhaul to local direct links. D2D wireless networks are considered one of the candidates for future 5G networks. A comprehensive review of D2D communications is available in [119]. Similarly, a survey of the existing methodologies related to aspects such as interference management, network discovery, proximity services, and network security in D2D networks is presented in [120]. In the next few paragraphs, we briefly highlight the major research works relevant to the context of emerging 5G wireless communications.

Major recent research activities in D2D include game theoretic pricing schemes [121], channel measurement and modelling [122], proximity services (ProSe) [123], and public safety networks [119].

An overview of the major challenges in a two-tier cellular network that involves a macrocell tier (i.e. BS-to-device communications) and a device tier (i.e. D2D) with some proposed pricing schemes for different types of device relaying are discussed in [121]. Scheduling algorithms for effectively sharing multimedia content using D2D communications are proposed in [124]. In addition, [125] describes Network Assisted Routing (NAR) algorithm for D2D communication in 5G cellular architectures with the goal to extend the coverage of BSs.

More importantly, extremely low latency, high energy efficiency and scalability are vital to 5G networks. Thus, it is essential to decrease the control signalling and end-to-end latency in network assisted D2D communications [126].

C. Big Data – Definition, Impacts and Challenges

With the propagation of smart, M2M and IoT technologies, perhaps, unsurprisingly big data comes on the scene. There is basically a cornucopia of ‘real-world’ ubiquitous data about people and what they do, applications, artefacts and things, albeit the data can be characteristically viewed as being dirty, noisy, poorly defined, redundant, unstructured, semi-structured, multidimensional, disparate and volatile, to mention a few [94].

In today’s world of comprehensive data, there is a perceived universal picture of the physical world more than ever previously attainable with the advent of computing and internet technologies such as the cloud and IoT. Interestingly, our smart, IoT-enabled mobile phones and devices of today are able to tell us estimated time of arrival to reach our homes, offices or attend a concert. We get reminders and alerts informing us when it is best to leave for the airport in order to meet a flight in time, or some wearable device telling us how well our hearts have performed in a particular day or over a period of time.

In simple terms, big data refers to huge and complex sets of data whose scale is in the order of quintillion (10^{18}) bytes i.e. EBs or higher. The sheer scale or volume of data generated on a daily basis in the modern world has gone beyond the levels of megabytes (MBs), gigabytes (GBs), terabytes (TBs) and petabytes (PBs) to a few EBs, and as a result, it is becoming increasingly difficult for traditional data management tools and technologies to efficiently handle such scales of data [94], [96].

More importantly, we use the term big data when a quintet of unique features or characteristics (i.e. volume, velocity, variety, variability and veracity – the five Vs) can be attributed to a particular dataset. Volume apparently indicates the massive amount of data which need to be seamlessly manipulated or handled to maintain high availability for storage, processing and retrieval. Velocity relates to the speed of data ingress and egress which could complicate the processing and analyses of the data from a resource management (load balancing) viewpoint. Variety and variability centre on the range and types of data sources which could be too varied or great to assimilate, prove, transform or track for structural consistency and future reuse. Veracity focuses on ensuring ingested, processed, analysed or manipulated data is of great and proven quality (i.e. sufficiently meets referential integrity, consistency, validation, reconciliation and provenance tests) given the high-volume, high-velocity and high-variety characteristics of the data [94]–[97].

Enterprises are increasingly investing in big data analytics platforms to identify market trends and gaps in their bids to remain highly competitive amongst their peers. However, translating real world ubiquitous data into useful, meaningful formats suitable for orchestrating a prescriptive, predictive, diagnostic or descriptive analytics may be very challenging given the multiple disparate sources of the data, reflected in fundamental discrepancies in models or structures, which could range from schema-free to flat, relational, nested, object-oriented, hierarchical, or non-flat data and so on. Therefore, with big data there are bound to be data quality issues and challenges. This is particularly so considering the vastly complex and heterogeneous environments of most enterprises looking to utilise and maximise the power of big data analytics solutions [94]–[97].

While there are several software platforms, packages, tools and database technology systems designed and used for big data analytics e.g. MS Azure, Cloudera, Apache Hadoop, Couchbase, MongoDB, RStudio etc, in addition to various algorithms and techniques suitable for handling the five Vs e.g. statistical machine learning (SML), neural networks (NN), Bayesian, Bandit, principal component analysis (PCA), support vector machine (SVM), and single value decomposition (SVD) amongst others [94]; there is no commercial package till date that is able to provide an all-encompassing ‘out-of-the-box’ solution to the main problems in big data and data analytics (i.e. prescriptive, predictive, diagnostic or descriptive) in general. These key problems relate to areas including (but not limited to) classification (categorisation), regression, clustering, dimensionality reduction, multi-collinearity, reproducibility/replicability, hyper-parameter optimisation, algorithmic computability, pre-/post-processing of covariates (input variables) and target responses (output variables) amongst others [94]–[95]. Extensive experimentations would be vital in tackling these challenges which could aid the development of quantitative and qualitative techniques and processes for managing big data.

Owing to an incessant proliferation of data from smart, M2M, IoT and cloud technologies, which may be collectively viewed as a form of technology convergence, it is highly unlikely that R&D into big data solutions (techniques, methodologies, platforms etc) will cease in the near future.

VI. Challenges of 5G and Radically Changing Technologies in the Modern Era

A. *Performance*

In order for 5G technology to be successful, [98] highlights the need to move from a traditionally-driven viewpoints of coverage, reliability, reduced capital expenditure (CAPEX) and operational expenditure (OPEX) to a more value creation and SLA perception. It further adds that a fundamental change in design approach from a predominant measured key performance indicator (KPI) (QoS-centric) to a more perceived KPI (QoE-centric) will be important in achieving 5G's key goals and performance targets.

Providing concurrent seamless wireless communications, ultra-connectivity, mobility on demand, immersive multimedia experience, ultra-responsiveness (TI), MEC, NS, and NFV capabilities etc – (the intriguing characteristics and enabling technologies of 5G) – will encounter technical hurdles and performance bottlenecks. There are bound to be challenges in the design, build and implementation for a sustainable anticipated 5G ecosystem (2020) and beyond. The role of R&D in resolving these challenges cannot be overemphasised.

For example, as promising as massive MIMO is (given the capabilities it is able to offer); it attracts several challenges (e.g. scheduling complexity, link adaption problems etc), which require substantial experimental efforts to overcome. Although work has been carried out by experts in the field, some of these challenges are exacerbated as the number of antennas at the BSs increases. This is partly due to the fact that the accuracy of the channel state information (CSI) is a prerequisite for achieving reasonable performance improvements from a massive MIMO system. The CSI has direct impact on precoding design, coding or modulation scheme (definition, acquisition, or assignment), signal modulation and demodulation. These factors can result in FDD challenges such as pilot and feedback overheads, CSI estimation complexity and quantization. While several researchers have undertaken a detailed study of massive MIMO techniques, applying and adapting the concept in ultra-fast data rate systems, whilst also providing minimal complexity and highly accurate CSI quantization remains an ongoing research challenge [31], [81], [98].

Also, due to the growing trend of bringing intelligence in proximity to the edge of cellular networks (MEC, C-RAN, SDN, NFV, NS), cutting-edge networking design and approaches will be required in overcoming several challenges [36], [62], [65] such as, providing 1ms round-trip latency and an enhanced haptic perception (multisensory or multimodal feedback), establishing the optimal waveform selection and robust modulation methodologies (physical layer challenges), designing CP and UP separation and coordination techniques, building optimal resource management and task allocations schemes (online or offline scheduling). In resolving some of these challenges, R&D into artificial intelligence (advanced machine learning and predictive analytics) will be vital as several critical factors will need to be considered in designing workable 5G-enabled solutions. Some of the factors include the control server and processing mechanisms for sensors and actuators (smart, MTC, IoT and MEC technologies), real-time operating system for ultra-low latency capability, and multithreading capability for multiple transmissions over parallel communication channels.

Typically, cloud computing technologies are known to attract technical challenges in the areas of data migration and integration, resource management (due to dynamically changing and varying workloads), data privacy and security (data centre location based on deployment model), QoS and performance. As cloud technologies and computing methodologies continue to proliferate and now being introduced to the budding 5G wireless technology, therefore R&D into several innovative techniques for orchestration and management (including QoS and QoE) will be necessary. These will include designing and implementing novel methods for establishing optimal load balancing required for user experience optimisation and network capacity given the possibility of multiple users around a boundary zone having a great network coverage, but sub-optimal requested data rate [34]. In addition, to ensure efficient QoS and QoE, new techniques and solutions will be required such as defining resilient, error-free thresholds for detecting, monitoring and correlating traffic, whilst also ensuring the network's capability to achieve best possible performance (in respect of latency, throughput, spectral efficiency, error-rate, connectivity management etc) without compromising security.

B. *Trust, privacy and security*

Security is vital in any technological feat achieved and will arguably remain the biggest concern in the TMT sector and the wider IT industry in the modern era. In [99], it was considered that a technology may not be regarded as top-class until it amply and capably fulfils all the core requirements of modern cryptography, including confidentiality, authentication, integrity and non-repudiation. Security becomes particularly vital in 5G technology given the major use cases (mMTC, URLLC and eMBB) the technology is envisioned to propagate.

The introduction of cloud computing, virtualisation, softwarization etc to telecommunications or cellular networking will incur new, significant challenges. Therefore, R&D in areas such SDN, NFV, MEC, NS etc will be fundamental to 5G's success. Admittedly, virtualisation will enable flexible slicing, mobility management and the ability to customise radio resources for different applications and UEs [62] while softwarization will help achieve low costs (streamlined CAPEX and OPEX), since it will be possible establish ubiquitous connectivity using minimal physical infrastructure, increasing cost competitiveness between network operators (including mobile and virtual operators). However, new methodologies will need to be developed to ensure an assuring security can be provided, whilst also taking into consideration the potential consequential impacts from both socio-economic and techno-economic perspectives.

In underlining the need for privacy and security in 5G networks, [32] talks about the need for the development of innovative trust and service delivery models to handle radically changing and evolving threat landscape, highlighting several key areas of R&D with regards to 5G security. These include identity management, radio network security, security assurance, energy-efficient security, cloud security, and flexible/scalable security architecture. In addition to these, the sought solutions would need to take into consideration all the germane standards and regulations (e.g. data protection act (DPA), general data protection regulation (GDPR)) without jeopardising the end-to-end availability, performance (e.g. relaying in D2D set-up), resilience, reliability; and other 5G essential requirements.

C. Flexibility

Typically, the overall network topology in any telecommunication system is often subject to change due to the heterogeneity of network nodes. If we take into account the addition of a distributed system of server clusters (MEC) working in unison with SDN and NFV to facilitate , then the design of a practicable and cost-effective 5G architecture will need to be able to successfully operate in a manner that is 'agnostic' to the heterogeneity of the network irrespective of the mode of operation (NSA or SA), otherwise the envisioned capabilities through the use of MEC, NFV, SDN etc may not be realised in a performant and flexible manner.

In [76], architecture modularisation and slicing for 5G networks was proposed so that the projected set of use cases could be achieved with adequate flexibility across a HetNet landscape. This was centred on the use of six fundamental building blocks – access function, connectivity management, mobility management, flow management, security and AAA management, and context generation and handling function (CGHF). While the modularisation concept appears promising, significant R&D work would need to be undertaken in the design (e.g. device selection and association), instantiation and operation of network slices (e.g. concurrent connectivity to multiple slices, context information sharing etc) in line with the core building blocks.

D. Energy efficiency

As alluded to earlier in the paper, one of the key features, capabilities and performance targets of 5G is to provide support for low-cost, minimal energy consuming devices (a prolonged battery life without having to recharge). Energy efficiency of the 5G network is an essential ingredient in lowering operational costs which in turn could significantly increase the reach of wireless connectivity to very remote areas [33]. Energy efficiency has been gaining more attention lately, [22] proposes a new power consumption model for LTE/LTE-A wireless network, however further R&D need to be undertaken in this area (such as designing and creating dynamically controlled, adaptive machine learning-based power management systems [98]) in order to achieve an all-round success of 5G wireless access by 2020 and beyond.

VII. Conclusion

Having recapped the key highlights from 1G through 4G technology, this paper has presented some of the salient points regarding 5G technology in terms of its key features, capabilities, performance targets and technology components. Key fascinating characteristics and enabling technologies have been underlined. To a reasonable extent, it may be inferred that 5G is the next big thing in cellular and wireless systems and telecommunications in general.

Aside from the inadequately resolved, unaddressed technical challenges, and the limitations of the prior generation of technologies, we may arguably consider the insatiable nature of technology users and service consumers, often desirous additional capabilities (such as high speed data rate systems and extremely low or no latency systems) has generally resulted in an astronomical rise in the demand and consumption of wireless communication radio resources in the modern era. As a consequence, technology developers and service providers have had to innovate a new generation of technology atop 3G and 4G.

5G technology seeks to achieve a 1ms round-trip latency for major use cases and critical applications. In addition, it is anticipated that ultra-connectivity, ultra-high reliability and extremely high availability etc, reflected in everything and/or everyone interconnected and able to seamlessly share information harmoniously, efficiently, reliably and securely irrespective of time or location, whilst also prohibiting overall system uncertainty will be achieved. On the presumption that all of these (and many more as discussed in the paper) will be feasibly achieved, would these suffice in fulfilling the demands of the contemporary users or would there be another generation of technology within the next decade and subsequently? Perhaps a more appropriate question ought to be: are we passably prepared on all accounts for the looming 5G technology, including socio-economic, techno-economic, regulatory, and sustainability facets?

It is apparent that 5G technology will radically change the cellular, wireless and telecommunications landscape in the near future and possibly beyond. The advent of 5G will pave way for the development and deployment of new services and cost-effective telecommunications infrastructure. However, to the extent that the distinct characteristics and enabling technologies of 5G encounter trade-offs during deployment and implementation (e.g. between performance and security which is always a significant concern), the long term viability and sustainability of 5G technology will remain questionable.

Be it gradual or rapid in propagation, it will surely be a matter of time to discover whether 5G technology will ultimately bring about the supposed networked society. Will 5G be ultimately able to provide the multifarious technical capabilities and benefits as envisioned or will it create a situation where the associated risks and technical challenges markedly outweigh the perceived benefits and capabilities? In any case, as we look forward to the impending era of 5G technology and beyond, R&D of novel techniques and enabling technologies in the area of mobile, cellular and wireless communications will certainly not cease, at least, in the short term.

APPENDIX

Definitions of some key terms and some common acronyms used (including acronyms that are not defined elsewhere in the paper) are presented in tables V and VI respectively.

TABLE V
DEFINITION OF TERMS

Term	Definition
Augmented Reality (AR)	A technology whereby computer-aided (i.e. digital) information based on user context is integrated with the user's environment in real-time, such that the information is graphically enriched or augmented to the display. AR finds its applications in healthcare, retail, public safety, oil and gas, etc [31], [51].
Availability	Refers to a system's endurance against potential outage scenarios [29].
Centimetre wave (CmWave)	A type of wave produced by/from a 5G small cell deployed in a 6GHz – 30GHz band using about 500MHz BW [34].
Cognitive radio	"Cognitive radio is an intelligent wireless communication system that is aware of its surrounding environment (i.e., outside world), and uses the methodology of understanding-by-building to learn from the environment and adapt its internal states to statistical variations in the incoming RF stimuli by making corresponding changes in certain operating parameters (e.g., transmit-power, carrier-frequency, and modulation strategy) in real-time, with two primary objectives in mind: highly reliable communications whenever and wherever needed; efficient utilization of the radio spectrum" [134]
Connection density	Refers to the total number of devices fulfilling a specific QoS per square kilometre [59].
Converged network	Refers to the seamless coexistence of telephony and multimedia within a single network.
Downlink	Also called forward link, it refers to the signal transmission in a direction from the BS to the MS [3], [100]–[101].
Fronthaul	Refers to the connection between a new network architecture of centralised baseband controllers and remote standalone radio heads at cell sites [79]. It is kind of analogous to the backhaul, which connects the mobile network to a wired network.
Haptic information	In tactile internet, haptic information comprises two separate types of feedbacks – kinesthetic feedback (e.g. force, torque, position, velocity) and tactile feedback (e.g. surface texture, friction) [62].
Quality of Experience (QoE)	Influenced by several factors such as content, service, network, application, device and usage context, QoE "refers to the degree of delight or annoyance of the user of an application or service" [46]. ITU-T defines QoE as the "overall acceptability of an application or service, as perceived subjectively by the end user" [47]. In [48], QoE intelligence is considered to be highly valuable to telecommunication network operators (including MNOs).
Quality of Service (QoS)	Refers to a collection of characteristics regarding the performance of any connection or network, which contributes to the level of satisfaction, observed and derived by an end user vis-à-vis the network's performance. More technically, QoS means that which a flow strives to achieve. Peculiar features of a flow include latency, jitter, reliability, bandwidth, and data rate. Notably, QoS delivery to a network may be nonguaranteed (best-effort delivery) or guaranteed [16].
Scheduling	Refers to the dynamic allocation of resources to UEs in both DL and UL [41].
Synchronised Reality (SR)	Coined by researchers at King's College London UK (http://www.pressreleasepoint.com/worlds-first-5g-end-end-network-debuted), SR as the term implies, allows the synchronisation of the real, virtual and mental worlds far beyond VR and AR in isolation. In other words, SR creates a technology where VR, AR (the physical surroundings) and mental stimulations are synchronised i.e. seamlessly combined
Throughput	Measured in bits per second (bps), it describes the rate at which data (text, voice, video, multimedia etc) is transported through a network [16].
Uplink	Opposite to DL and also called reverse link, it refers to the signal transmission in a direction from the MS to the BS [3], [100]–[101].
Virtual Reality (VR)	A technology whereby physical (i.e. real) presence is simulated by computer graphics, allowing the user to interactively interact with the simulated elements (i.e. virtually). VR creates a wholly artificial environment. Applications of VR can be found in immersive sports broadcasting, medicine, architecture, entertainment, arts etc [31], [52]. Both AR and VR help users to interact as if they are in the same location. From an experience perspective, while VR wholly replaces a user's audio and visual sensations, AR enriches this by providing further information that is germane to the surroundings [32].

TABLE VI
ACRONYMS

Acronym	Full form
1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
AUC	Authentication Centre
BSC	Base Station Controller
BSS	Base Station Subsystem
BTS	Base Transceiver Station
eCoMP	Evolved CoMP
ECSD	Enhanced Circuit-Switched Data
EGPRS	Enhanced GPRS
eIMT-A	Enhanced International Mobile Telecommunications Advanced
EIR	Equipment Identity Register
eLTE	Evolved LTE
eMIMO	Evolved MIMO
ETSI	European Telecommunications Standards Institute
EV-DO	Evolution Data Only/Optimised
GGSN	Gateway GPRS Support Node
GMSC	Gateway MSC
GMSK	Gaussian Minimum Shift Keying
HLR	Home Location Register
HSDPA	High Speed Downlink Packet Access
HSUPA	High Speed Uplink Packet Access
IEC	International Electrotechnical Commission

IMT	International Mobile Telecommunications
ISDN	Integrated Services Digital Network
ITU	International Telecommunication Union
ITU-R	ITU Radio communication standardisation sector
ITU-T	ITU Telecommunication standardisation sector
JTACS	Japanese Total Access Communication System
LTE-A	LTE Advanced
LTE-LAA	LTE-A Licence-Assisted Access
LTE-LWA	LTE-A Wi-Fi Link Aggregation
LTE-M	LTE for Machines (MTC, M2M, IoT)
LTE-U	LTE-A in Unlicensed Spectrum
MSC	Mobile (Services) Switching Centre
MSS	Mobile Station Subsystem
MTSO	Mobile Telephone Switching Office
NMT	Nordic Mobile Telephony
NSS	Network Switching Subsystem
NX	New Flexible air interface for 5G
OMC	Operations and Maintenance Centre
PLMN	Public Land Mobile Network
PSK	Phase Shift Keying
PSTN	Public Switched Telephone Network
QPSK	Quadrature Phase Shift Keying
RNS	Radio Network Subsystem
SGSN	Serving GPRS Support Node
SISO	Single Input Single Output
SU-MIMO	Single-User MIMO
TACS	Total Access Communication System
TD-SCDMA	Time Division Synchronous CDMA
TIA-EIA	Telecommunications Industries Association - Electronic Industries Association
UTRAN	UMTS Terrestrial RAN
VLR	Visitor (Visited) Location Register

REFERENCES

[1] K. Campbell, J. Diffley, B. Flanagan, B. Morelli, B. O'Neil and F. Sideco, "The 5G economy: How 5G technology will contribute to the global economy," The IHS 5G economic impact study, IHS Economics and IHS Technology, January 2017.

[2] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020," White Paper, February, 2016.

[3] D. T. C. Wong, P. Kong, Y. Liang, K.C. Chua, and J.W. Mark, "Wireless Broadband Networks," John Wiley & Sons, April 2009.

[4] M. Shafi, et al. "5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice," IEEE Journal on Selected Areas in Communications Vol. 35, No. 6, pp. 1201-1221, 2017.

[5] V. Nguyen, A. Brunstrom, K. Grinnemo, and J. Taheri, "SDN/NFV-based Mobile Packet Core Network Architectures: A Survey," IEEE Communications Surveys & Tutorials, Vol. 19, No. 3, pp. 1567-1602, Third Quarter, 2017.

[6] A. Gupta and R. K. Jha, "A Survey of 5G Network: Architecture and Emerging Technologies," IEEE Access Special Section on Recent Advances in Software Defined Networking for 5G Networks, Volume 3, pp. 1206 – 1232, 2015.

[7] J. H. Schiller, "Mobile Communications," 2nd Edition, Pearson Education Limited, 2003.

[8] W. C. Y. Lee, "Wireless and Cellular Telecommunications," 3rd Edition, McGraw-Hill Companies, Inc., 2006.

[9] I. Stojmenovic, "Handbook of Wireless Networks and Mobile Computing," John Wiley & Sons Inc., 2002.

[10] J. Eberspächer, H-J Vögel, C. Bettstetter and C. Hartmann, "GSM – Architecture, Protocols and Services," 3rd Edition, John Wiley & Sons Ltd, 2009.

[11] J. Bannister, P. Mather and S. Coope, "Convergence Technologies for 3G Networks - IP, UMTS, EGPRS and ATM," John Wiley & Sons Ltd, 2004.

[12] P. Chandra, "Bulletproof Wireless Security - GSM, UMTS, 802.11 and Ad Hoc Security," Elsevier Inc., 2005.

[13] M. Stamp, "Information Security – Principles and Practice," John Wiley & Sons, Inc., 2006.

[14] T. S. Rappaport, "Wireless Communications - Principles and Practice," 2nd Edition, Prentice Hall Inc., 2002.

[15] L. Goleniewski and K. W. Jarrett, "Telecommunications Essentials, Second Edition: The Complete Global Source," 2nd Edition, Pearson Education Inc., October 2006.

[16] B. A. Forouzan, "Data Communications and Networking," Fourth Edition, McGraw-Hill, 2007.

[17] H. Holma and A. Toskala, "WCDMA for UMTS, Radio Access for Third Generation Mobile Communications," Third Edition, John Wiley & Sons Ltd, 2004.

[18] T. Halonen, J. Romero and J. Melero, "GSM, GPRS and EDGE Performance: Evolution Towards 3G/UMTS," 2nd Edition, John Wiley & Sons Ltd, 2003.

[19] M. Rumney (Editor-in-Chief), "LTE and the Evolution to 4G Wireless – Design and Measurement Challenges," Agilent Technologies Inc., 2009.

[20] F. Khan, "LTE for 4G Mobile Broadband - Air Interface Technologies and Performance," Cambridge University Press, 2009.

[21] H. Holma and A. Toskala, "LTE for UMTS, OFDMA and SC-FDMA Based Radio Access," John Wiley & Sons, 2010.

[22] O. T. Eluwole and M. Lohi, "Coordinated Multipoint Power Consumption Modeling for Energy Efficiency Assessment in LTE/LTE-advanced Cellular Networks," IEEE 19th International Conference on Telecommunications (ICT), Jounieh, Lebanon, April, 2012.

[23] Nokia Networks White Paper, "LTE Release 12 and Beyond," 2014. [Online]. Available: https://resources.ext.nokia.com/asset/200174.

[24] 4G Americas, "Understanding 3GPP Release 12 Standards for HSPA+ and LTE-Advanced Enhancements", February 2015. [Online]. Available: http://www.5gamericas.org/files/6614/2359/0457/4G_Americas_-_3GPP_Release_12_Executive_Summary_-_February_2015.pdf

[25] T. Nakamura, "Proposal for Candidate Radio Interface Technologies for IMT-Advanced Based on LTE Release 10 and Beyond (LTE-Advanced), 3GPP ITU-R WP 5D 3rd Workshop on IMT-Advanced, 15 October 2009.

[26] 3GPP TR 36.913, V8.0.0, Technical Specification Group RAN, "Requirements for Further Advancements for Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced) (Release 8)," June 2008.

[27] 3GPP TR 36.913, V9.0.0, Technical Specification Group RAN, "Requirements for Further Advancements for Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced) (Release 9)," December 2009.

[28] L. Gupta, R. Jain, and H. A. Chan, "Mobile Edge Computing – An Important Ingredient of 5G Networks," IEEE Software Defined Networks Technical Community, Newsletter, March 2016.

[29] O. N. C. Yilmaz, "Ultra-Reliable and Low-Latency (URLLC) 5G Communication," European Conference on Networks and Communications (EuCNC), June 2016.

[30] Ericsson, "5G - key Component of the Networked Society (RWS-150009)," 3GPP RAN Workshop on 5G, Phoenix, AZ, USA, 17 – 18 September 2015.

[31] Samsung Electronics White Paper, "5G Vision," DMC R&D Center, August 2015. [Online]. Available: http://www.samsung.com/global/business-images/insights/2015/Samsung-5G-Vision-2.pdf

[32] Ericsson White Paper, "5G systems – Enabling the Transformation of Industry and Society," UEN 284 23-3251 rev B, January 2017. [Online]. Available: https://www.ericsson.com/assets/local/publications/white-papers/wp-5g-systems.pdf

[33] Ericsson White Paper, "5G Radio Access – Capabilities and Technologies," Uen 284 23-3204 Rev C, April 2016. [Online]. Available:https://www.ericsson.com/assets/local/publications/white-papers/wp-5g.pdf

[34] Nokia Networks, "Ten key rules of 5G deployment, Enabling 1 Tbit/s/km2 in 2030," White Paper, May 2015. [Online]. Available: https://networks.nokia.com/file/39891/ten-key-rules-of-5g-deployment

[35] P. Sorrells, "5G: Advancing the world of connectivity," CommScope White Paper, September 2016. [Online]. Available: http://www.commscope.com/5g/wp-advancing-the-world-of-connectivity/

[36] M. Maier, M. Chowdhury, B. P. Rimal, and D. P. Van, "The Tactile Internet: Vision, Recent Progress, and Open Challenges," IEEE Communications, Volume 54, Issue 5, pp. 138 – 145, May 2016.

[37] N. Kundargi, "5G Standardization: 3GPP Status Update and Overview," National Instruments. [Online]. Available: ni.com/5g

[38] G. Brown, "Designing 5G-Ready Mobile Core Networks," Heavy Reading White Paper, September 2016. [Online]. Available: https://www.affirmednetworks.com/wp-content/themes/renden/pdf/5G_Whitepaper_Heavy_Reading.pdf

[39] G. Brown, "Service-Oriented 5G Core Networks," Heavy Reading White Paper. [Online]. Available:

http://carrier.huawei.com/~/media/CNBG/Downloads/track/HeavyReadingWhitepaperServiceOriented5GCoreNetworks.pdf
[40] J. Kima, D. Kimb, and S. Choi, "3GPP SA2 architecture and functions for 5G mobile communication system," The Korean Institute of Communications Information Sciences, Publishing Services by Elsevier B.V, ICT Express, Volume 3, Issue 1, pp. 1–8, March 2017.
[41] 3GPP TR 38.804 V14.0.0, "Technical Specification Group Radio Access Network; Study on New Radio Access Technology; Radio Interface Protocol Aspects (Release 14)," March 2017.
[42] 3GPP TR 38.801 V14.0.0, "Technical Specification Group Radio Access Network; Study on new radio access technology: Radio access architecture and interfaces (Release 14)," March 2017.
[43] 3GPP TS 23.501 V1.4.0, "Technical Specification Group Services and System Aspects; System Architecture for the 5G System; Stage 2 (Release 15)," September 2017.
[44] 3GPP TR 23.799 V14.0.0, "Technical Specification Group Services and System Aspects; Study on Architecture for Next Generation System (Release 14)," December 2016.
[45] 3GPP TR 22.891 V14.2.0, "Technical Specification Group Services and System Aspects; Feasibility Study on New Services and Markets Technology Enablers (SMARTER); Stage 1 (Release 14)," September 2016.
[46] Qualinet, "White Paper on Definitions of Quality of Experience," Output from the fifth Qualinet meeting, Version 1.2, Novi Sad, March 12, 2013.
[47] ITU-T Rec. P.10, "Vocabulary for performance and quality of service, Amendment 2: New definitions for inclusion in Recommendation ITU-T P.10/G.100," Int. Telecomm. Union, Geneva, 2008.
[48] E. Liotou, N. Passas and L. Merakos, "The Emergence of Experience Packages in the 5G Era," IEEE 5G Tech Focus, Volume 1, Number 3, September 2017.
[49] T. Shimojo, A. Umesh, D. Fujishima and A. Minokuchi, "Future Core Network for the 5G Era," NTT DOCOMO Technical Journal, vol. 17, no. 4, pp. 50 – 59, 2016. [Online]. Available: https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/rd/technical_journal/bn/vol17_4/vol17_4_006en.pdf
[50] R. Kirichek, "5G and Tactile Internet. Network and Services Decentralization," ITU Workshop on Innovative ICT Technologies, Tashkent, Uzbekistan, 21-22 September 2016.
[51] TechTarget, "augmented reality (AR)". [Online]. Available: http://whatis.techtarget.com/definition/augmented-reality-AR
[52] Virtual Reality Society (VRS), "What is Virtual Reality". [Online]. Available: https://www.vrs.org.uk/virtual-reality/what-is-virtual-reality.html
[53] M. Dohler and K. O'Brien, "The Internet of Things," online course, FutureLearn, 2016.
[54] D. D. Coleman and D. A. Westcott, "Certified Wireless Network Administrator, CWNA Study Guide," Wiley Inc., 2006.
[55] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for Next Generation Wireless systems," IEEE Communications Magazine, Volume 52, Issue 2, pp. 186–195, February 2014.
[56] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," IEEE Journal of Selected Topics in Signal Processing, Volume 8, Issue 5, pp. 742 - 758, October 2014.
[57] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, M. Tanno, "Coordinated Multipoint Transmission/Reception Techniques for LTE-Advanced [Coordinated and Distributed MIMO]," IEEE Wireless Communications, Volume 17, Issue 3, pp. 26–34, June 2010.
[58] Anritsu, "5G Standardization Status in 3GPP," No. 91, March 2016. [Online]. Available: http://dl.cdn-anritsu.com/ja-jp/test-measurement/reffiles/About-Anritsu/R_D/Technical/91/91-02-5g-2.pdf
[59] ITU-R SG05 Contribution 40, "Draft new Report ITU-R M.[IMT-2020.TECH PERF REQ] - Minimum requirements related to technical performance for IMT-2020 radio interface(s)," February 2017. [Online]. Available: https://www.itu.int/md/R15-SG05-C-0040/en
[60] BBC Technology News, "5G researchers manage record connection speed," 25 February 2015. [Online]. Available: http://www.bbc.com/news/technology-31622297
[61] I. Grigorik, "High Performance Browser Networking," O'Reilly Media, Inc., 2013.
[62] A. Aijaz, M. Dohler, A. H. Aghvami, V. Friderikos, and M. Frodigh, "Realizing the Tactile Internet: Haptic Communications over Next Generation 5G Cellular Networks," IEEE Wireless Communications, Volume 24, Issue 2, pp. 82 – 89, April 2017.
[63] P. Popovski, "Ultra-Reliable Communication in 5G Wireless Systems," IEEE 1st International Conference on 5G for Ubiquitous Connectivity (5GU), pp. 146 – 151, November2014.
[64] M. Dohler, "The Tactile Internet, IoT, 5G and Cloud on Steroids," IET 5G Radio Technology Seminar, Exploring Technical Challenges in the Emerging 5G Ecosystem, London, UK, March 2015.
[65] ITU-T Technology Watch Report, "The Tactile Internet," ITU, August 2014. [Online]. Available: https://www.itu.int/dms_pub/itu-t/oth/23/01/T23010000230001PDFE.pdf
[66] IEEE Talks 5G, "Three Questions for Gerhard Fettweis," January 2017. [Online]. Available: https://5g.ieee.org/publications/ieee-talks-5g/50-ieee-talks-5g-three-questions-for-gerhard-fettweis
[67] M. Dohler, "5G Tactile Internet – Lab Experimentation @ King's," UK 5G Testbed Meeting, London, UK, 5 July 2016. [Online]. Available: https://www.commnet.ac.uk/wp-content/uploads/2016/08/05Jul16-KCL-Mischa-Dohler.pdf
[68] G. P. Fettweis, "The Tactile Internet: Applications and Challenges," IEEE Vehicular Technology Magazine, Volume 9, Issue 1, pp. 64 – 70, March 2014.
[69] A. B. Grant and O. T. Eluwole, "Cloud Resource Management – Virtual Machines Competing for Limited Resources," IEEE 55th International Symposium ELMAR, Zadar, Croatia, pp. 269–274, September, 2013.
[70] M. H. Ghahramani, M. C. Zhou, and C. T. Hon, "Toward Cloud Computing QoS Architecture: Analysis of Cloud Systems and Cloud Services," IEEE/CAA Journal of Automatica Sinica, vol. 4, no. 1, pp. 6–18, January 2017.
[71] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," National Institute of Standards and Technology, Information Technology Laboratory, Version 15, Oct. 7, 2009.
[72] A. Manzalini, C. Buyukkoc, P. Chemouil, S. Kuklinski, F. Callegati, A. Galis, M. Odini, C. Lin I, J. Huang, M. Bursell, N. Crespi, E. Healy, and S. Sharrock, "Towards 5G Software-Defined Ecosystems: Technical Challenges, Business Sustainability and Policy Issues," IEEE SDN Initiative White Paper, IEEE Future Directions, July 2016.
[73] K. Pretz, "The "Softwarization" of Telecommunications Systems," The Institute, The IEEE news source, 26 July 2016.
[74] A. Manzalini and Noel Crespi, "SDN and NFV for Network Cloud Computing: A Universal Operating System for SD Infrastructures," IEEE Fourth Symposium on Network Cloud Computing and Applications (NCCA), 11 – 12 June 2015.
[75] L. Gorrati, "The 5G Infrastructure Public private partnership (5G-PPP) Project SESAME: Bringing Virtualization, Control and Intelligence to the Network Edge". [Online]. Available: http://www.charisma5g.eu/wp-content/uploads/2016/07/5G-PPP-Project-SESAME.Bringing-Virtualization-Control-and-Intelligence-to-the-Network-Edge.pdf
[76] X. An, R. Trivisonno, H. Einsiedler, D. V. Hugo, K. Haensge, X. Huang, Q. Shen, D. Corujo, K. Mahmood, D. Trossen, M. Liebsch, F. Leitao, C. Phan and F. Klamm, "End-to-End Architecture Modularisation and Slicing for Next Generation Networks". [Online]. Available: https://arxiv.org/ftp/arxiv/papers/1611/1611.00566.pdf
[77] N. Sprecher, "Mobile Edge Computing, a key technology towards 5G," 5G World, London, UK, 2016. [Online]. Available: http://www.etsi.org/component/rsfiles/download-file/files?path=MEC+Presentations+at+Industry+Events%5C201606+-+5G+World+-+ETSI+Mobile+Edge+Computing.pdf
[78] iGillottResearch (iGR) Inc., "Moving to Cloud RAN and the Potential of Edge Data Centers," White Paper, Version 1.0, October 2016. [Online]. Available: www.commscope.com/Docs/iGR_cRAN_White_Paper.pdf
[79] D. Jones (Mobile Editor), "What is [Bleep] Fronthaul?". [Online]. Available: http://www.lightreading.com/what-the--bleep--is-fronthaul/a/d-id/707868
[80] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!," IEEE Access, Volume 1, pp. 335 – 349, May 2013.
[81] A. Bader and M. Alouini, "5G "On The Fly"," IEEE 5G Tech Focus, Volume 1, Number 3, September 2017.
[82] IGI Global, "What is Smart Technology," Online dictionary. [Online]. Available: https://www.igi-global.com/dictionary/smart-technology/38186

[83] W. Derzko, “Smart Technologies,” World Future Society, 29 July 2006. [Online]. Available: http://archives.ocediscovery.com/2007/presentations/Session3WalterDrezkoFINAL.pdf
[84] TechTarget, “machine-to-machine (M2M)”. [Online]. Available: http://internetofthingsagenda.techtarget.com/definition/machine-to-machine-M2M
[85] C. Wang, “M2M Service Architecture: Delivering M2M Services Over Heterogeneous Networks,” IEEE Communications Quality & Reliability 2012 International Workshop, May 17, 2012.
[86] Z. Abbas and W. Yoon, “A Survey on Energy Conserving Mechanisms for the Internet of Things: Wireless Networking Aspects,” Sensors Journal, no. 15, pp. 24818 – 24847, September 2015.
[87] M. Weiser, “The Computer for the 21st Century,” Scientific American Journal, September 1991.
[88] Postscapes, “Internet of Things (IoT) History”. [Online]. Available:https://www.postscapes.com/internet-of-things-history/
[89] K. D. Foote, “A Brief History of the Internet of Things,” DATAVERSITY Education LLC, 16 August 2016. [Online]. Available: http://www.dataversity.net/brief-history-internet-things/
[90] K. Ashton, “That 'Internet of Things' Thing,” RFID Journal. [Online]. Available: http://www.rfidjournal.com/articles/view?4986
[91] TechTarget, “IoT analytics guide: Understanding Internet of Things data”. [Online]. Available: http://searchbusinessanalytics.techtarget.com/essentialguide/IoT-analytics-guide-Understanding-Internet-of-Things-data
[92] SRI Consulting Business Intelligence (SRIC BI), Conference Report CR 2008-07, “Disruptive civil technologies: Six technologies with potential impacts on US interests out to 2025,” April 2008. [Online]. Available: https://fas.org/irp/nic/disruptive.pdf
[93] O. T. Eluwole, N. S. Udoh and O. D. Ojo, “The Impact of Internet on African Education and Culture,” International Journal of Business, Humanities and Technology (IJBHT), vol. 4, no. 3, pp. 69-77, May 2014.
[94] K. Mengersen (Lead Educator), “Big Data: Statistical Inference and Machine Learning”, online course, FutureLearn, 2016.
[95] L. Xu, C. Jiang, J. Wang, J. Yuan, and Y. Ren, “Information Security in Big Data: Privacy and Data Mining”, IEEE Access, Volume 2, pp. 1149 – 1176, October 2014.
[96] Tableau Software, “Top Ten Big Data Trends for 2017”, White Paper, 2017. [Online]. Available: https://www.tableau.com/learn/whitepapers
[97] V. Krishnan, “Solving the Internet of Things’ Last-Mile Problem”, Tableau Software White Paper, 2017. [Online]. Available: https://www.tableau.com/learn/whitepapers
[98] The IEEE 5G Initiative, “IEEE 5G and Beyond Technology Roadmap White Paper”, 2017. [Online]. Available: https://5g.ieee.org/images/files/pdf/ieee-5g-roadmap-white-paper.pdf
[99] O. T. Eluwole, M. Lohi and N. S. Udoh, “Smart Card Technology: Contactless Payment System and Near Field Communication,” SDPS 19th International Conference on Transformative Research in Science and Engineering, Business and Social Innovation, Sarawak, Malaysia, pp. 19-26, June 2014.
[100]N. Patwari, “Wireless Communication Systems,” Department of Electrical and Computer Engineering, University of Utah, Fall 2011. [Online]. Available: http://www.ece.utah.edu/~npatwari/pubs/lectureAll_ece5325_6325_f11.pdf
[101]A. Manikas, “Communication theory,” Department of Electrical and Electronic Engineering, Imperial College of Science, Technology and Medicine, November 2001. [Online]. Available: http://skynet.ee.ic.ac.uk/notes2007/CS%20Part%203(1)%202001%20-%20Mobile%20Communication%20Systems%20(GSM).pdf
[102]M. Ojo, D. Adami and S. Giordano, "A SDN-IoT Architecture with NFV Implementation," 2016 IEEE Globecom Workshops (GC Wkshps), Washington, DC, USA, 2016.
[103] M. Bellanger, et al. "FBMC physical layer: a primer." PHYDYAS, January 25.4 (2010): 7-10.
[104] L. Dai, et al. "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends." IEEE Communications Magazine 53.9 (2015): 74-81.
[105]C. Sun, et al. "Beam division multiple access transmission for massive MIMO communications." IEEE Transactions on Communications 63.6 (2015): 2170-2184.
[106]F. Schaich and W. Thorsten "Waveform contenders for 5G—OFDM vs. FBMC vs. UFMC." Communications, Control and Signal Processing (ISCCSP), 2014 6th International Symposium on. IEEE, 2014.
[107] N. Bhushan, et al. "Network densification: the dominant theme for wireless evolution into 5G." IEEE Communications Magazine 52.2 (2014): 82-89.
[108] J. Liu, et al. "Network densification in 5G: From the short-range communications perspective." IEEE Communications Magazine 55.12 (2017): 96-102.
[109] L. Yang, and Z. Wei. Interference Coordination for 5G Cellular Networks. Springer, 2015.
[110] W. Nam, et al. "Advanced interference management for 5G cellular networks." IEEE Communications Magazine52.5 (2014): 52-60.
[111] M. O. Ojo and O. O. Aramide. "Various interference models for multicellular scenarios: A comparative study." Digital Information and Communication Technology and its Applications (DICTAP), 2015 Fifth International Conference on. IEEE, 2015.
[112] Z. Ding, et al. "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users." IEEE Signal Processing Letters 21.12 (2014): 1501-1505.
[113] R. Baldemair, et al. "Ultra-dense networks in millimeter-wave frequencies." IEEE Communications Magazine 53.1 (2015): 202-208.
[114] Y. Mehmood, et al. "M2M communications in 5G: state-of-the-art architecture, recent advances, and research challenges." IEEE Communications Magazine 55.9 (2017): 194-201.
[115] E. Yaacoub, and D. Zaher. "On using relays with carrier aggregation for planning 5G networks supporting M2M traffic." Wireless and Mobile Computing, Networking and Communications (WiMob), 2014 IEEE 10th International Conference on. IEEE, 2014.
[116] R. Ratasuk, et al. "Recent advancements in M2M communications in 4G networks and evolution towards 5G." Intelligence in Next Generation Networks (ICIN), 2015 18th International Conference on. IEEE, 2015.
[117] S. Guirguis, et al. "M2M communications: Enablement in 4G LTE, deployment considerations and evolution path to 5G." Wireless Networks and Mobile Communications (WINCOM), 2017 International Conference on. IEEE, 2017.
[118] Y. Zhang, et al. "Cognitive machine-to-machine communications: visions and potentials for the smart grid." IEEE network 26.3 (2012).
[119] A. Asadi, Q. Wang and V. Mancuso. "A survey on device-to-device communication in cellular networks." IEEE Communications Surveys & Tutorials 16.4 (2014): 1801-1819.
[120] R. I. Ansari, et al. "5G D2D Networks: Techniques, Challenges, and Future Prospects." IEEE Systems Journal (2017).
[121] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu. "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions." IEEE Communications Magazine 52.5 (2014): 86-92.
[122] X. Cheng, et al. "Device-to-device channel measurements and models: a survey." IET communications 9.3 (2015): 312-325.
[123] X. Lin, et al. "An overview of 3GPP device-to-device proximity services." IEEE Communications Magazine 52.4 (2014): 40-48.
[124] U. Abbasi and H. Elbiaze. "Multimedia streaming using D2D in 5G ultra dense networks." Consumer Communications & Networking Conference (CCNC), 2018 15th IEEE Annual. IEEE, 2018.
[125]A. V. Bastos, C. M. Silva and D. C. da Silva Jr. "Assisted Routing Algorithm for D2D Communication in 5G Wireless Networks."
[126]O. N. C. Yilmaz et al., “Smart mobility management for D2D communications in 5G networks,” in Proc. IEEE Wireless Commun. Netw. Conf. Workshops (WCNCW), 2014, pp. 219–223.
[127]Z. Gao, et al. "MmWave massive-MIMO-based wireless backhaul for the 5G ultra-dense network." *IEEE Wireless Communications* 22.5 (2015): 13-21.
[128]H. Q. Ngo, E. Larsson, and T. Marzetta, “Energy and spectral efficiency of very large multiuser MIMO systems,” *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013
[129]E. Bjornson, E. G. Larsson, and M. Debbah, “Optimizing multicell massive MIMO for spectral efficiency: How many users should be scheduled?” in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, 2014, pp. 612–616.
[130]T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
[131]F. Rusek *et al.*, “Scaling up MIMO: Opportunities and challenges with very large arrays,” *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013

[132]J. Hoydis *et al.*, “Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?” *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
[133]K. E. Skouby and P. Lynggaard. "Smart home and smart city solutions enabled by 5G, IoT, AAI and CoT services." *Contemporary Computing and Informatics (IC3I), 2014 International Conference on*. IEEE, 2014.
[134]S. Haykin "Cognitive radio: brain-empowered wireless communications." *IEEE Journal on Selected Areas in Communications* 23.2 (2005): 201-220.