ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΔΙΑΤΜΗΜΑΤΙΚΟ ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΠΛΗΡΟΦΟΡΙΚΗ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗ ΒΙΟΙΑΤΡΙΚΗ

ΚΑΤΕΥΘΥΝΣΗ ΠΛΗΡΟΦΟΡΙΚΗΣ

«ΠΛΗΡΟΦΟΡΙΚΗ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗΝ ΑΣΦΑΛΕΙΑ, ΔΙΑΧΕΙΡΙΣΗ ΜΕΓΑΛΟΥ ΟΓΚΟΥ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΠΡΟΣΟΜΟΙΩΣΗ»

Μελέτη και Έρευνα για Δίκτυα 5ης Γενιάς

Νικόλαος Γουδρουμανίδης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Επιβλέπων

Σταμούλης Γεώργιος

Λαμία, 2017

«Υπεύθυνη Δήλωση μη λογοκλοπής και ανάληψης προσωπικής ευθύνης»

Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, και γνωρίζοντας τις συνέπειες της λογοκλοπής, δηλώνω υπεύθυνα και ενυπογράφως ότι η παρούσα εργασία με τίτλο "**A Survey of 5G Networks"** αποτελεί προϊόν αυστηρά προσωπικής εργασίας και όλες οι πηγές από τις οποίες χρησιμοποίησα δεδομένα, ιδέες, φράσεις, προτάσεις ή λέξεις, είτε επακριβώς (όπως υπάρχουν στο πρωτότυπο ή μεταφρασμένες) είτε με παράφραση, έχουν δηλωθεί κατάλληλα και ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή. Αναλαμβάνω πλήρως, ατομικά και προσωπικά, όλες τις νομικές και διοικητικές συνέπειες που δύναται να προκύψουν στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δεν μου ανήκει διότι είναι προϊόν λογοκλοπής.

Ο ΔΗΛΩΝ

Νικόλαος Γουδρουμανίδης

19 Μαρτίου 2017

Τριμελής Επιτροπή:

Σταμούλης Γεώργιος

Λουκόπουλος Αθανάσιος

Ευμορφόπουλος Νέστορας

Επιστημονικός Σύμβουλος:

Τζιρίτας Νικόλαος

A Survey of 5G Networks

Nikolaos Goudroumanidis

I| INTRODUCTION

Information technologies have become an integral part of our society, having a profound socio-economic impact and enriching our daily lives with a plethora of services from media entertainment (e.g. video) to more sensitive and safety-critical applications (e.g. e-commerce, e-Health, first responder services, etc.). If analysts' prognostications are correct, just about every physical object we see (e.g. clothes, cars, trains, etc.) will also be connected to the networks by the end of the decade (Internet of Things). Also, according to a Cisco forecast of the use of IP (Internet Protocol) networks by 2017, Internet traffic is evolving into a more dynamic traffic pattern. The global IP traffic will correspond to 41 million DVDs per hour in 2017 and video communication will continue to be in the range of 80 to 90% of total IP traffic. This market forecast will surely spur the growth in mobile traffic with current predictions suggesting a 1000x increase over the next decade.

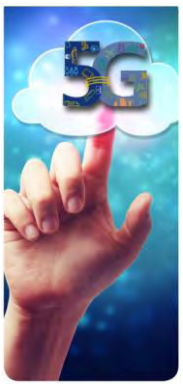
On the other hand, energy consumption represents in today's network a key source of expenditure for operators that will reach alarming levels with the increase in mobile traffic, as well as a factor that is widely expected to diminish market penetration for next-generation handsets as they become more sophisticated and power hungry.

These two attributes in synergy have urged operators to rethink the way they design, deploy and manage their networks in order to take significant steps towards reducing their capital and operating expenditures (Capex and Opex) in next-generation mobile networks – what is generally referred to as 5G, or more specifically 5G mobile.

In order to be ready for the 5G challenge, key mobile stakeholders are already preparing the 5G roadmap that encompasses a broad vision and envisages design targets that include:

- 10–100x peak-rate data rate
- 1000x network capacity
- 10x energy efficiency and
- 10–30x lower latency paving the way towards Gigabit wireless

Early prominent scenarios are starting to emerge, where industrial stakeholders are proposing disruptive ideas towards shifting the market to their customer base and expertise. All the ideas are promising and could play a paramount role in the deployment of 5G mobile networks,



with many of these concepts generated through white papers, international research efforts and technology fora. However, the work reported so far is fragmented and lacks cohesion, based on evolving specific scientific and technology strands such as small cells, network coding or even cloud networking, to name a few. Metaphorically, these works can be perceived as pieces of the 5G jigsaw but, without a holistic perspective in place, it becomes difficult to 'envisage and build the jigsaw' without adopting an interdisciplinary design approach, it becomes even more difficult to even fit two pieces together. It is clear that without a concerted view on the fundamentals of 5G, we will end up building a system that is sporadic and disjointed, providing incremental improvement at best. So, what are the fundamentals of 5G? Well, in essence, if we abstract the technological details, these are the basic building blocks or axioms on which we can build to evolve incremental improvements and represent the most basic platform on which to deliver new services and applications. While building upon 4G systems, in the most basic sense, 5G is an evolution considered to be the convergence of Internet services with legacy mobile networking standards leading to what is commonly referred to as the 'mobile Internet' over Heterogeneous Networks (HetNets), with very high-speed broadband. Green communications also seem to play a pivotal role in this evolutionary path with key mobile stakeholders driving momentum towards a greener mobile ecosystem through cost-effective design approaches. Therefore, in essence, the scope of 5G is not only the mobile and wireless pieces, but also includes the wide area coverage network; or in other words, the Internet will also play a pivotal role in the fabric of the 5G technology ecosystem. Understanding the Internet today, its limitations and the way forward, will assist us with our interdisciplinary design and place a fence around the 5G mobile system solution space based on the requirements and mechanics of the overlay networks. Indeed, if we can take a step back and take a snapshot of the 'holistic picture' then we are able to nicely design and shape the pieces of our jigsaw, so that they fit together seamlessly, and engineer the system that we had originally intended in the right timeframe.

Cellular networks are undergoing a major shift in their deployment and optimization. New infrastructure elements, such as femto/pico base stations, fixed/mobile

relays, cognitive radios and distributed antennas are being massively deployed, thus making future 5G cellular systems and networks more heterogeneous. In this emerging networking environment, small cells could play a fundamental role for the successful deployment of 5G systems. However, despite the recent popularity of small cells on a smaller scale, there is no single technological advance today that can meet the projected traffic demand for 2020. In fact, today's technology roadmaps depict different blends of spectrum (Hertz), spectral efficiency (bits per Hertz per cell) and small cells (cells per km2) as a stepping stone towards meeting the 5G challenge. Therefore, as we migrate towards the 5G era, with advances in small-cell technologies aggregated with supplementary techniques based on advanced antennas (mmWave and massive MIMO (multiple input multiple-output) among others) Multiple-Input Multiple-Output – MIMO) and additional spectrum, we can potentially arrive at a candidate solution for 5G mobile networks. In the absence of any disruptive technologies, once cell-densification limits are reached, and given no further increases in spectral efficiency levels, wider spectrum and a more efficient utilization of available resources and sharing remains the way forward.

Beyond 4th-Generation (4G) wireless technologies, the introduction of heterogeneous networks (HetNets) shifts the interest towards short- and medium-range communications inside the macro cells, motivating further the concept of node cooperation. To that end, we investigate how cooperation can play a major role in 5G mobile networks towards enhancing link reliability and promoting energy efficiency. We base our idea on the notion of Decode-and-Forward (DF), a traditional cooperative technique that has attracted great interest, especially after the widespread adoption of mechanisms such as Automatic Repeat reQuest (ARQ) and Network Coding (NC), which are facilitated by the DF operation. Although DF has been extensively studied in the literature, the increasing number of wireless devices as we head towards 5G, along with the dense urban environment, triggers the need for a Medium Access Control (MAC) layer that can harness the underlying benefits of cooperation and NC through inter- layer design, without neglecting the physical layer impact.

Driven by consumer demand, an astounding 1000x increase in data traffic is expected in this decade. This sets the stage for enabling 5G technology that delivers fast and cost-effective data connectivity, whilst minimizing the deployment cost. Despite the success of small cells and MIMO in 4G systems, these in synergy have not advanced far enough to meet the projected traffic demand. In fact, the future aims towards the aggregate combination of spectrum, spectral efficiency and small cells to work in synergy to deliver the targeted gains. However, how we can exploit legacy spectrum more effectively, as well as introduce new sources of spectrum to cater for additional traffic demands and scenarios, deserves mention. particularly as we are experiencing an era where spectral resources are at a premium. A number of technologies and techniques have been identified as enablers for exploiting clean and new spectrum opportunities in 5G wireless network under the umbrella of cognitive radios (CRs).

5G mobile networks are expected to provide support for ubiquitous mobility, symmetrical and asymmetrical data transmission, broadband connectivity at any time, in any place on any device, but concerningly at the expense of power consumption. In fact, reduced power consumption – or in other words, energy efficiency – will be of paramount importance since future handsets will become increasingly power hungry, leading to hot devices with reduced battery lifetime, affecting the possible market uptake of any new so called '5G i-phone'. Therefore, there is a need both in the infrastructure and, beyond that, in the user terminals, to adopt a more holistic design approach towards harnessing the energy gain from different constituents of the 5G technology ecosystem. These phones, or more likely handset 'devices', will be energy-efficient multi-standard radio transceivers, with common base-band functionality serving several standards and all radio modes integrated onto a reduced chip set.

A| HISTORY OF WIRELESS COMMUNICATIONS

A new generation of cellular system appears every 10 years or so, with the latest generation (4G) being introduced in 2011. Following this trend, the 5G cellular system is expected to be standardized and deployed by the early 2020s. The standardization of the new air interfaces for 5G is expected to gain momentum after the International Telecommunication Union- Radio Communication Sector's (ITU-R) meeting at the next World Radio Communication Conference (WRC), to be held in 2015. Although IMT requirements for 5G are yet to be defined, the common consensus from academic researchers and industry is that in principle it should deliver a fiber-like mobile Internet experience with peak rates of up to 10Gbps in static/low mobility conditions, and 1 Gbps blanket coverage for highly mobile/cell edge users (with speeds of > 300 km/h). The round-trip time latency of the state-of-the-art 4G system (Long-Term Evolution – Advanced; LTE-A) is around 20 ms, which is expected to diminish to less than 1 ms for 5G.

To understand where we want to be in terms of 5G, it is worthwhile to appreciate where it all started and to mark where we are now. The following provides a roadmap of the evolution towards 5G communications:

Before 1G (<1983): All the wireless communications were voice-centric and used analogue systems with single-side-band (SSB) modulation.

1G (1983–): All the wireless communications were voice-centric. In 1966, Bell Labs had made a decision to adopt

analogue systems for a high-capacity mobile system, because at that time the digital radio systems were very expensive to manufacture. An analogue system with FM radios was chosen. In 1983, the US cellular system was named AMPS (Advanced Mobile Phone Service). AMPS were called 1G at the time.

2G (1990–): During this period, all the wireless communications were voice-centric. European GSM and North America IS-54 were digital systems using TDMA multiplexing. Since AT&T was divested in 1980, no research institute like Bell Labs could develop an outstanding 2G system as it did for the 1G system in North America. IS-54 was not a desirable system and was abandoned. Then, GSM was named 2G at the time when 3G was defined by ITU in 1997. Thus, we could say that moving from 1G to 2G means migrating from the analogue system to the digital system.

2.5G (1995–): All the wireless communications are mainly for high-capacity voice with limited data service. The CDMA (code division multiple access) system using 1.25 MHz bandwidth was adopted in the United States. At the same time, European countries enhanced GSM to GPRS and EDGE systems.

3G (1999–): In this generation, the wireless communications platform has voice and data capability. 3G is the first international standard system released from ITU, in contrast to previous generation systems. 3G exploits WCDMA (Wideband Code Division Multiple Access) technology using 5 MHz bandwidth. It operates in both frequency division duplex (FDD) and time division duplex (TDD) modes. Thus, we could say that by migrating from 2G to 3G systems we have evolved from voice-centric systems to data-centric systems.

4G (2013–): 4G is a high-speed data rate plus voice system. There are two 4G systems. The United States has developed the WiMAX (Worldwide Interoperability for Microwave Access) system using orthogonal frequency-division multiplexing (OFDM), evolving from WiFi. The other is the LTE system that was developed after WiMAX. The technology of LTE and that of WiMAX are very similar. The bandwidth of both systems is 20 MHz. The major cellular operators are favorable to LTE, and most countries around the world have already started issuing licenses for 4G using current developed LTE systems. The cost of licensing through auction is very high. Thus, we could say that migrating from 3G to 4G means a shift from low data rates for Internet to high-speed data rates for mobile video.

5G (2021–): 5G is still to be defined officially by standardization bodies. It will be a system of super high-capacity and ultra-high-speed data with new design requirements tailored towards energy elicited systems and reduced operational expenditure for operators. In this context, 5G envisages not only one invented technology, but a technology ecosystem of wireless networks working in synergy to provide a seamless communication medium to the end user. Thus, we can say that moving from 4G to 5G means a shift in design paradigm from a single-discipline system to a multi-discipline system.

B| EVOLUTION OF LONG-TERM EVOLUTION (LTE)

A summary of IMT-Advanced requirements for 4G is as follows:

- Peak data rate of 100 Mbps for high mobility (up to 360 km/h) and 1 Gbps for stationary or pedestrian users.
- User-plane latency of less than 10 ms (single-way UL/DL (uplink/downlink) delay).
- Scalable bandwidth up to 40 MHz, extendable to 100 MHz.
- Downlink peak spectral efficiency (SE) of 15 bit/s/Hz.
- Uplink peak SE of 6.75 bit/s/Hz.

Paving the way to 5G entails both evolutionary and revolutionary system design. While disruptive radio access technologies (RATs) are needed to provide a step up to the next level of performance capability, we also need to improve the existing RATs. In this regard, we need to further improve the LTE system to beyond 4G (B4G). First targeting the IMT-Advanced requirements, LTE standard Release (R)-8 was unable to fulfil the requirements in the downlink direction (although it could meet all the requirements in the uplink direction) with a single antenna element at the User Equipment (UE) and four receive antennas at the Evolved Node B (eNB). In contrast, LTE-A is a true 4G technology (meeting all the IMT-Advanced requirements), requiring at least two antenna elements at the UE. As such, it was accepted as IMT-Advanced 4G technology in November 2010 [1]. **Figure 1** illustrates the evolution of the LTE standard by the 3rd Generation Partnership Project (3GPP) to beyond 4G. The innovations on this roadmap mainly include improving the SE and the area capacity while reducing the network operational cost to ensure fixed marginal cost for the operators.

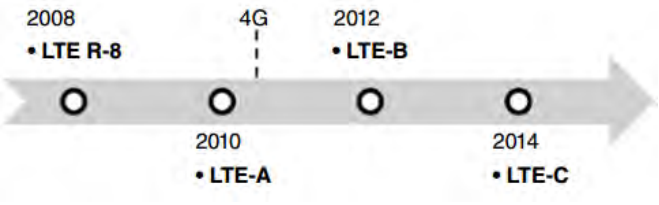


Figure 1: Evolution of LTE standard to beyond 4G

C| WHAT IS 5G?

From analogue through to LTE, each generation of mobile technology has been motivated by the need to meet a requirement identified between that technology and its predecessor. For example, the transition from 2G to 3G was expected to enable mobile internet on consumer devices, but whilst it did add data connectivity, it was not until 3.5G that a giant leap in terms of consumer experience occurred, as the combination of mobile broadband networks and smartphones brought about a significantly enhanced mobile internet experience which has eventually led to the app-centric interface we see today. From email and social media through music and video streaming to controlling your home appliances from anywhere in the world, mobile broadband has brought enormous benefits and has fundamentally changed the lives of many people through services provided both by operators and third party players [2].

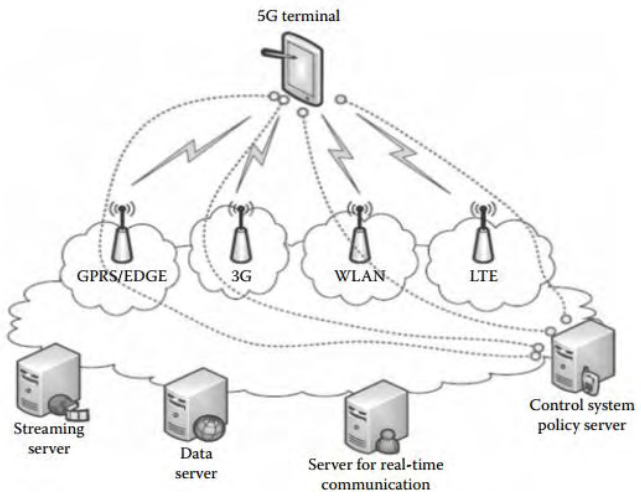


Figure 2: 5G Structure

More recently, the transition from 3.5G to 4G services has offered users access to considerably faster data speeds and lower latency rates, and therefore the way that people access and use the internet on mobile devices continues to change dramatically. Across the world operators are typically reporting that 4G customers consume around double the monthly amount of data of non-4G users, and in some cases three times as much. An increased level of video streaming by customers on 4G networks is often cited by operators as a major contributing factor to this.

5G is envisioned as an end-to-end ecosystem that enables a fully mobile and connected society. 5G will enable a "hyper-connected" world in which the network is highly heterogeneous, converging multiple types of access technologies (both fixed and wireless) across both licensed and unlicensed spectrum, offering unprecedented user experience continuity. Additionally, it will be modular in nature, allowing it to be deployed and scaled on demand to accommodate multiple types of devices, as well as multiple types of user interactions [3].

5G networks are anticipated to be more spectrally efficient than their predecessors, support substantially more users and higher device connection densities, offer higher data rates, and deliver services that are contextual, personalized, responsive and real-time. Additionally, wider network coverage, reduced latency and prolonged battery life of connected devices are also projected.

Consumer and business needs have evolved. The one-size-fits-all model no longer works. Looking back, the first generation of mobile was primarily voice; users paid for the calls they made, and they owned their phones. People used mobile to talk to each other. Today, it's all about bundles aimed to keep users loyal. These include incentives like free voice and SMS in your home network, low-priced devices, application-specific bandwidth deals, data storage, images and video support, as well as self-service portals and cloud-based seamless backup.

People no longer use networks to just talk to each other; now they send each other messages, pictures and links. People share stuff that is bandwidth hungry; they watch movies on the way to work, and they access up to the minute traffic reports before they hit the morning jams.

This is what we are doing today – just five years after data traffic overtook voice in mobile networks. By 2020, it won't be just people and systems creating and sharing data; billions of things will also be an integral part of the communications infrastructure. This infrastructure will encompass cloud computing, virtualization, massive uptake of mobile broadband, more data than ever, more and new devices, and new business models [4].

II| 5G ARCHITECTURE

With the requirements of sub-millisecond latency and bandwidth limitation in traditional wireless spectrum, cellular networks are now poised to break the Base Station (BS) centric network paradigm. Figure 3 depicts this gradual movement from BS centric to a device centric network. The increase in demand by wireless industry motivated the advancement towards much smaller cell deployment from the initial macro hexagonal coverage. Researchers these days are focused on ways to design user centric networking. User is no longer the final resolution of the wireless network but is expected to participate in storage, relaying, content delivery and computation within the network [5].

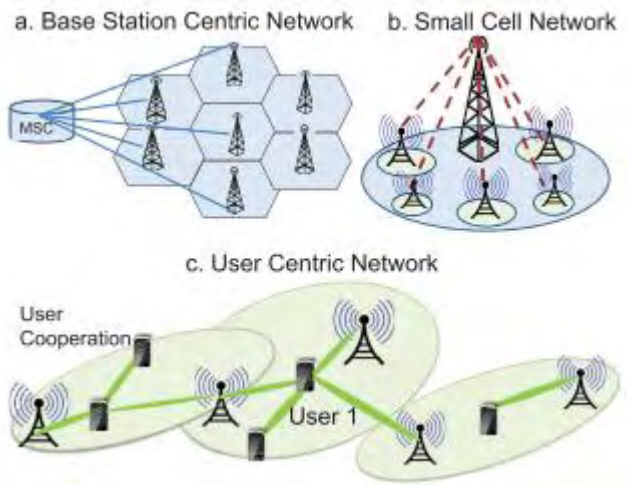


Figure 3: Shift from BS Centric to User Centric Architecture

Future networks are expected to connect diverse nodes in different proximity. Small, micro, pico and femto cell deployment is already underway. Thus, dense 5G networks will have high co-channel interference, which will gradually render the current air interface obsolete. This pushes in the concept of sectorized and directional (energy focused) antennas, as opposed to the age-old omnidirectional antennas. Therefore, Space Division Multiple Access (SDMA) and efficient antenna design are utmost necessary. Decoupling of user and control planes, along with seamless interoperability between various networks are expected to strengthen the foundation for 5G systems. In this section, we discuss the requirements for 5G network architecture, changes in air interface and design of smart antennas. Emerging technologies, like SDN, Cloud-RAN and HetNets are also discussed.

A| RADIO NETWORK EVOLUTION

Overall layout of 5G wireless networks breaks the rules of BS centric cellular concept and moves towards a device centric topology. 5G network proposes the use of higher frequencies for communication. The propagation and penetration of mm-wave signal in outdoor environment is quite limited. Thus, node layout cannot follow traditional cellular design or even any definite pattern. Rappaport and his group propose site specific node layout for 5G radio network design. For instance, ultra-dense deployment is necessary in areas requiring high data rates, like subway stations, malls and offices [6, 7, 8]. Line of Sight (LOS) communication is undisputed preference over Non-Line of Sight (NLOS) communication. Alternately, reflected, scattered and diffracted signals still might have sufficient energy, which needs to be explored when LOS is completely blocked. 5G cellular technology needs to work with an enormous number of users, variety of devices and diverse services. The primary concern therefore, is the integration of 5G BSs with the legacy cellular networks (e.g. 4G, 3G and 2G) [8, 9, 10]. Different configurations like, mm-wave BS grid systems, mm-wave integrated with 4G systems and mm-wave standalone systems are proposed by Farooq and his team at Samsung Electronics. Large beamforming gains extend the coverage, while reducing interference and improving link quality at the cell edges. This feature enables mm-wave BS grids to provide low latency and cost effective solutions. Figure 4 shows a hybrid system of mm-wave (5G) and legacy 4G network. It proposes a dual-mode modem, enabling the user to switch between the two networks for better experience. Alternately, mm-wave spectrum can also be used only for data communications, while control and system information can be transmitted by using traditional 4G networks. On the other hand, as shown in Figure 5, standalone 5G systems operate exclusively on mm-waves. Such systems envision the use of same mm-wave spectrum for both backhaul and wireless access links. The concept of narrow beams allows acceptable spectrum overlap and also improves link quality between BS grids and large number of users [10, 11, 12]. Thus, the radio networking in 5G communications is expected to be much different from legacy networks. Evolution in radio would also change the schematics of the air interface.

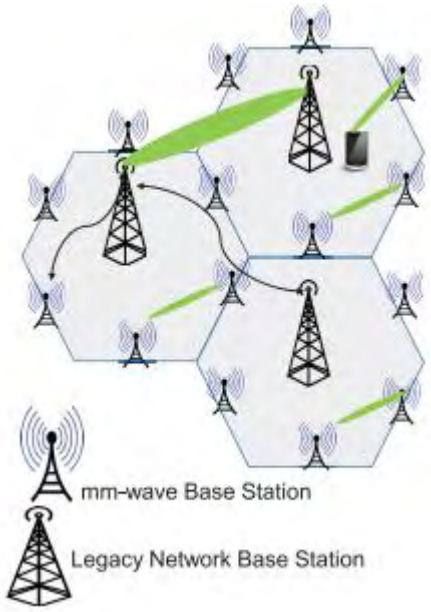


Figure 4: Hybrid Network with both Mm-Wave Small Cell Network

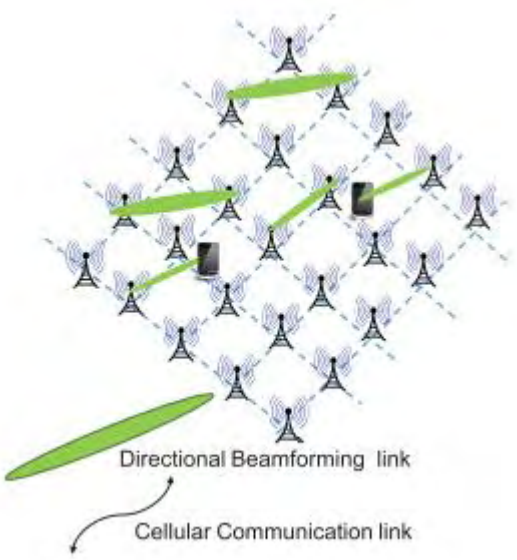


Figure 5: Stand Alone Mm-Wave Small Cell Network

B| ADVANCED AIR INTERFACE

Small radio wave lengths of mm-wave propagation demand small antenna sizes. This enables the use of large number of smaller antennas. Controlling phase and amplitude of signal, using array antennas, helps in enhancing electromagnetic waves in the desired direction, while cancelling in all other directions. This necessitates the introduction of directional air interfaces. Figure 6 shows this change of air interface from omni-directional transmission to a directional one. Highly directional radiation patterns could be secured by using adaptive beamforming techniques, resulting in the introduction of Spatial Division Multiple Access (SDMA). Effective SDMA improves frequency reuse for beamforming antennas at both transmitter and receiver [12, 13, 14]. We defer the details of antenna training, beamforming and SDMA till Section III and Section IV respectively.

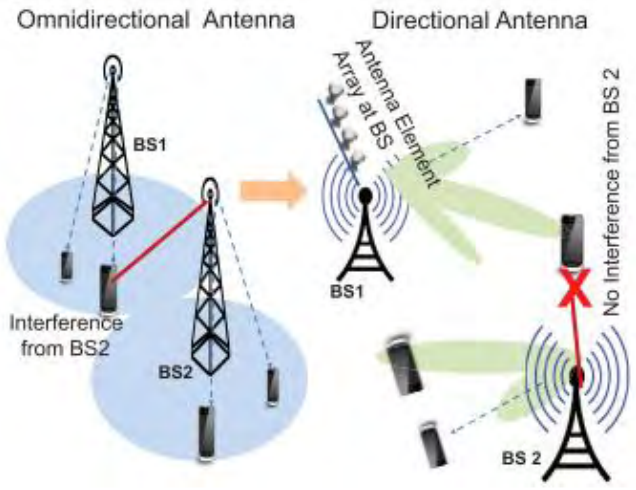


Figure 6: Conventional Omnidirectional Antennas and Smart Beamforming Directional Antennas

However, the hardware challenges, more precisely, the high-power consumption by mixed signal components might constraint these advantages. It might not be possible to connect every antenna to high rate Analog to Digital (A/D) and Digital to Analog (D/A) converters. Hybrid architecture, integrating analog and digital beamforming, with the optimal beamforming weights, can provide possible solutions. Details of analog and digital beamforming forming concepts are discussed in. Sectorization of BS into multiple sectors also relaxes the hardware constraints. However, this raises further challenges in synchronization and data transmission, which needs to be resolved. Optimal antenna configurations for different beamforming techniques enhance performance. For instance, horn antennas at transmitter, patch antennas at receiver and special antenna arrays in high rise urban environment for vertical steering of the beam, would enable efficient communication. Vast BS deployment and need for LOS communication could be eased by the separation of uplink and downlink. Multiple nodes can facilitate different transmissions to use different communication paths at different channel conditions. Understanding of the fundamental techniques of directive air interface along with its advancements would lay strong foundation for the efficient 5G communication [15-19].

C| NEXT GENERATION SMART ANTENNA

Successful deployment of 5G networks depends on the effective antenna array design. This exploits the advantages of change in air interface. The multi-beam smart antenna array system should be used to realize SDMA capabilities. Smart antennas help in interference mitigation, while maintaining the optimal coverage area and transmit power reduction of both mobile handset and BS [16]. Moreover, for the same physical aperture size, more energy can be transmitted at higher frequency by the use of narrow beams [17].

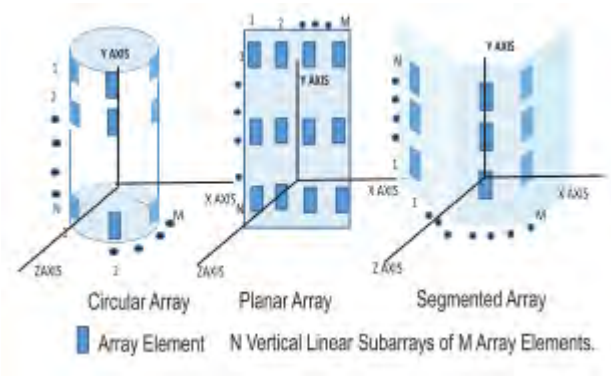


Figure 7: Three Major Array Configurations for Smart Antennas

Smart antenna implementation enables the same channel to be used by different beams. This reduces one of the major problems of wireless communications: co-channel interference. Use of beamforming antennas, with fractional loading factor, further dilutes the co-channel interference problem. Application of highly directional beams do not necessarily require any fractional loading. Infrastructure expenses and complex operations impede indiscriminate use of directional antennas. However, even less complex antennas are capable of providing considerable capacity gains [18, 19]. Therefore, a smart antenna design, optimized over directional gains, cost and complexity is very important for development of 5G wireless communications.

Vertical planar subarrays steer the beams in horizontal plane by varying the weights associated with the subarray elements [19]. The subarray configurations are crucial for beam steering. Figure 7 demonstrates three different possibilities to arrange an antenna subarray: (i) circular, (ii) planar and (iii) segmented. Better coverage of circular subarray makes it more suitable for wireless communications. While curvature allows wider beam steering, linear configurations have better directivity, but limited scan-angle range. Instead of circular or linear, simple segmented configurations can also be carefully designed to achieve the required level of directivity and scan range [19]. Generally, horn antennas have higher gains over all other antennas. An array of horn antennas provides high power output required at BS [20]. The space, size and power are constraints at the mobile device. Hence, more simple patch antennas are suitable candidate for devices. Generally, space, not the size, limits the deployment of sophisticated smart antennas at both BS and MS. However, Samsung's experiments at 28 GHz bands with patch antennas in popular handset have shown promising results.

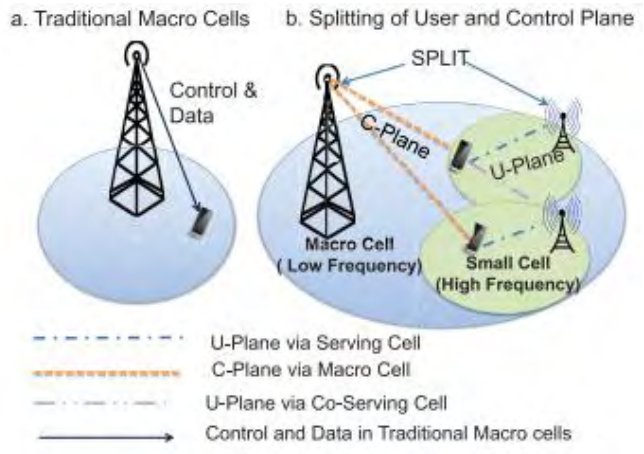


Figure 8: Control Plane and User Plane Separation

D| AGILITY AND RESILIENCE BY SPLITTING OF PLANE – SDN

The changes in architecture and air interface emphasizes on small cells and increased number of antennas. Configuration and maintenance of many servers and routers, in such a dense 5G deployment, is a complex challenge. Software Design Network (SDN) offers a simplified solution for this complex challenge. SDN considers a split between control and data planes, thereby introducing swiftness and flexibility in 5G networks [20, 21]. Figure 8 depicts the segregation of user and control signals. Increase in user plane capacity thus becomes independent of control plane resources. This endows the 5G network with high data at the required locations, without incurring control plane overhead. SDN decouples the data and control planes by using the software components. These software components are responsible for managing the control plane, thereby reducing hardware constraints. Interaction between the two planes is achieved using open interfaces, like Open Flow. It also facilitates switching between different configurations [21, 22, 23 ,24].

SDN can step over OSI layers to remodel networks for a complete automated administration. Redundant interfaces are reduced by controllers, which assign policy to routers for monitoring functions. SDN applied to Radio Access Networks (RAN) presents itself as a SON solution. SON algorithms optimize RAN by control plane coordination at a coarse granularity, while leaving the fine granular data plane unaffected. Although, SON provides high gains, improvement in data plane requires cooperation of multiple BS for data transmissions. Coordinated Multi Point (CoMP) transmission facilitates cooperative data transmission at a very fine time scale. Cloud RAN also offers a viable solution by decentralizing the data plane. Data and control signals can be routed through different nodes, different spectrum and even different technologies to manage the network density and diversity [24].

E| CENTRALISED ARCHITECTURE – CLOUD RAN

Cloud Radio Access Network (C-RAN) resolves some of the major problems associated with increasing demands for high data rates [25]. Wireless industry is working on measures to enhance network capacity by adding more cells, implementing MIMO techniques, establishing complex structure of HetNets and small cell deployment. However, inter-cell interference, CAPital EXpenditure (CAPEX) and OPerating Expenditure (OPEX) impedes these efforts. C-RAN offers to improve system architecture, mobility, coverage performance and energy efficiency while at the same time reducing the cost of network deployment and operation. C-RAN is based on fundamentals of centralization and virtualization. The

baseband resources are pooled at Baseband Unit (BBU), situated at remote central office (not at the cell sites). In traditional cellular networks, the Internet Protocol, Multi-protocol functionality and Ethernet are extended all the way to remote cell sites [26]. Figure 9 shows a typical C-RAN architecture, with BBUs from many remote sites centralized at a virtual BBU pool. This results in statistical multiplexing gains, energy efficient operations and resource savings [25]. Virtual BBU pools further facilitate scalability, cost reduction, integration of different services and reduction in time consumption for field trials. Remote Radio Heads (RRH), comprising of transceiver components, amplifiers and duplexers enable digital processing, analog-digital conversions, power amplification and filtering. RRHs are connected to BBU pool by single mode fiber of data rate higher than 1 Gbps. This simplified BS architecture is paving the way for dense 5G deployment by making it affordable, flexible and efficient. Powerful cloud computing ability can easily handle all complex control processes [23].

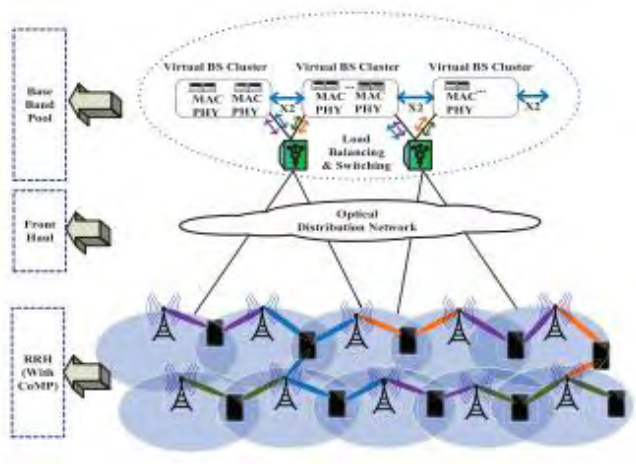


Figure 9: Cloud Radio Access Network (C-RAN) Architecture

China Mobile is strongly advocating C-RAN as it improvise fundamentals for network construction, deployment, cost structure and flexible end user services. Infrastructure sharing protocol, proposed by Mohammad Banikazemi of IBM, provides cost effective solutions for dense deployment, along with backward compatibility. By shifting the RF frontend to BBUs, radio frequencies are generated in the BBU itself. Transmissions are carried out by a shared cloud-radio over fiber infrastructure. This enables the use of analog RF, aiding many services and operators to coexist without any significant interference. Moreover, SDN creates options to seamlessly merge cloud applications with wireless networks through programmable interfaces. Recent researches have proposed SDN based, virtual networks with cloud as a backbone [29]. Small cell deployment can sometimes be difficult, expensive and constraint by site topology. This makes backhaul network for small cells a critical infrastructure. A heterogeneous backhaul technology integrating both, the fixed broadband access and wireless LOS backhaul is the most suitable. Thus, a standardized interface is needed for designing and optimizing RAN along with backhaul network. Dynamic adaptation of routing nodes [30], proposed by Peter Rost of NEC Europe Laboratories, considers RAN as a service (RANas) for flexibility in RAN centralization. RANas concept proposes centralized cloud platform with packaging and delivering functions, depending on the actual network requirements. capacity limits and system-level optimization of uplink C-RAN are studied under practical finite-capacity backhaul constraints in [31]. Compress-and-forward relay strategy is considered for transmitting compressed version of received signals from BSs to central processor. Level of quantization noise introduced by the compression is considered key parameter in backhaul design [31]. Cloud computing based radio access encourages shared pool of configurable resources enabling minimal deployment, management and operational efforts.

F| HETEROGENOUS APPROACH – HETNETS

Another way to handle the wireless traffic explosion, expected in 5G communication, is deployment of large number of small cells giving rise to Heterogeneous Networks (HetNets) [32]. HetNets are typically composed of small cells, having low transmission power, besides the legacy macrocells. By deploying low power small BSs, network capacity is improved and the coverage is extended to coverage holes [33, 34]. Moreover, the overlap of all small, pico, femto cells with the existing macro cells, leads to improved and efficient frequency reuse. Figure 10 shows the concept of HetNets. Deployment of HetNets calls for a coordinated operation between traditional macro cells and small cells for mutual interference reduction. Researchers at University of Manitoba, Canada emphasize on multi-tier networks and interference upgradation for 5G communications. Various interference management challenges in 5G hybrid networks are addressed in [35, 36]. The interference between the macro and second tier cells in 5G HetNets is addressed by reverse Time Division Duplex (TDD) protocol in [37]. It facilitates local estimation of both the intra-tier and inter-tier channels. In reverse TDD mode BS is in downlink operation when the Small-Cell Access (SCAs) operate in uplink and vice versa.

Researchers from Qualcomm Technologies Inc. and Samsung Mobile Solutions Lab emphasize on both network and device side interference management techniques. Advanced receiver with capabilities to take advantages of interference signals structure (including modulation constellation, coding scheme, channel and

Figure 10: Coordinating Cells in HetNet Architecture

resource allocation) are considered as key drivers. Inappropriate Radio Access Technology (RAT) can generate unnecessary signaling overhead [60]. To mitigate these issues in multi-RAT, efficient RAT handover decisions and optimized partitioning of common resources are proposed in [38, 39]. Concurrent utilization of multiple RATs improves capacity and connectivity. However, joint use of multiple networks has not received much research attention. Smart coupling between multiple RATs promises, further capacity and coverage improvement in HetNets [40]. In [41], authors introduce various radio resource management schemes for femtocell enabled HetNets. Cross-tier and co-tier interferences are addressed, while maintaining optimized radio resource utilization, fairness and QoS. Various frequency scheduling algorithms and frequency reuse techniques enhance HetNet performance. Spectral resource allocation strategies also present the potential to solve interference problem. Cooperative and distributed radio resource management algorithms for enabling random HetNet deployment are discussed in [42]. The work in [43] proposes optimization of BS and device association in downlink HetNets under proportional fairness criterion. An overall framework of green HetNets, for balancing energy efficiency and spectral efficiency, is provided in [44]. Two-tier heterogeneous network, proposed in [37], promises improved network performance by co-locating Massive MIMO BS and low-power SCAs. While massive MIMO reaps conventional benefits by ensuring outdoor mobile coverage, SCAs equipped with cognitive and cooperative functionalities act as main capacity-drivers for indoor and outdoor low mobility users. However, backhaul represents one of the major bottleneck in dense SCA deployment. In contrast to legacy wired backhaul, SCAs are likely to be connected via an unreliable wireless backhaul infrastructure. Characteristics of error rate, delay, capacity and deployment cost are expected to vary from case to case. Thus, wireless backhaul links provide a viable and economical alternative. For simplified deployment, operation, management and round-the-clock optimization of HetNets, cloud assisted platform is advocated in [32]. Moreover, cloud based intelligent handoff and location management can ensure seamless connectivity in HetNets. Thus, we believe heterogeneous connectivity of small cells is the major building block of the emerging 5G architecture. The directivity and small cell design, together with advances in resource allocation are promising for higher coverage and data rates of 5G communication.

III| SECURITY FOR 5G COMMUNICATIONS

Nowadays, the trend towards a ubiquitous computing environment, as envisioned by [45], has led to mobile networks characterized by continuously increasing demand for high data rates and mobility. The most prominent technology that has emerged to address these issues is 5G mobile and a lot of effort has been put into developing it over the past few years with the vision of it being deployed by 2020 and beyond. 5G communications aim at providing big data bandwidth, infinite capability of networking and extensive signal coverage in order to support a rich range of high-quality personalized services to the end users. Towards this aim, 5G communications will integrate multiple existing advanced technologies with innovative new techniques. However, this integration will

lead to tremendous security challenges in future 5G mobile networks [46].

Particularly, it is expected that a wide spectrum of security issues will be raised in 5G mobile networks due to a number of factors including: (i) the IP-based open architecture of the 5G system, (ii) the diversity of the underlying access network technologies of the 5G system, (iii) the plethora of interconnected communicating devices, which will also be highly mobile and dynamic, (iv) the heterogeneity of device types in terms of their computational, battery power and memory storage capabilities, (v) the open operating systems of devices, and (vi) the fact that the interconnected devices are usually going to be operated by non-professional users in security issues. Consequently, 5G communications systems will have to address more and much stronger threats than the current existing mobile communications systems.

However, despite the fact that the upcoming 5G communications systems will be the target of many known and unknown security threats, it is not clear which threats will be the most serious and which network elements will be targeted most frequently. Since such knowledge is of utmost importance for the provision of guidance in ensuring security for the next generation mobile communications systems, the objective of this chapter is to present the potential security issues and challenges for the upcoming 5G communications systems.

A| SECURITY ISSUES AND CHALLENGES

The most attractive targets for future attackers in the upcoming 5G communications systems will be the *User Equipment*, *access networks*, *mobile operator's core network* and the *external IP networks*. To help understand the future security issues and challenges affecting these 5G system components, we present representative examples of possible threats and attacks specific to these components. To derive these examples, we explore threats and attacks against legacy mobile systems (i.e. 2G/3G/4G) that may affect the upcoming 5G communications systems by exploiting specific features in this new communication platform. For the example attacks, we also discuss potential mitigation techniques derived from the literature, in order to provide a roadmap towards the deployment of more enhanced countermeasures.

B| USER EQUIPMENT

In the 5G communications era, User Equipment (UE), such as powerful smartphones and tablets, will be a very important part of our daily life. Such equipment will provide a wide range of appealing features to enable end users to access a plethora of high-quality personalized services. However, the expected growing popularity of the future UE, combined with the increased data transmission capabilities of 5G networks, the wide adoption of open operating systems and the fact that the future UE will support a large variety of connectivity options (e.g. 2G/3G/4G, IEEE 802.11, Bluetooth) are factors that render the future UE a prime target for cyber-criminals. Apart from the traditional SMS/MMS-based Denial of Service (DoS) attacks, the future UE will also be exposed to more sophisticated attacks originated from mobile malware (e.g. worms, viruses, trojans) which will target both the UE and the 5G cellular network. The open operating systems will allow end users to install applications on their devices, not only from trusted but also from untrusted sources (i.e. third-party markets). Consequently, mobile malware, which will be included in applications made to look like innocent software (e.g. games, utilities), will be downloaded and installed on end user's mobile devices exposing them to many threats. Mobile malware can be designed to enable attackers to exploit the stored personal data on the device or to launch attacks (e.g. DoS attacks) against other entities, such as other UE, the mobile access networks, the mobile operator's core network and other external networks connected to the mobile core network. Hence, compromised future mobile devices will not only be a threat to their users, but also to the whole 5G mobile network serving them [47].

C| MOBILE MALWARE ATTACKS TARGETING UE

As future UE in the 5G era will be a personal device storing everything from phone contacts to banking information and taken almost everywhere by the end user, it will serve as a single gateway to the end user's digital identity and activities. Thus, the UE will be increasingly vulnerable to mobile malware targeting the stored personal and sensitive information, such as bank credentials, SMSs/MMSs, audio/video files, emails, contacts and GPS coordinates, that attackers can exploit and misuse for financial gain. The malicious software will gain unauthorized access to the end user's stored information, collect it and forward it to the owner of the malware through all of the UE's communication channels [48-50].

Additionally, the future UE will be vulnerable to mobile malware causing disruption to normal service operations. To achieve disruption, the installed malicious software may use all available CPU cycles for junk computations leading to huge power consumption that will rapidly cause the depletion of the UE's power source. This attack falls in the category of DoS attacks against UE [48].

However, the above attacks can be also executed by mobile botnets in order to target many mobile end users at the same time and in an automated way. Thus, mobile botnets are expected to be a significant means for

attackers to gain financial benefits on a larger scale in the 5G era.

C| 5G MOBILE BOTNETS

In the 5G communications environment, mobile botnets are expected to be increasingly used by attackers, since future mobile devices will be ideal remote controlled machines due to their specific features. In particular, 5G mobile devices will support different connectivity options and increased uplink bandwidth, and will tend to be always turned on and connected to the Internet. Thus, future attackers will be enabled to deploy mobile botnets for 5G communications networks in many efficient ways [49, 50].

Similar to mobile botnets in legacy mobile networks [47], future mobile botnets for 5G networks will be networks of compromised mobile devices under the control of malicious actors commonly referred to as bot-masters. For example, a centralized 5G mobile botnet, where the compromised mobile devices will be controlled by the attacker through central Command and Control (C&C) servers, is illustrated in Figure 11. This centralized 5G mobile botnet will consist of the following actors [49]:

Bot-master: will be the malicious actor that can access and manage the botnet remotely via the bot-proxy servers (i.e. central C&C servers). The bot-master will be responsible for choosing the mobile devices that will be compromised by malware and turned into bots. Specifically, the bot-master will exploit security vulnerabilities (e.g. operating system and configuration vulnerabilities) of the chosen mobile devices and compromise them. In current mobile botnets, the bot-masters can use similar http techniques to those used by the PC-based botnets, as well as new techniques specific to mobile devices' features, such as SMS messages, in order to distribute their commands. Since 5G UE will support a large variety of connectivity options, it will also be possible for the bot-masters of future 5G mobile botnets to make use of additional techniques in order to command and control their bots.

Bot-proxy servers: will be the means of communication that the bot-master will use to command and control the bots indirectly.

Bots: will be programmed and instructed by the bot-master to perform a variety of malicious activities, such as Distributed Denial of Service (DDoS) attacks against network elements in the mobile network, mass distribution of spam, and the theft and further distribution of sensitive data, as well as installation of malware on other mobile devices.

D| ACCESS NETWORKS

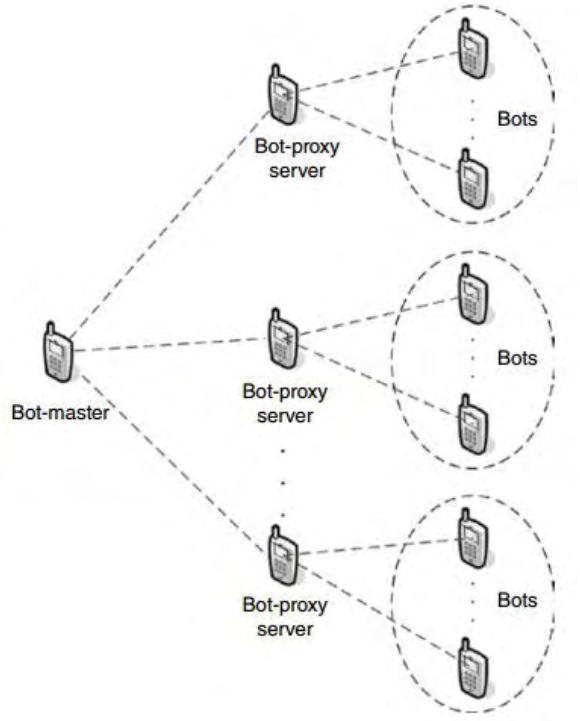


Figure 11: Centralized 5G Mobile Botnet

In 5G communications, access networks are expected to be highly heterogeneous and complex, including multiple different radio access technologies (e.g. 2G/3G/4G) and other advanced access schemes, such as femtocells, so that service availability will be guaranteed. For instance, in the absence of 4G network coverage, the UE should be able to establish a connection over 2G or 3G networks. However, the fact that 5G mobile systems will support many different access networks leads them to inherit all the security issues of the underlying access networks that they will support [51].

During the evolution from 4G communications to 5G communications, enhanced security mechanisms should be implemented to counter emerging security threats on 5G access networks. To address this issue, potential security threats to the future 5G access networks should be firstly identified. Thus, in this section, we focus on existing attacks on current 4G access networks and HeNB femtocells which could be also possible attacks on the 5G access networks.

E| ATTACKS ON 5G NETWORKS

UE Location Tracking: Tracking the UE presence in a specific cell or over multiple cells is a security issue for LTE networks that can seriously affect subscriber's privacy. Two techniques that can be used by attackers to achieve UE location tracking in future 5G access networks are those for LTE networks described in references [52, 53].

They are based on the Cell Radio Network Temporary Identifier (C-RNTI) and the packet sequence numbers.

UE Location Tracking Based on C-RNTI: The C-RNTI provides a unique and temporary UE identification (UEID) at the cell level. It is assigned by the network via an RRC control signal when a UE is associated with the cell. However, the C-RNTI is transmitted in the L1 control signal in plain text. Thus, an adversary is able to determine whether the UE using the given C-RNTI is still in the same cell or not. According to [53], periodic C-RNTI re-allocation is a potential solution. Periodic C-RNTI re-allocation for a UE staying for a long time on the same cell can make it more difficult for an attacker to obtain information related to its presence in the cell. Additionally, it will make it more difficult for the attacker to distinguish if indeed a new UE has arrived at the cell or if it is the same UE that has refreshed its C-RNTI.

Moreover, UE location tracking can be achieved by tracking the combination of the C-RNTI with handover signals. This combination allows UE location tracking across multiple cells. During the handover process, a new C-RNTI is assigned to the UE via the Handover Command message. Thus, in a case where the allocation of C-RNTI itself is not confidentiality protected, an attacker can link the new C-RNTI in the Handover Command message and the old C-RNTI in the L1 control signal. To mitigate this type of attack, encryption of RRC messages, such as the Handover Command message and the Handover Confirm message, is proposed in reference [53]. Encryption of these messages prevents an attacker from associating the RRC messages with a C-RNTI and mapping them together during handover processes.

UE Location Tracking Based on Packet Sequence Numbers: The use of continuous packet sequence numbers for the user plane or control plane packets before and after a handover can enable an attacker to determine the mapping between the old and the new C-RNTIs [52]. UE tracking based on packet sequence numbers can also be applicable to the idle-to-active mode transitions if the sequence numbers are kept continuous. Then, an attacker can track the UE based on the continuous packet sequence numbers of packet streams. To address UE tracking based on sequence numbers, the authors in reference [53] propose that the sequence numbers over the radio should be discontinuous in handover processes and possibly also in the state transitions between idle and active modes. Particularly, they propose the use of a random offset in order to make the user and control plane sequence numbers discontinuous on the radio link. Finally, another solution also proposed in reference [53] is the use of fresh keys for each eNB, which allows setting the sequence number to any random value and thus makes it discontinuous.

Attacks Based on False Buffer Status Reports: In LTE networks, an attacker can exploit the buffer status reports, which are used as input information for packet scheduling, load balancing and admission control algorithms, to achieve his malicious intents. Particularly, the attacker can send false buffer status reports on behalf of the legitimate UE in order to change the behavior of these algorithms on the eNBs and cause serving issues towards the legitimate UE [52, 53].

By changing the behavior of the packet scheduling algorithm, the attacker is able to steal bandwidth. To achieve that, the attacker can make use of C-RNTIs of other legitimate UEs and send false buffer status reports. This can make the eNB think that the legitimate UEs have no data to transmit. Consequently, the packet-scheduling algorithm in the eNB will allocate more resources for the attacker's UE and fewer or no resources for the legitimate UEs, and lead to DoS.

Furthermore, by changing the behavior of load-balancing and admission control algorithms in the eNBs, DoS can be experienced by the new arriving UE in the cell. To achieve that, the attacker can send a wide range of false buffer status reports from various UE claiming that they have more data to send than they actually have. This makes the eNB falsely assume that there is a heavy load in this cell and new arriving UE cannot be accepted. To address the attacks based on false buffer status reports, the use of a one-time access token within the MAC-level buffer status report message is proposed in reference [53]. According to this solution, the UE will have to present this token to the eNB to get the access right. The token is different for each buffer status report sent during a Discontinuous Reception (DRX) period.

Message Insertion Attack: Message insertion attack is another type of attack for LTE networks and is described in references [52, 53]. In LTE networks, the UE is allowed to stay in active mode but turn off its radio transceiver to save power consumption. This is achieved through the DRX period. However, during a long DRX period, the UE is still allowed to transmit packets because the UE may have urgent traffic to send. This feature can potentially cause a security breach. An attacker can inject control protocol data units (C-PDU) into the system during the DRX period to achieve DoS attack against the new arriving UE. According to [53], a solution for mitigating the message insertion attack is the request for capacity through the uplink buffer status report.

F| HeNB FEMTOCELL ATTACKS

The physical size, material quality, lower-cost components and IP interface of HeNB femtocells make them more vulnerable to attacks compared to eNBs. In this subsection, it's presented the main categories of the potential attacks related to HeNB femtocells, with specific examples of attacks for each category. Additionally,

countermeasures for these attacks are discussed. An extensive and detailed list of all possible attacks related to HeNB femtocells and corresponding mitigations can be found in reference [59].

Physical Attacks on HeNB: Physical tampering with HeNB is an attack where a malicious actor can modify or replace HeNB components. With this type of attack, it is possible to affect both end users and mobile operators. For example, modified RF components of an HeNB may interfere with other wireless devices of an eHealth tele-monitoring system in the patient's environment and cause them to malfunction. This can result in health risks for the patient. On the operator's side, an HeNB with modified RF components can impact harmfully on the surrounding macro network. Thus, it is obvious that HeNB should be physically secured in order to prevent easy replacement of its components. In addition, trusted computing techniques should be used to detect when modifications on critical components of an HeNB have occurred. Furthermore, booting HeNBs with maliciously modified software can lead to further security breaches for end users and operators. This can be achieved in HeNBs supporting user-accessible boot code update methods. As a result, eavesdropping on communication and identity fraud are two possible security issues that end users have to address. Also, DoS attacks against the network operators are possible. A mitigation approach is to secure the booting process by using cryptographic means, such as a Trusted Platform Module (TPM) [59].

Attacks on HeNB Credentials: In this category of attacks, the compromise of HeNB authentication credentials is included. According to this attack, an attacker obtains a copy of the authentication credentials from the wires of the targeted HeNB. Then, any malicious device can use them and impersonate the given HeNB. Thus, the attacker can mount masquerade attacks against the end user and the operator. The success of obtaining a copy of the credentials of the targeted HeNB depends on the implementation. Consequently, the credentials should be stored in a protected domain, such as a TPM module, in order not to be compromised easily [59].

Configuration Attacks on HeNB: A possible attack of this category is the mis-configuration of the Access Control List (ACL) of the targeted HeNB. Firstly, the attacker gains access to the ACL, including the Closed Subscriber Group (CSG) list. Then, he modifies the ACL so that illegitimate devices can access the network. In addition, the attacker can modify the ACL to prevent legitimate devices from accessing the network, as well as change the level of access for different devices. As a result, legitimate end users can experience the effects of DoS attacks, and some other malicious end users can make use of services free of charge if the billing is based on the HeNB. Hence, it is essential to ensure secure creation, maintenance and storage of the ACL [59].

Protocol Attacks on HeNB: The protocol attacks category includes man-in-the-middle attacks on HeNB first network access, which can have a very harmful impact on end users. HeNBs are vulnerable to this type of attack when they do not have unique authentication credentials. In these cases, during the first contact of the targeted HeNB to the core network over the Internet, the operator is not able to identify it. Thus, an attacker on the Internet can intercept all traffic originating from the HeNB and get access to private information and exploit it further. To address the man-in-the-middle attacks, authentication credentials should be used by the HeNB in the very first contact with the network. The use of UICC (Universal Integrated Circuit Card) or certificates can be potential solutions towards mitigating these attacks. In UICC-based solutions, the UICC is inserted in the HeNB by the point of sales or the customer, and mutual authentication between the HSS (Home Subscriber Server) and the UICC takes place. On the other hand, in certificate-based solutions, the certificate is stored on the HeNB at the manufacturing phase of the HeNB and used for mutual authentication between the first contact node (i.e. Security Gateway) and the HeNB [59].

Attacks on Mobile Operator's Core Network: DoS attacks can be launched, through malicious traffic originating from compromised HeNBs, against core network elements. Two categories of DoS attacks which can be directed to the core network, but not to the HeNBs, are the following: (i) IKEv2 (Internet Key Exchange Version 2) attacks (e.g. IKE_SA_INIT flood attacks, IKE_AUTH attacks) that can be launched against the initial establishment of the IKEv2 tunnel between the HeNB and the Security Gateway; and (ii) layer 5–7 volume attacks and IKEv2 volume attacks when a high volume of signalling traffic or IKEv2 tunnel setup traffic overwhelms the infrastructure. To mitigate these attacks, the Security Gateway should remain secure and available as first contact point in the core network. Furthermore, this category encompasses HeNB location-based attacks such as the changing of the HeNB location without reporting. A malicious actor may relocate the HeNB and make the provisioned location information invalid. As a result, this can cause emergency calls emanating from the relocated HeNBs not to be reliably located or routed to the correct emergency centres. Furthermore, lawful interception position reporting is impossible. A location locking mechanism is a potential solution to prevent these attacks [59].

User Data and Identity Privacy Attacks: Eavesdropping on another end user's E-UTRAN (Evolved Universal Terrestrial Radio Access Network) user data is a very harmful attack of this category against the privacy of the end user. The attacker installs his own HeNB and configures it to the open access mode. Then, the targeted end user makes use of this malicious HeNB in order to connect to the core network without knowing that this HeNB is compromised. Hence, the attacker is able to eavesdrop on all data flowing

between the targeted end user and the network. This attack exploits the unprotected user traffic in some part of the HeNB. For that reason, unprotected user data should never leave a secure domain inside the HeNB to avoid this eavesdropping attack. Furthermore, the end users should be notified when they are connected to a closed- or an open-type HeNB [59].

Attacks on Radio Resources and Management: Radio resource management tampering is an attack where the HeNB provides incorrect radio resource information. To achieve this, the malicious actor has to get access to the HeNB and modify its resource management aspects. At least, he should be able to modify the power control part of the HeNB. An example of the consequences of this type of attack is increased handover. Thus, the configuration interface of the HeNB should be adequately secured [59].

G| MOBILE OPERATOR'S CORE NETWORK

Due to their IP-based open architecture, 5G mobile systems will be vulnerable to IP attacks that are common over the Internet. DoS attacks, which are a major threat on the Internet today, are going to be present on the future 5G communications systems targeting entities on the mobile operator's core network. However, the 5G mobile operator's core network may be also affected by DDoS attacks targeting external entities, but transferring their malicious traffic over it. Potential attacks include:

DDoS Attacks Targeting the Mobile Operator's Core Network: DDoS attacks will be very serious incidents impacting the availability of the targeted future 5G mobile core network. Since 5G mobile networks are going to be used by millions oφ users, the consequences of DoS and DDoS attacks against the core network will be severe. In the 5G communications environment, DDoS attacks can be launched by a botnet including a large number of infected mobile devices. In this subsection, two representative DDoS attacks against a 4G mobile operator's core network are presented. These two examples of attacks can be also expanded to the 5G core network.

Signaling Amplification: A DDoS attack example for a future 5G mobile operator's core network might be the signaling amplification attack that 4G networks face as described in reference [54]. This attack could be performed by a botnet of multiple infected mobile devices within the same cell in order to deplete the network resources, leading to service degradation. This attack exploits the signaling overhead required to set up and release dedicated radio bearers in LTE networks. Thus, a large number of dedicated bearer requests will be initiated simultaneously, forcing the different network entities to follow the heavy signaling dedicated bearer activation procedure for each bearer. After obtaining the dedicated bearers, the bots will not use them, and after the expiration of the inactive bearer timeout, the bearers will be deactivated following the dedicated bearer deactivation procedure, which incurs heavy signaling as well. Then, the malicious devices of the botnet will execute the same steps over and over again to amplify the attack and degrade the network performance. Finally, the proposed detection technique for this attack is based on features such as the inter-setup time and the number of bearer activations/deactivations per minute. The setting of a lower bound threshold for inter- etup time determines the performance of the detection technique. A high value for the inter-setup time threshold would result in too many false positives. On the other hand, a low value for this threshold might lead to undetected exploits. Furthermore, a high number of bearer activations/deactivations per minute indicates malicious activity and should be discovered and stopped by the operator [51, 54].

HSS Saturation: The HSS is an essential node of the Evolved Packet Core (EPC) since it comprises the master database for a given user and it contains the subscription-related information to support the network entities handling calls/sessions. The HSS also provides support functions in user authentication and access authorization. A Home Network may contain one or more HSSs based on the number of mobile subscribers, the capacity of equipment and the organization of the network [55, 56]. Thus, a DDoS attack against this key node can potentially reduce the availability of the mobile core network significantly.

In reference [57], some research work has already explored the possibility of overloading a Home Location Register (HLR), which is a key component of the HSS, by exploiting a botnet of mobile devices. The results of this research showed that the reduction of the throughput is dependent on the size of the botnet. Moreover, it is worthwhile to mention that in this type of attack, the legitimate users of the infected mobile devices are unlikely to be aware of the occurrence, since these attacks are executed by quietly launching network service requests and not by a flood of phone calls. Finally, according to this research work, basic filtering and shedding are two possible mitigation techniques against such attacks. However, the implementation of mechanisms intelligent enough to respond to more dynamic attacks remains a challenging task. Particularly, it is difficult for a provider to distinguish attacks from other traffic, since a significant amount of context is lost as messages are exchanged between the mobile devices and the HLR (e.g. granularity of location). Furthermore, filtering in the core network may occur too late to prevent legitimate users from experiencing DoS, due to the large overhead related to the first hop of communications in mobile networks [57].

DDoS Attacks Targeting External Entities over a Mobile Operator's Core Network: In future, upcoming 5G mobile networks may also serve as a gateway for DDoS attacks

against targets in other external networks (e.g. enterprise networks) connected to the mobile core network. In this scenario, a botnet of mobile devices may be used to generate a high volume of traffic and transmit it to the victim, located in the external network's infrastructure, over the mobile core network. Although the target of these attacks will not be the core network itself, the fact that they inject large traffic loads into the core network can impact its performance. The recent DDoS attacks against Spamhaus over the Internet proved how the high volume of attack traffic can affect the availability of the underlying communication network employed to transmit it to the specific target [51].

H| EXTERNAL IP NETWORKS

In 5G communications systems, external IP networks can also be the target of DDoS attacks, where mobile botnets generate a high volume of traffic and transmit it to the target over the mobile core network. In addition, external IP networks, such as enterprise networks, can be a soft target for being compromised by malware through infected mobile devices accessing them. In this subsection, we present a representative scenario, based on [58], of how an enterprise network can be compromised through the infected 5G mobile device of an employee. A solution against this threat, proposed in reference [58], is also discussed.

Compromised Enterprise Networks: The current wide adoption of smartphones has already led many employees to bring their own smartphone devices into the work environment and use them to access information assets located in isolated enterprise networks or enterprise networks with strict access control. This trend is expected to continue and accelerate in the upcoming 5G era. However, many security concerns will be raised for the enterprise networks accessed by employees' smartphones due to the potential susceptibilities of smartphones to mobile malware [58]. The potential vulnerabilities can be exploited by attackers to compromise an otherwise secure enterprise network. For example, mobile malware, such as Dream Droid that recently infected the Android Market, may be used by attackers to get unauthorized access to enterprise networks through employees' future smartphones.

Another characteristic of employees' future smartphones that may be exploited by attackers to compromise enterprise networks will be the diversity of their connectivity capabilities. They will support not only mobile communication technologies (2G/3G/4G/5G), but also other connectivity technologies such as WiFi, Bluetooth, NFC (Near Field Communication) and USB (Universal Serial Bus). Thus, the multiple connectivity technologies may be abused by attackers as mobile malware propagation channels. In other words, employees' smartphones may work as bridges for attackers between the enterprise network and the outside world. Thus, an employee's smartphone may be compromised through a mobile communication channel or a short-range communication channel and become a wormhole to the target enterprise network or bring the malicious payload directly to it through another communication channel supported by the smartphone.

In an attack scenario, we consider that the employee's smartphone is connected to a desktop PC through USB and the desktop PC is connected to the internal enterprise network. Then, the bot-master can be connected to a backdoor on the employee's smartphone via WiFi or the 4G mobile network and inject the malicious payload to the internal enterprise network through the USB connection.

To avoid security breaches for the enterprise network arising from the use of employees' smartphones inside the work environment, a very common approach is to periodically scan all employees' smartphones with anti-malware software. However, this approach is intrusive and too costly in terms of energy. Thus, innovative solutions providing a balance between security responsiveness and cost effectiveness are required. In reference [58], strategic sampling is proposed as a method to address this requirement by identifying and periodically sampling the security representative smartphones. Then, the sampled devices are checked for malware infections. Smartphones' security representativeness is measured by the employees' interests and the co-location logs on their devices. The probabilities used in the strategic sampling method are derived from a lottery tree reflecting the smartphones' security representativeness.

IV| EMERGING APPLICATIONS

A wide variety of new emerging applications is the major guiding force behind the commercial roll out of 5G wireless systems. 5G architecture is expected to provide network solutions for a wide range of public and private sectors, like energy, agriculture, city management, health care, manufacturing and transport, with improved software services. Apart from the enormous number of connections, 5G networks also have to support diverse nature of devices and their associated service requirements [60]. Although research and development in some of these applications are already underway in 4G wireless, original 4G LTE standards, 3GPP LTE Release 8.0 [61] did not include support to any of these applications. Rather, these applications were spawned later, and started explosive increase in wireless data usage, thereby imposing additional burden on resource constrained 4G wireless networks. Naturally, later versions of 4G LTE networks, often termed as "LTE Advanced" gradually started to include these applications. On the other hand, it is expected that massive bandwidth of 5G mm-

wave communications will provide a native, de-facto support for these emerging applications. In this section, we present some of these killer applications, like, D2D communications, M2M communications, IoV, IoT and Healthcare.

A| D2D COMMUNICATION

Device centric nature of 5G wireless is expected to enable the devices in proximity to communicate directly bypassing the cellular BS [62] for sharing relevant contents. Figure 12 shows different D2D communication scenarios. A comprehensive review of D2D communications is available in [62]. 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 [63], social networking prototypes (e.g. Qualcomms FlashLinQ) [64], public safety networks and maximum allowable distance estimation for commercial roll out [66].

An adhoc D2D network of 5G wireless devices, using group key agreement and routing control process, is proposed in [67]. Low latency, energy efficiency and scalability are vital to 5G networks. Thus, it is essential to decrease the control signaling and end to end latency in network assisted D2D communications. Nokia Research Center proposes, smart mobility management, D2D-aware handover and D2D triggered handover solutions [68]. These system level improvements can support a reliable vehicle to vehicle communications in 5G wireless. Spectrum sharing, interference management, multi-hop communications and energy efficiency are major challenges in hyper dense 5G mobile environment and require further investigation [69-70].

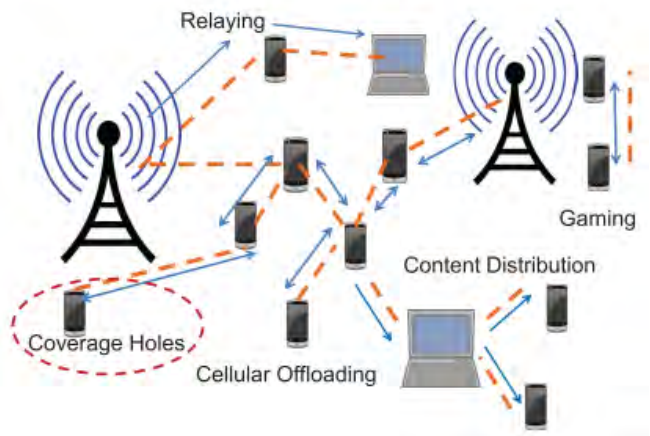


Figure 12: Use Cases of Device to Device Communications

B| M2M COMMUNICATION

Like D2D communications, M2M communications are also expected to have native support in 5G wireless. Major features of M2M communications involve automated data generation, processing, transfer and exchange between intelligent machines, with minimum human intervention [71]. Figure 13 delineates that unlike local D2D communications, M2M communications connect massive number of devices, like smart metering, sensors and smart grid equipment's, along with wide coverage areas.

Figure 13: Use Cases of Device to Device Communications

M2M communications envision umpteen number of devices with small data, sporadic transmissions, high reliability, low latency and real time operation. Major reviews of existing M2M research works include various commercial, hardware and research platforms [72] as well as major architectural enhancements, network functionalities and implementation challenges [73]. We provide a short overview of the major M2M research works relevant to our context. Joint use of carrier aggression and relay station in OFDMA-based 5G wireless is proposed in [74]. Latest advances and developments in architecture, protocols, standards and security for M2M evolution from 4G to 5G wireless are discussed in [75]. Network uncertainty and mobility often lead to complex interference within M2M networks themselves, as well as between M2M networks and cellular networks [76]. We expect the cognitive radio to emerge and assist in developing novel cognitive M2M architecture for sensing and using the available frequency bands. Cognitive Radio technology driven Smart Objects (CRSOs), with high energy efficiency and environmental knowledge are expected to improve M2M communications performance, required for IoT technology [77].

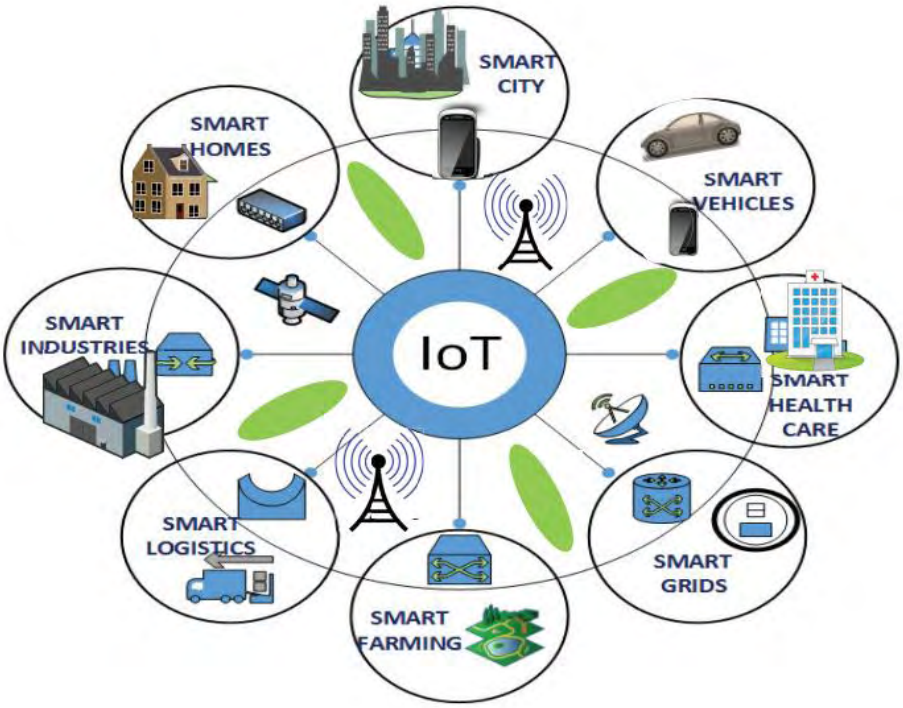


Figure 14: Internet of Things (IoT): Connecting "Anything, Anyone, Anytime, Anyplace"

C| INTERNET OF THINGS (IoT)

As shown in Figure 14, IoT envision millions of simultaneous connections, involving a variety of devices, connected homes, smart grids and smart transportation systems. This vision could be eventually realized only with the advent of high bandwidth 5G wireless networks. IoT enables internet connections and data inter-operability for numerous smart objects and applications [78]. Six unique challenges [140] of IoT include **(i)** Automated sensor configuration, **(ii)** context discovery, **(iii)** acquisition, modeling and reasoning **(iv)** selection of sensors in 'sensing-as-a-service' model [79] (v) security privacy-trust and **(vi)** context sharing. Implementation of IoT is complex, as it includes cooperation among massive, distributed, autonomous and heterogeneous components at various levels of granularity and abstraction. The concept of cloud, offering large storage, computing and networking capabilities, can be integrated with diverse IoT enabled devices [81]. A high-level design of cloud assisted, intelligent, software agent-based IoT architecture is proposed in [80]. Smart objects, enabled with 5G wireless are expected to form the basis of large scale IoT design and roll out. More recently, Social Internet of Things (SIoT) is also coming up for exploring the relationship between objects and form social networks [82]. Concepts, reviews and challenges of SIoT are presented in [83-84]. We expect IoT will gradually transform the current Internet from the human centric interactions to a M2M platform equipped with 5G wireless. This ubiquitous connectivity of autonomously communicating, IoT- enabled devices is the basis of 5G wireless. To advocate IoT on a global scale, ITU-T's IoT Global Standards (IoT-GSI), proposes unified approach for technical standard developments [85].

D| ADVANCED VEHICULAR COMMUNICATIONS

Development in IoT automatically leads to the evolution of Internet of Vehicles (IoV) [86] a network of interconnected vehicles for robust traffic management and reduced collision probabilities. High bandwidth, pervasive availability, and low latency of 5G wireless is assuring smart and intelligent vehicular communications. Emerging vehicular cloud is responsible for all essential services and applications, like content search routing, spectrum sharing and dissemination [87]. IoV involves very huge spatio temporal data (Big Data), which needs to be processed and delivered with high safety and security. IoV is also expected to explore roadside cooperative, as well as non-cooperative relay nodes. Cooperative and non-cooperative Bayesian coalition games, using learning automata, is conducted in IoV for VANET. Vehicles, as smart and interactive social objects, form the basis of Social IoV (SIoV) [88]. SIoV leverage VANETS and develops a vehicular social networking platform, based on cyber physical architecture. Intelligent Internet of Vehicles Management System (IIOVMS), with cloud assisted data processing, over a wide number of vehicles helps in traffic management [89]. Dedicated Short Range Communication Working Group (DSRC), presents IEEE 1609 standards for Wireless Access in Vehicular Environment (WAVE) [90-91]. Society of Automotive Engineer (SAE) standards, along with IEEE standards for

vehicular communications are elaborated by John B. Kenney of Toyota Info Technology Center in [92]. To address the limitations of WAVE, Vehicular IP in WAVE (VIP-WAVE) framework is also proposed in [93].

Providing a reliable high data rate on High Speed Trains (HST) is another crucial challenge [106]. It consists of many digital units, multiple Radio Units (RU) and Terminal Equipment (TE). Communication links are formed between the RUs of the digital units and of the TEs [106]. A distributed architecture with an access point in every carriage, for preventing any temporary outage and enhancing the QoE, is proposed for 5G connectivity [94].

E| HEALTH CARE AND WEARABLE

Advancements in sensing and communication technology have opened up new possibilities for health monitoring [95]. US census bureau projects an ageing world in less than 30 years from now. Wearable technology promises to provide health care solutions to growing world strained by the ballooning ageing population [96]. Devices with capabilities of measuring multiple physiological signals in ambulance like environment are being developed in [97]. The record of multiple physiological signals over a long-time period helps in understanding the disease pathophysiology. Improved addressing, extended security services and higher bandwidth enables new possibilities of healthcare [98]. Emerging 5G wireless and Body Area Network (BAN) are facilitating a paradigm shift in real-time remote patients' health monitoring. The major constraint in real time data collection and monitoring is bandwidth limitation. Higher bandwidth and data rates of 5G wireless are expected to resolve this bandwidth constraints. Comfort, physical, psychological and social aspects of wearable devices are discussed in [99]. These capabilities require huge data processing, storage and real time communications. An IoT based system, endowed with big data and cloud computing concepts, for emergency medical services is presented in [78]. 5G wireless is expected to resolve big data challenges of real-time health care applications bringing benefits to the mankind.

F| MISCELLANEOUS APPLICATIONS

Apart from the above-mentioned applications, the financial industry, with increasing businesses and customers, also requires strong computing and data processing [100]. Application of grid computing in financial industry is discussed in [101]. 5G based future mobile networks have a huge potential to transform different financial services [102], like banking, payments, personal finance management, social payments, peer to peer transaction and local commerce.

Sensing, communication and control increases efficiency and reliability of power grids, thereby modernizing them to Smart Grids (SGs). SGs use wireless networks for energy data collection, power line monitoring, protection and demand/response management [102]. Comparisons between smart and existing power grids are given in [103]. Smart information and smart communication subsystem are integral to smart grids. Smart grids seamlessly link physical components and wireless communications representing large-scale cyber-physical systems. Wireless technology is already being explored for efficient real time Demand-Response (DR) management [104]. High bandwidth and low latency of proposed 5G are expected to resolve many challenges associated with SG demand response.

Similarly, smart homes, with roots in automation, embedded systems, entertainment, appliances, efficiency and security is an active technical research area [106]. Smart cities, with fundamentals of sustainability are gaining momentum. Major concepts of IoT, M2M, Cloud computing, integrated with 5G are very persuasive in these research areas.

V| QUALITY OF SERVICE (QoS)

By developing the previous generations of mobile networks, the 3rd Generation Partnership Project (3GPP) has successfully standardized the principles and models of service quality management at the network level. Moreover, the new feature of service quality management has been introduced in 3GPP networks. Ensuring the quality of service (QoS) in 3GPP networks by their evolution from high-speed packet access (HSPA) technology to Long-Term Evolution (LTE)– Advanced technology is based on the following principles [107]:

- Provision of service management by operator
- Differentiation of service quality and users
- Minimal involvement of the user terminal in the service quality management process
- Support of QoS for client applications that are invariant to the access network
- Rapid establishment of sessions
- Continuity of quality management function with mobile networks of previous generations
- Convergence of services in the interaction of mobile networks with fixed-access networks
- Rapid introduction of new services to the market

The implementation of QoS management principles at the network level suggests a steady increase in the number of

mobile applications that control QoS based on the service quality requirements and the creation of the necessary high-level data exchange by bearer services.

Fourth-generation (4G) networks based on QoS model management at the network level have implemented new types of QoS management that can use QoS network model. In these cases, these old applications have to be refreshed. However, we can find some terminals where QoS terminal model management was used. This means that in some years, two QoS management models coexisted in mobile terminals. The situation when two QoS management models were used and the evolution of

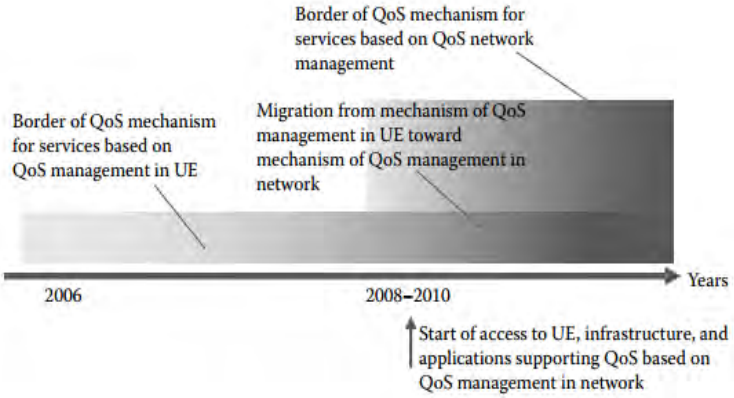


Figure 15: Two QoS management models in mobile networks

these models are shown in Figure 15. the period of 2008–2010 is considered to have been the point of transition from a model of QoS management based on user terminals to the network QoS management model.

To date, there are 3GPP requirements, in particular for general packet radio service (GPRS) networks and packet switched networks, to maintain QoS management at both the user terminal layer and the network layer to provide a smooth transition to QoS management only at the network level.

The implementation of 3G users' requirements to ensure QoS in the "end user– end user" (E2E) chain begins with the activation of the QoS parameter negotiation procedure in the network. Tis procedure depends on the parameters of the user subscription to services stored in the home subscriber server (HSS) database and the current availability of network resources to 3G subscribers, which allows the final compound for a subscriber to be guaranteed. The procedure of QoS parameter approval and QoS management in the 3G network begins with sending a signaling message of session control by the user terminal at the nonaccess stratum (NAS) layer.

In 4G networks, unlike packet connections in second-generation (2G)/3G networks, a typical service of data exchange with a predetermined class of QoS is ready to form a connection to the packet network when a subscriber terminal is connecting to the network. QoS options for data exchange services are determined by the QoS parameters in the user profile, which are stored in the Subscription Profile Repository (SPR) database. Tis situation is very similar to the QoS management in GPRS/3G networks. However, in 4G networks, after transmission of the first data packet from the user terminal, this packet is routed to the packet data network (PDN), where the policy and charging rules function (PCRF) node that manages network policies and billing analyzes the quality class of the requested service in the E2E chain. Depending on the requested service class, the PCRF node can use different modifications of the QoS parameters to all nodes involved in the management of QoS data services. An LTE user terminal, unlike a 2G/3G user terminal, has no opportunity to request a particular QoS class, and only the LTE network is responsible for managing QoS. Similarly, a 4G network subscriber cannot request information about the QoS parameters, as is done, for example, through the use of a secondary context in the 3G network.

One feature of QoS management in the 4G network is that one user terminal can simultaneously support a variety of active services in the E2E chain, and each of these services will have its own individual QoS profile. A 4G user terminal may have up to 256 Evolved UMTS Terrestrial Radio Access Network (E-UTRAN) Radio Access Bearers (E-RAB) (communication services between the access terminal (AT) and the serving gateway [S-GW] service connections) by using E-UTRAN protocols, while in 3G networks only 15 different RAB-IDs are identified.

For the realization of QoS management in the network, we have to define the main QoS parameters for future 5G networks that will enable quality management for the new technology.

A| QoS AS A FACTOR OF THE TRUST IN 5G NETWORKS

Currently, leading organizations in the international standardization and development of telecommunication technologies, such as the International Telecommunication Union (ITU), 3GPP, the Institute of Electrical and Electronics Engineers (IEEE), and the European Telecommunications Standards Institute (ETSI), have not formulated a strict definition of *trusted network*. However, the trust in the communication network significantly affects consumers' choice of communication operator, the regulation of operators' activities by state bodies, and the market demand for communication services and equipment.

The trust in network or communication technology has market and regulatory aspects that can contribute to the

development of the network and the technology and increase the attractiveness of the services. Therefore, networks and communication technologies should correspond to both market and regulatory requirements of trust.

Given the many factors affecting the trust in 5G networks, in this chapter we will briefly review the major factors and examine in detail the impact of service quality on the trust in 5G networks.

The existing understanding of a trusted network is based on the concepts adopted by the developers of computer networks, which traditionally include [108]:

- Secure guest access: Guests obtain restricted network access without threatening the host network.
- User authentication: A trusted network integrates user authentication with network access to better manage who can use the network and what they are allowed to do.
- End point integrity: A trusted network performs a health check for devices connecting to the network. Devices out of compliance can be restricted or repaired.
- Clientless end point management: A trusted network offers a framework to assess, manage, and secure clientless end points connected to the network, such as Internet protocol (IP) phones, cameras, and printers.
- Coordinated security: Security systems coordinate and share information via the Interface for Metadata Access Points (IF-MAP) standard, improving accuracy and enabling intelligent response.

According to the Kaspersky Internet Security Company definition [109], a trusted network is a network that can be considered absolutely safe, within which your computer or device will not be subjected to attacks or unauthorized attempts to gain access to your data.

The proposed comprehensive review of the issue of trusted communication networks complements the concepts of computer network developers with the views of consumers, which also include the QoS provided by trusted networks. Subscribers' and regulators' views on the quality aspects of a trusted network are not always taken into account when creating a new mobile technology that reduces trust in the network.

To implement a systematic approach to the trusted communication network, the trust of two major players in the telecommunications market should be considered: consumers, who provide market demand for communication services, and regulators, who monitor the effectiveness of operators' network infrastructure. As can be seen from Figure 16, consumers' and regulators' requirements for a trusted mobile communication network may either coincide or differ. The main factors

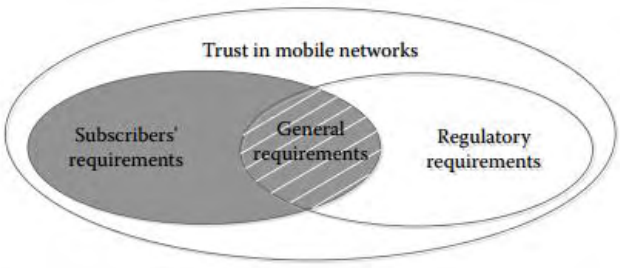


Figure 16: **Domains of trust in mobile networks**

affecting the trust of the subscriber and the regulator are shown in Figure 17, taking into account their importance in descending order.

Most factors are the same for consumers and regulators, but factors determining consumer trust, according to the authors' evaluation, have the dominant influence on the mobile network. Traditional factors influencing consumers' and regulators' trust in 5G networks are information security of confidential user data, security of subscribers' devices, and network infrastructure. The basis for such security is resistance to physical attacks on subscriber devices, such as illegal substitution of identification modules (Universal Subscriber Identity Module [USIM] cards), installation of malicious software on the user device and its impact on the user device configuration, resistance to network attacks on user devices and network infrastructure, such as denial of service (DoS) attacks and "man in the middle" attacks, and resistance to attacks on confidential user data.

Consumer	*Regulator*
Quality of service	Network security
Quality of experience	Information security
Information security	Network performance
Network performance	Network reliability
Network reliability	Quality of service
Convenience and security of subscriber's equipment	

Figure 17: Main factors affecting the trust of the subscriber and the regulator in the network

In addition to the security performance, the trust of users and regulators in 5G networks will depend on quality performance, since the security of the mobile network itself does not guarantee that the communication service will be provided without interruption and with the stated quality. A reduced quality in 5G networks will lead to a decrease of trust in them, and as a result, in an outflow of subscribers. Also, given that the 5G network

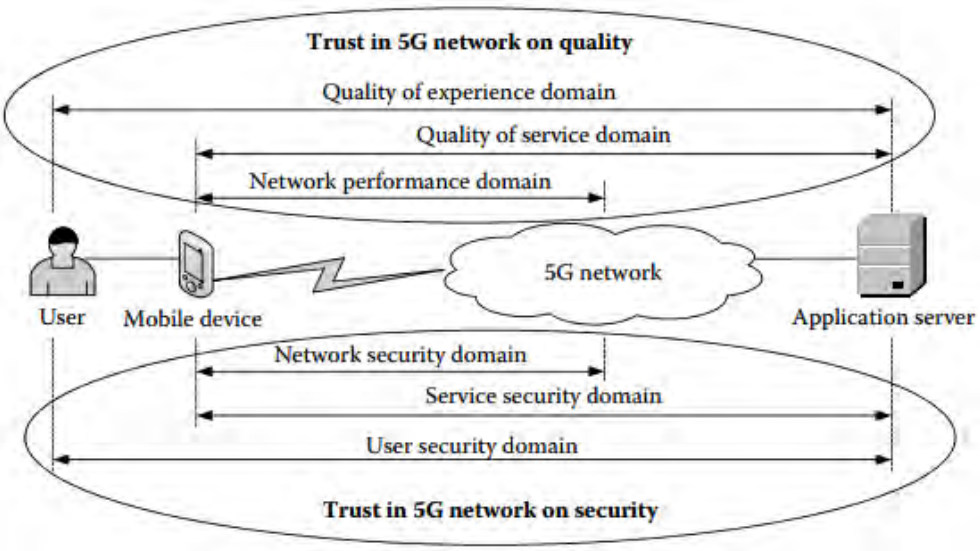


Figure 18: Quality and security levels of trust in a mobile network

will be used in a variety of financial systems, public safety systems, and traffic and energy management systems, a deterioration in their quality could lead to the loss of human life, environmental disasters, and financial fraud.

The quality parameters of 5G networks can be divided into three levels: network performance (NP), QoS, and quality of experience (QoE), as shown in Figure 18. NP and QoS are objective indicators that can be measured using specialized analyzers, while QoE indicators are subjective, estimated by users on the basis of their personal experience. The deterioration of QoS and NP will primarily lead to lowering the trust of regulators and business-to-business (B2B) and business-to-government (B2G) customers in 5G networks, while QoE deterioration will lead to lowering the trust of the mass market.

B| SERVICES AND TRAFFIC

To meet the QoS requirements in future 5G networks, we can take advantage of the analysis of future requirement levels for services such as high-density video and machine-to-machine communications, and then transfer these levels into 5G QoS requirements. Mobile and Wireless Communications Enablers for the Twenty-Twenty Information Society (METIS) projects consider three basic business models of 5G services: extreme Mobile Broadband (xMBB), massive machine-type communications (M-MTC), and ultrareliable machine-type communications
(U-MTC) [110].

Forecasts from the leading specialists working on international 5G projects [111–113] show that video services, such as high-definition (HD) and ultra-high definition (UHD) video, with high-quality resolution will have a dominant position among services rendered in 5G networks. According to reports by leading 4G network operators, video services dominate in the subscribers' traffic and will continue to dominate in the content of 5G networks.

For instance, the current traffic volume of video services is estimated by different operators to be from 66% to 75% of the total traffic in 4G networks, including 33% for YouTube services and 34% for clear video, as well as closed circuit television (CCTV) monitoring (video surveillance) in machine-to-machine (M2M) networks. In addition, by 2020, the volume of mobile M2M connections will grow with a compound annual growth rate (CAGR) index of 45% [114], up to 2.1 billion connections. Given the growing mass scale of M2M services in all industries, they will dominate over basic services (voice and data) in 4G and 5G networks.

The 5G European development strategy also aims to enable subscribers by 2025 to choose how to connect to TV broadcasts, via a 5G modem or antenna with digital video broadcasting–terrestrial (DVB-T), so this will require appropriate quality management mechanisms.

Therefore, the efforts of developers to improve quality management mechanisms will focus on video and M2M services traffic, the improvement of quality checking algorithms, and the creation of new quality assessment methods. When developing requirements for QoS in 5G networks, two key traffic models should be firstly considered: high-speed "server–subscriber" video flow and massive M2M. Video transmission services will be an important stimulus to development and a rapidly growing segment of 5G network traffic. In 2013, the volume of video services in the total traffic of 4G network subscribers already exceeded 50%, and by 2019, it is forecast to

increase by at least 13 times [115]. Tus, we can already observe the first wave of the oncoming "tsunami" of subscribers' traffic in 4G networks. The monthly consumption of data transmission traffic in 4G networks has already reached 2.6 GB, and the monthly consumption of traffic in 5G networks will exceed 500 GB.

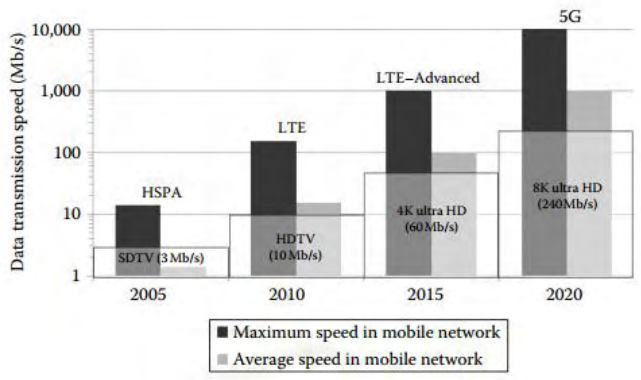


Figure 19: Technological capabilities of video transfer for mobile networks of various generations

The growth of video services traffic volume will be associated with the implementation of various technologies of video services image quality, from standard definition (SD) TV to UHD TV (8K), which in its turn requires a data transmission speed of up to 10 Gb/s in the network. The technological capabilities of mobile networks of various generations to broadcast video with various qualities of video image are shown in Figure 19. The capability of video broadcasting depends on the data transmission speed in the radio access network (RAN) [116].

According to the forecasts shown in Figure 20, in 2018, the number of M2M connections in the networks of mobile operators will exceed 1.5 billion [116], which is five times higher than the current rate, and in 2022, mobile operators will have more than 2.6 billion M2M connections. At the same time, the share of M2M connections out of the total number of connections in the mobile operators' networks will increase from the current 5% to 15% in 2018 and to 22% in 2022.

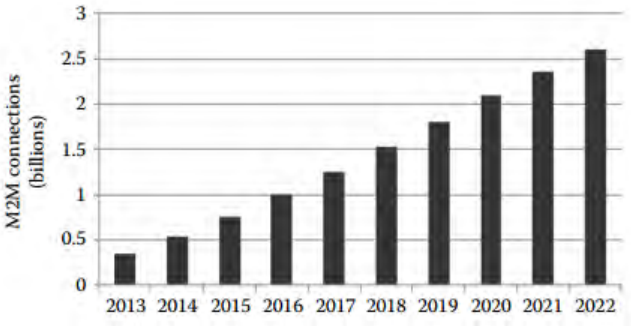


Figure 20: Number of M2M connections in mobile networks

The strategies of M2M operators are aimed at creating universal M2M platforms capable of operating in multiple vertical economic sectors. Tis will lead to the possibility of implementing approaches, tools, and processing methods for structured and unstructured Big Data derived from M2M networks. According to ABI Research forecasts, the M2M Big Data and analytics industry will grow by a robust 53.1% over the next 5 years, from US$1.9 billion in 2013 to US$14.3 billion in 2018. Tis forecast includes revenue segmentation for the five components that together enable analytics to be used in M2M services: data integration, data storage, core analytics, data presentation, and associated professional services.

M2M services require much smaller data rates than video services, and generally do not require a guaranteed data rate. However, many M2M services, especially those used in the management of industrial systems, are critical of delays in mobile networks. Therefore, M2M services will also affect the quality of 5G networks.

C| QoS PARAMETERS

Quality control and management in mobile networks are based on the use of the key QoS parameters, such as bit rate, latency, and packet loss.

In the current generation of mobile networks, there are two major types of network bearers: guaranteed bit rate (GBR) and nonguaranteed bit rate (non-GBR). GBR bearers are used for real-time services, such as rich voice and video. A GBR bearer has a minimum amount of bandwidth that is reserved by the network, and always consumes resources in a radio base station regardless of whether it is used or not. If implemented properly, GBR bearers should not experience packet loss on the radio link or the IP network due to congestion. GBR bearers will also be defined with the lower latency and jitter tolerances that are typically required by real-time services.

Non-GBR bearers, however, do not have a specific network bandwidth allocation. Non-GBR bearers are used for best-effort services, such as file downloads, e-mail, and Internet browsing. These bearers will experience packet loss when a network is congested. A maximum bit rate for non-GBR bearers is not specified on a per-bearer basis. However, an aggregate maximum bit rate (AMBR) will be specified on a per-subscriber basis for all non-GBR bearers.

Packet Delay Budget (PDB): Tis parameter identifies a maximum acceptable end-to-end delay between the user equipment (UE) and the packet data network gateway (PDN-GW). The purpose of using the PDB parameter is to support the queues of planning process and network functions at the connection level. The maximum delay budget (MDB) parameter is interpreted as the maximum packet delay with a confidence level of 98%. The PDB parameter defines the time limit for packet delay, for

which the "final" package of the session will be transmitted with a delay of not greater than a predetermined value of the PDB. In this case, the packet should not be dropped.

Packet Error Loss Rate: This is the proportion of packets lost due to errors when receiving data packets. The maximum value of this parameter specifies the largest number of data packets lost during transmission over the network.

According to the assumptions of this analysis, these QoS parameters will be used in the process of developing 5G QoS requirements supporting the three main business models of 5G.

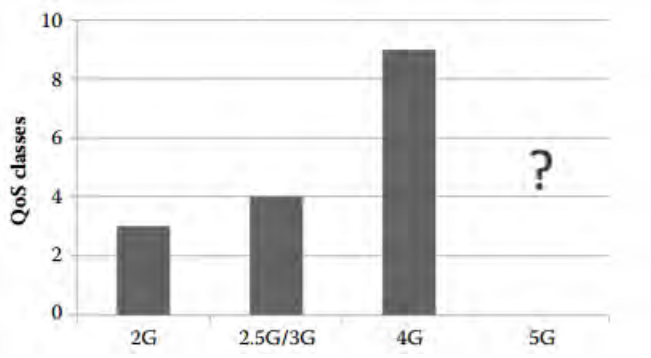


Figure 21: Evolution of QoS classes in mobile networks

D| QUALITY REQUIREMENTS IN 5G NETWORKS

5G mobile technologies that are expected to appear on the market in 2020 should significantly improve customers' QoS in the context of the snowballing growth of data volume in mobile networks and the growth in the number of wireless devices and the variety of services provided [111]. It is expected that mobile communication networks built on the basis of 5G technologies will provide a data transfer speed of more than 10 Gb/s.

Figure 21 shows the evolution of the QoS class from 2G to 4G, which has doubled the QoS class. These trends raise the question of how many QoS classes will be enough for 5G.

Previous 4G technologies (LTE/LTE–Advanced) provide flexible QoS management based on the division of data transfer characteristics into nine classes. These classes cover both of the 4G quality principles: services provision without quality assurance (best effort or non-GBR) and guaranteed QoS provision (GBR).

Unfortunately, these LTE technological advances in the field of QoS management cover only part of the E2E chain, in particular 5G–5G and 4G–4G intra-network connections. The quality management system does not extend to connections between 5G subscribers and other mobile 2G/3G/4G and fixed networks. The absence of the possibility of coordinated and flexible quality management in fixed IP and mobile networks of previous generations will for a long time remain a brake on the new level of subscribers' service quality in 5G networks.

The METIS project has identified 12 use cases for 5G networks: virtual reality office, dense urban information society, shopping mall, stadium, tele protection in smart grid network, traffic jam, blind spots, real-time remote computing for mobile terminals, open-air festival, emergency communications, massive deployment of sensors and actuators, and traffic safety and efficiency, and has developed QoE requirements for them [116]. The QoE performance requirements that provide trust in network 5G are presented in Figure 22. The highest requirements for experienced user throughput are developed for the "virtual reality office" use case. End users should be able to experience data rates of at least 1 Gbps in 95% of office locations and for 99% of the busy period. Additionally, end users should be able to experience data rates of at least 5 Gbps in 20% of office locations, for example, at their actual desks, for 99% of the busy period. The highest requirements for network latency are developed for the "dense urban information society" use case, in which deviceto-device (D2D) latency is less than 1 ms. The highest requirements for availability and reliability of the 5G network are identified for the "traffic safety and efficiency" use case: 100% availability with a transmission reliability of 99.999% is required to provide services at every point on the road.

During the evolution of QoS management mechanisms in 3GPP (Global System for Mobile Communications [GSM]/Universal Mobile Telecommunications System [UMTS]/LTE) networks, there was a migration from QoS management at the UE level to QoS management at the network level. Tis approach to QoS management will be maintained in 5G networks as well. QoS management mechanisms in 5G networks should provide video and voice over IP (VoIP) traffic prioritization toward web-search traffic and other applications tolerant to quality.

QoE Indicators	*Requirements*
Experienced user throughput	5 Gbps in data level (DL) and user level (UL)
Latency	D2D latency less than 1 ms
Availability	≈100%
Reliability	99.999%

Figure 22: QoE Performance Requirements for 5G Networks

The service of streaming video transfer without buffering is very sensitive to network delay, so one of the most important parameters that determine QoS requirements is the total PDB, which is formed on the RAN air interface and is treated as the maximum packet delay with a confidence level of 98%.

QoS Terms	Packet Delay Budget (ms)		
	3G	4G	5G
Without quality assurance	Not determined	100–300	Not determined
With guaranteed quality	100–280	50–300	1

Figure 23: Requirements for Delay in 3G/4G/5G Networks

Figure 23 lists the requirements for delay in 3G/4G/5G networks formed in 3GPP [116] and the METIS project. These data demonstrate that with the increase in mobile network generation, the requirements for the lower boundary of the total data delay across the network decline. Also, an analysis of the requirements for the overall 5G network delay revealed that given the accumulation effect, the delay in the 5G RAN network should be less than 1 ms.

A comparison of the requirements for delay in the control and user planes for signaling traffic and user traffic, respectively, presented in Figure 24, shows that the requirements for 5G networks will be twice as rigid for traffic in the user plane and 10 times more rigid in the subscriber traffic plane [114].

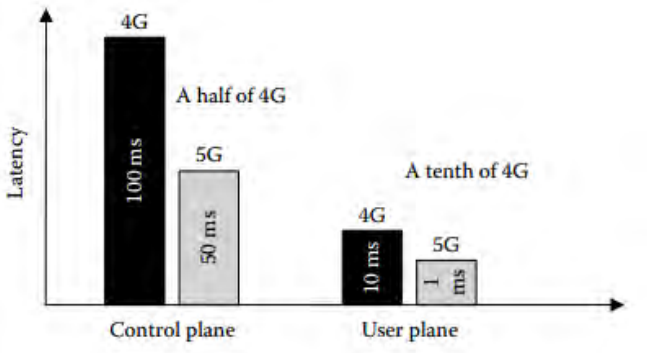


Figure 24: Requirements for delay in control and user planes for 4G/5G networks

The emergence of 5G networks in the market in 2020 will focus on a significant improvement in the characteristics of mobile networks, including QoS, which will provide a high level of trust in these networks.

A one-sided view of trusted 5G networks from the security position alone will limit the growth in trust of customers and regulators. Developing high-level requirements in the QoS field will allow 5G developers to obtain trust in 5G at an early stage.

Given that the principles of QoS control will be preserved during the transition from 4G to 5G, the main effort of 5G developers should be focused on the virtualization of network functions that are responsible for the management and control of QoS in the network. Also, the QoS architecture of 5G should provide information exchange between the QoS manager and the Spectrum Manager for the effective management of spectrum resources, with the benefit of ensuring QoS and trust in 5G networks.

VI| ENERGY EFFICIENCY FOR 5G NETWORKS

The need for media rich input/output, computation and communication forces users to charge their mobile devices more often. On the other side, network operators have been adding more and more base stations (BSs) to meet a higher service demand. A large portion of the network operators' operational costs are due to energy consumption of the wireless systems for both mobile users and service providers. The fifth-generation infrastructure, when defined as the ultra-broadband network, will be associated with the true revolution into communication field, taking forward new services to everyone and everything. The 5G mobile networks are required to have tremendous spectral efficiency (SE) and energy efficiency (EE) improvement simultaneously. As a result, these efficiencies together with cost efficiency have been widely used like three obligatory 5G evolution metrics. In addition, joint energy and communication can maximally save cost by applying both energy cooperation on the supply side and communication operation on the demand side. As for 5G networks, the primary goal is to satisfy a variety of users' needs in a more energy efficient manner. The question often arises is where the energy can be further saved and what available information in the network can be explored.

A| ENERGY SAVING APPROACH

The number of subscribers and thus the increased traffic intensity in cellular networks has moved the limits of capacity and energy consumption because mobile equipment is working in all days of a 24/7/365 regime. On the other hand, energy saving approach is one of the way to improve the efficiency of cellular networks. For network operators, it is difficult to maintain the capacity growth utilize bandwidth, decrease delay and to

limit at the same time energy consumption. The following framework has to be taken into account [117]:

Deployment-energy tradeoff to balance the deployment cost, throughput and energy consumption in the whole network

Spectrum-energy tradeoff in order to balance the accessible rate and energy consumption of the system

Bandwidth-power tradeoff in order to balance the bandwidth utilized and the required power for transmission

Delay-power tradeoff to balance the average end-to-end delay and the average consumed power in the transmission.

Using these tradeoffs in different research aspects, energy saving approach can be easily obtained. Network infrastructure, sharing and BS sharing off are promising solutions for energy and cost reduction bringing for mobile operators the decrement of capital and operational expenditures associated with the deployment and the operation of the cellular networks [118]. There exist three types of sharing [119]: passive, active and roaming-based. Passive sharing understands the joint use of sites, masts, building among mobile network operators (MNOs). In active sharing the MNOs share antennas, switches, and backhaul network equipment, while in roaming-based sharing one MNO relies on the coverage of another one on a permanent basis in a region.

B| ENERGY AND INFORMATION FLOW IN MOBILE NETWORKS

Applying both energy and communication cooperation maximally save cast can be achieved. Energy cooperation is on the supply side, while communication one is on the demand side. An example of a model of cellular networks with energy and communication cooperation among BSs [120] is shown in Figure 25.

The energy trading and sharing is enabled using aggregator. This is upper energy cooperation layer [121]. In principle, aggregators enable to cluster BSs into a finite number of groups. An aggregator serves as an intermediary device to control each group of BSs for the grid. In that case, two-way energy flow between the grid and BS groups will be provided. Communication cooperation serves to a cost saving on the demand side and seeks to minimize the total energy cost by optimally utilizing cheap renewable energy and reliable on-grid energy. In the case of joint operation, the BSs share the energy information by using the two-way information flow through smart meters. The communication information is exchanged through the backhaul connections. It can be seen that the joint energy and communication cooperation is more complex compared to energy or communication operation separately.

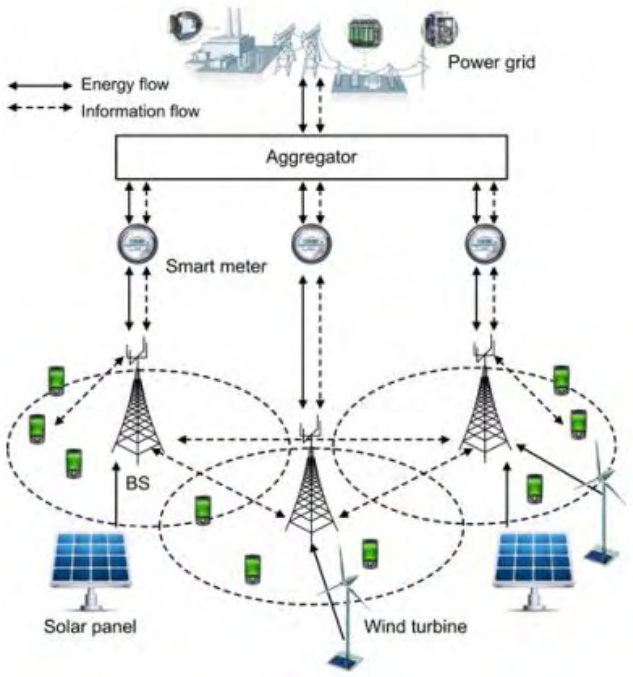


Figure 25: Energy and communication cooperation in mobile networks

C| ADVANCES IN ENERGY EFFICIENCY FOR MOBILE NETWORKS

EE for wireless networks is the metric of interest when taking into account energy consumption. Among other definitions, a satisfactory operational one which was proposed by Lawrence Berkeley National Laboratory says that EE "is using less energy to provide the same service". The ability to adopt the transmission strategy according to the traffic demands will become an important design aspect of EE. Network infrastructure should be regarded as a resource that can be released or occupied on demand. To handle the exponential growth of mobile data traffic while alleviating the huge cost of infrastructure investment, massive multiple-input multiple-output (MIMO), small cells, and device-to-device (D2D) communications have been proposed for Long Term Evolution Advanced (LTE-A) networks. The full application of these technologies could be expected in 5G systems.

Network energy performance is one of a crucial requirement of 5G networks, because of reduced cost ownership and facilities in the process of extensions network connectivity to remote areas. 5G systems with high energy performance should be built on the principle to only be active and transmit when and where needed. To only be active when needed implies an always available approach with dynamic activation on several levels: nodes,

functionality, subsystems and components. To only transmit when needed refers in particular to minimized transmissions not directly related to the delivery of user data. To only be active where needed covers the spatial domain of "always available" and may refer both to the same levels as above but with addition of extra dimension to distributed structures.

One of the goals of 5G networks is transmitting as high data rate as possible taking into account users' need in a more energy efficient manner. In connection with this, it should be noted the importance of identifying where energy can be saved. Outlines for advances in 5G mobile network technology can be presented in a step by step form in Figure 26.

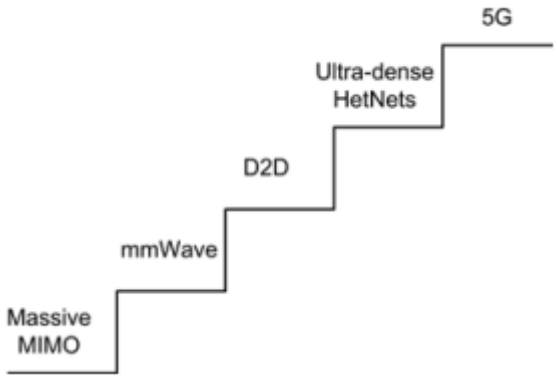


Figure 26: Outlines for advances in 5G mobile networks

Traditional MIMO systems improve EE because of the array gain, spatial diversity and/or spatial multiplexing gains. On the other side, massive (large-scale) MIMO is a form of multi-user MIMO in which the number of antennas at the BS is much larger comparing to the number of devices per signaling resource. This concept allows for order of magnitude improvement in SE and EE using relatively simple processing [122, 123]. Massive MIMO can increase the capacity 10 times or more and simultaneously improve the radiated EE on the order of 100 times. Finally, a small cell is formed using low-power and low-cost micro, pico and femto BSs. With an improved frequency reuse factor, a small cell is able to enhance SE. At the same time, with reduced distance between the user and the BS, the required transmit power to overcome path loss, fading and noise is also reduced. As a result, both uplink and downlink EE can be improved [124].

There is a great potential for 5G networks to improve both SE and EE, leading toward ultra-dense heterogeneous network (HetNet) environment [125]. As for EE, the deployment of massive MIMO and HetNets for same coverage is still under discussion. Control in HetNets can provide more flexible and higher EE for a large number of small cells, while a massive MIMO cell performs better than a small number of small cells due to its large array gain [126]. By reducing the size of the cell, area SE is increased through higher frequency reuse, while transmit power can be reduced such that the power lost through propagation will be lower. This approach can provide a flexible coverage and improved SE and EE.

An approach to solve the highly dense network problems will be through direct mobile-to-mobile communication in order to either share their radio access connection, or to exchange information. D2D communications can reduce interferences especially in unlicensed frequency bands. Together with small cells, D2D communication will support low-cost architecture. In that way D2D communications can significantly improve EE for devices. Also, more freedom for D2D users in the sense of directly transmission of data, directly reusing the resource of cellular users and by communicating with each other through the BS in the standard way, can be provided [127]. An example of D2D communications in HetNets is shown in Figure 27.

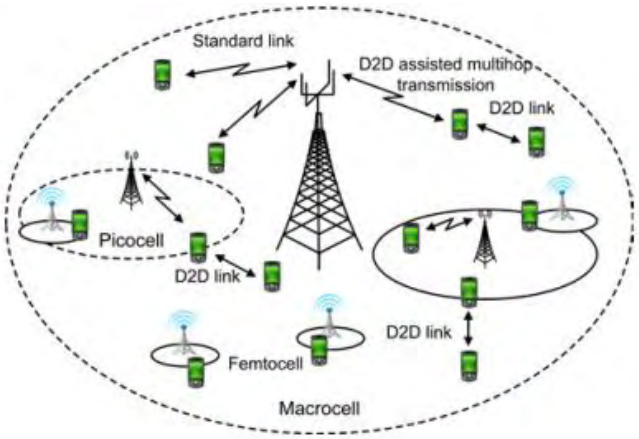


Figure 27: An example of D2D communications in HetNets

Users in cellular networks can transmit data directly to each other. Due to physical proximity, D2D communication provides proximity gain, reuse gain and hip gain. In that way D2D communication improves EE, which is one of significant performance metrics. The main characteristics are: energy constraints, bandwidth constraints, and multihop transmission [128]. In terms of energy constrained networks, the capacity per energy cost over a single link is considered in [129], while the bit per Joule capacity for wireless data delivery is proposed in [130]. For short distance communications, energy cost in the circuit becomes nontrivial, yielding a more complicated issue for the possible power region.

VII| CONCLUSION

Rapid penetration of wireless connectivity, almost exponential increase in wireless data (multimedia) usage and proliferation of feature-rich smart devices are gradually setting the stage for next major cellular evolution towards 5G. Next generation 5G wireless systems are already promising a manifold increase in data rate, connectivity and QoS. A plethora of new applications, like IoT, smart grids and IoV are expected to be supported under the umbrella of 5G systems. In this survey, we provide a comprehensive review of cellular evolution towards 5G networks. We begin with pointing out the new architectural paradigm shift, associated with the design of radio network layout, air interfaces, smart antennas, cloud and heterogeneous RAN. Subsequently, we give a detailed description of the underlying physical layer technology. This includes understanding of new physical channels, estimating new channel models with LOS/NLOS, novel antenna design, beamforming and massive MIMO. Next, we discuss the major MAC layer protocols and multiplexing schemes, like SDMA, IDMA and evolution of existing OFDM, required to efficiently support the new physical characteristics. Novel emerging applications, like D2D and M2M communications, IoT, Vehicular communications and Healthcare applications form the major driving force behind 5G. We digress into the details of these killer applications to understand the associated impact on the cellular evolution. As 5G is expected to offer a much better user experience, we highlight the new QoS, QoE and SON features, associated with the evolution of 5G networks. A major concern behind the massive roll out of 5G lies in increasing energy consumption and the associated greenhouse gas emissions and opex. This motivates us to make a review on energy aware BS, energy efficient backhaul and cost efficiency. In order to realize the current state of the implementation, we also look into the major 5G field trials, drive tests and simulations. Finally, we point out major existing research challenges and identify possible future research directions. We believe that our survey will serve as a guideline for major future research works in 5G wireless communications.

[1] Liu, L., Chen, R., Geirhofer, S. et al. (2012) Downlink MIMO in LTE-Advanced: SU-MIMO vs. MU-MIMO. IEEE Communications Magazine, 50(2), 140–147

[2] GSMA Intelligence, last accessed Dec. 2014 https://www.gsmaintelligence.com/research/?file=141208-5g.pdf&download

[3] Heavy Reading, White Paper, 5G: A Network Transformation Imperative http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/5g-a-network-transformation-imperative.pdf

[4] Ericsson.com, last accessed Oct. 2014, 5G: What Is It? https://www.ericsson.com/res/docs/2014/5g-what-is-it.pdf

[5] IMT – 2020 White Paper 2015 5G Network Architecture https://www.gsmaintelligence.com/research/?file=141208-5g.pdf&download

[6] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," IEEE Commun. Mag., vol. 52, no. 2, pp. 74–80, Feb. 2014

[7] T. S. Rappaport et al., "Millimeter wave mobile communications for 5G cellular: It will work!" IEEE Access, vol. 1, pp. 335–345, May 2013

[8] M. Olsson, C. Cavdar, P. Frenger, S. Tombaz, D. Sabella, and R. Jantti, "5GrEEn: Towards green 5G mobile networks," in Proc. IEEE Int. Conf. Wireless Mobile Comput. Netw. Commun., 2013, pp. 212–216

[9] T. S. Rappaport, F. Gutierrez, E. Ben-Dor, J. N. Murdock, Y. Qiao, and J. I. Tamir, "Broadband millimeter wave propagation measurements and models using adaptive beam antennas for outdoor urban cellular communications," IEEE Trans. Antennas Propag., vol. 61, no. 4, pp. 1850–1859, Apr. 2013

[10] R. Taori and A. Sridharan, "In-band, point to multi-point, mm-wave backhaul for 5G networks," in Proc. IEEE Int. Conf. Commun. Workshops, 2014, pp. 195–201

[11] Z. Pi and F. Khan, "System design and network architecture for a millimeter-wave mobile broadband (MMB) system," in Proc. IEEE Sarnoff Symp., 2011, pp. 1–6

[12] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," IEEE Commun. Mag., vol. 49, no. 6, pp. 101–107, Jun. 2011

[13] T. Korakis, G. Jakllari, and L. Tassiulas, "A MAC protocol for full exploitation of directional antennas in ad-hoc wireless networks," in Proc. ACM Int. Symp. Mobile Ad Hoc Netw. Comput., 2003, pp. 97–108

[14] J. Bae, Y. S. Choi, J. S. Kim, and M. Y. Chung, "Architecture and performance evaluation of mmWave based 5G mobile communication system," in Proc. Int. Conf. Inf. Commun. Technol. Convergence (ICTC), 2014, pp. 847–851

[15] S. Rajagopal, S. Abu-Surra, Z. Pi, and F. Khan, "Antenna array design for multi-gbps mmwave mobile broadband communication," in Proc. Global Telecommun. Conf. (Globecom), 2011, pp. 1–6

[16] Z. Feng and Z. Zhang, "Dynamic spatial channel assignment for smart antenna," Wireless Pers. Commun., vol. 11, no. 1, pp. 79–87, 1998

[17] Roh et al., "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," IEEE Commun. Mag., vol. 52, no. 2, pp. 106–113, Feb. 2014

[18] P. Cardieri and T. S. Rappaport, "Application of narrow-beam antennas and fractional loading factor in cellular communication systems," IEEE Trans. Veh. Technol., vol. 50, no. 2, pp. 430–440, Mar. 2001

[19] Kallnichev, "Analysis of beam-steering and directive characteristics of adaptive antenna arrays for mobile communications," IEEE Antennas Propag. Mag., vol. 43, no. 3, pp. 145–152, Jun. 200

[20] F. Lai, R. H. Hwang, H. C. Chao, M. Hassan, and A. Alamri, "A buffer-aware HTTP live streaming approach for SDN-enabled 5G wireless networks," IEEE Netw., vol. 29, no. 1, pp. 49–55, Jan. 2015

[21] Agyapong, M. Iwamura, D. Staehle, W. Kiess, and A. Benjebbour, "Design considerations for a 5G network architecture," IEEE Commun. Mag., vol. 52, no. 11, pp. 65–75, Nov. 2014

[22] Lara, A. Kolasani, and B. Ramamurthy, "Network innovation using openflow: A survey," IEEE Commun. Surv. Tuts., vol. 16, no. 1, pp. 493–512, First Quart. 2014

[23] H. Cho, C. F. Lai, T. K. Shih, and H. C. Chao, "Integration of SDR and SDN for 5G," IEEE Access, vol. 2, pp. 1196–1204, Oct. 2014

[24] Arslan, K. Sundaresan, and S. Rangarajan, "Software-defined networking in cellular radio access

networks: Potential and challenges," IEEE Commun. Mag., vol. 53, no. 1, pp. 150–156, Jan. 2015

[25] Checko et al., "Cloud RAN for mobile networks-a technology overview," IEEE Commun. Surv. Tuts., vol. 17, no. 1, pp. 405–426, First Quart. 2015

[26] Cvijetic, "Optical network evolution for 5G mobile applications and SDN-based control," in Proc. Int. Telecommun. Netw. Strategy Plann. Symp., 2014, pp. 1–5

[27] Chen and R. Duan, "C-RAN: The road towards green RAN," China Mobile Research Institute, Beijing, White paper, 2011

[28] Liu, J. Wang, L. Cheng, M. Zhu, and G. K. Chang, "Key microwavephotonics technologies for next-generation cloud-based radio access networks," J. Lightw. Technol., vol. 32, no. 20, pp. 3452–3460, 2014

[29] Banikazemi, D. Olshefski, A. Shaikh, J. Tracey, and G. Wang, "Meridian: An SDN platform for cloud network services," IEEE Commun. Mag., vol. 51, no. 2, pp. 120–127, Feb. 2013

[30] Rost et al., "Cloud technologies for flexible 5G radio access networks," IEEE Commun. Mag., vol. 52, no. 5, pp. 68–76, May 2014

[31] Zhou and W. Yu, "Optimized backhaul compression for uplink cloud radio access network," IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1295–1307, Jun. 2014

[32] Zhang, N. Cheng, A. T. Gamage, K. Zhang, J. W. Mark, and X. Shen, "Cloud assisted HetNets toward 5G wireless networks," IEEE Commun. Mag., vol. 53, no. 6, pp. 59–65, Jun. 2015

[33] M. Abd El-atty and Z. M. Gharsseldien, "On performance of HetNet with coexisting small cell technology," in Proc. IEEE Conf. Wireless Mobile Netw., 2013, pp. 1–8

[34] M. S. Huq, S. Mumtaz, M. Alam, J. Rodriguez, and R. L. Aguiar, "Frequency allocation for HetNet CoMP: Energyefficiency analysis," in Proc. Int. Symp. Wireless Commun. Syst., 2013, pp. 1–5

[35] Wang, H. Li, H. Wang, and S. Ci, "Probability weighted based spectral resources allocation algorithm in Hetnet under Cloud-RAN architecture," in Proc. Int. Conf. Commun. China Workshops, 2013, pp. 88–92

[36] Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," IEEE Wireless Commun., vol. 21, no. 3, pp. 118– 127, Jun. 2014

[37] Sanguinetti, A. L. Moustakas, and M. Debbah, "Interference management in 5G reverse TDD HetNets with wireless backhaul: A large system analysis," IEEE J. Sel. Areas Commun., vol. 33, no. 6, pp. 1187– 1200, Jun. 2015

[38] Nam, D. Bai, J. Lee, and I. Kang, "Advanced interference management for 5G cellular networks," IEEE Commun. Mag., vol. 52, no. 5, pp. 52–60, May 2014

[39] Talwar, D. Choudhury, K. Dimou, E. Aryafar, B. Bangerter, and K. Stewart, "Enabling technologies and architectures for 5G wireless," in Proc. MTT-S Int. Microw. Symp. (IMS), 2014, pp. 1–4

[40] Galinina et al., "Capturing spatial randomness of heterogeneous cellular/WLAN deployments with dynamic traffic," IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1083–1099, Jun. 2014

[41] L. Lee, T. C. Chuah, J. Loo, and A. Vinel, "Recent advances in radio resource management for heterogeneous LTE/LTE-A networks," IEEE Commun. Surv. Tuts., vol. 16, no. 4, pp. 2142–2180, Fourth Quart. 2014

[42] Xu et al., "Cooperative distributed optimization for the hyper-dense small cell deployment," IEEE Commun. Mag., vol. 52, no. 5, pp. 61–67, May 2014

[43] Shen and W. Yu, "Distributed pricing-based user association for downlink heterogeneous cellular networks," IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1100–1113, Jun. 2014

[44] Q. Hu and Y. Qian, "An energy efficient and spectrum efficient wireless heterogeneous network framework for 5G systems," IEEE Commun. Mag., vol. 52, no. 5, pp. 94–101, Feb. 2014

[45] Weiser, M., 'The computer for the 21st century', Scientific American, 265 (3), 94–104

[46] Bangerter, B., Talwar, S., Arefi, R. and Stewart, K., 'Networks and devices for the 5G era', IEEE Communications Magazine, 52(2), 90–96

[47] La Polla, M., Martinelli, F. and Sgandurra, D., 'A survey on security for mobile devices', Communications Surveys & Tutorials, IEEE, 15(1), 446–471

[48] Becher, M., Freiling, F. C., Hoffmann, J. et al., 'Mobile security catching up? Revealing the nuts and bolts of the security of mobile devices', Security and Privacy (SP), 2011 IEEE Symposium on (pp. 96–111)

[49] Arabo, A. and Pranggono, B., 'Mobile malware and smart device security: trends, challenges and solutions', Control Systems and Computer Science (CSCS), 2013 19th International Conference on (pp. 526–531). IEEE

[50] Flo, A.R. and Josang, A., 'Consequences of botnets spreading to mobile devices', Short-Paper
Proceedings of the 14th Nordic Conference on Secure IT Systems (NordSec 2009) (pp. 37–43)

[51] Piqueras Jover, R., 'Security attacks against the availability of LTE mobility networks: Overview and research directions', Wireless Personal Multimedia Communications (WPMC), 2013 16th International Symposium on (pp. 1–9)

[52] Seddigh, N., Nandy, B., Makkar, R. and Beaumont, J.F., 'Security advances and challenges in 4G wireless networks', Privacy Security and Trust (PST), 2010 Eighth Annual International Conference on (pp. 62– 71)

[53] Forsberg, D., Leping, H., Tsuyoshi, K. and Alanara, S., 'Enhancing security and privacy in 3GPP E- UTRAN radio interface', Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on (pp. 1–5)

[54] Bassil, R., Chehab, A., Elhajj, I. and Kayssi, A., 'Signaling oriented denial of service on LTE networks', Proceedings of the 10th ACM International Symposium on Mobility Management and Wireless Access (pp. 153–158)

[55] EPC (2014). 3GPP-The Evolved Packet Core http://www.3gpp.org/technologies/keywords-acronyms/100-the-evolved-packet-core

[56] 3GPP TS 23.002 V12.4.0. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Network architecture (Release 12), March 2014

[57] Traynor, P., Lin, M., Ongtang, M. et al., 'On cellular botnets: measuring the impact of malicious devices on a cellular network core', Proceedings of the 16th ACM Conference on Computer and Communications Security (pp. 223–234)

[58] Li, F., Peng, W., Huang, C. T. and Zou, X., 'Smartphone strategic sampling in defending enterprise network security', Communications (ICC), 2013 IEEE International Conference on (pp. 2155–2159)

[59] 3GPP TR 33.820 V8.3.0. 3rd Generation Partnership Project; Technical Specification Group Service and System Aspects; Security of H(e)NB (Release 8), December 2009

[60] P. Agyapong, M. Iwamura, D. Staehle, W. Kiess, and A. Benjebbour, "Design considerations for a 5G network architecture," IEEE Commun. Mag., vol. 52, no. 11, pp. 65–75, Nov. 2014

[61] 3GPP LTE Release 8.0, Overview of 3GPP release 8 V 0.3.3 (2014- 2009)
http://www.3gpp.org/specifications/releases/ 72-release-8

[62] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," IEEE Commun. Surv. Tuts., vol. 16, no. 4, pp. 1801–1819, Fourth Quart. 2014

[63] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: Challenges, solutions, and future directions," IEEE Commun. Mag., vol. 52, no. 5, pp. 86–92, May 2014

[64] X. Wu, S. Tavildar et al., "FlashLinQ: A synchronous distributed scheduler for peer-to-peer ad hoc networks,". IEEE/ACM Trans. Netw. (TON), vol. 21, no. 4, pp. 1215–1228, Aug. 2013

[65] N. Bhushan et al., "Network densification: The dominant theme for wireless evolution into 5G," IEEE Commun. Mag., vol. 52, no. 2, pp. 82–89, Feb. 2014

[66] H. Ding, S. Ma, and C. Xing, "Feasible D2D communication distance in D2D-enabled cellular networks," in Proc. IEEE Int. Conf. Commun. Syst., 2014, pp. 1–5

[67] Y. Jung, E. Festijo, and M. Peradilla, "Joint operation of routing control and group key management for 5G ad hoc D2D networks," in Proc. Int. Conf. Privacy Secur. Mobile Syst. (PRISMS), 2014, pp. 1–8

[68] 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

[69] N. Naderializadeh and A. S. Avestimehr, "ITLinQ: A new approach for spectrum sharing in Device-to-Device communication systems," IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1139–1151, Jun. 2014

[70] J. M. B. da Silva, G. Fodor, and T. F. Maciel, "Performance analysis of network-assisted two-hop D2D communications," in Proc. IEEE Globecom Workshops Broadband Wireless Access, 2014, pp. 1050– 1056

[71] Y. Zhang, R. Yu, M. Nekovee, Y. Liu, S. Xie, and S. Gjessing, "Cognitive machine-to-machine communications: Visions and potentials for the smart grid," IEEE Netw., vol. 26, no. 3, pp. 6–13, May/Jun. 2012

[72] J. Kim, J. Lee, J. Kim, and J. Yun, "M2M service platforms: Survey, issues, and enabling technologies," IEEE Commun. Surv. Tuts., vol. 16, no. 1, pp. 61–76, Fourth Quart. 2014

[73] F. Ghavimi and H. H. Chen, "M2M communications in 3GPP LTE/LTEA networks: Architectures, service requirements, challenges and applications," IEEE Commun. Surv. Tuts., vol. 17, no. 2, pp. 525–549, Second Quart. 2015.
AGIWAL et al.: NEXT GENERATION 5G WIRELESS NETWORKS 1653

[74] E. Yaacoub and Z. Dawy, "On using relays with carrier aggregation for planning 5G networks supporting M2M traffic," in Proc. IEEE Int. Conf. Wireless Mobile Comput. Netw. Commun. (WiMob), 2014, pp. 124–129

[75] R. Ratasuk, A. Prasad, Z. Li, A. Ghosh, and M. Uusitalo, "Recent advancements in M2M communications in 4G networks and evolution towards 5G," in Proc. Int. Conf. Intell. Next Gener. Netw., 2015, pp. 52–57

[76] M. Jo, T. Maksymyuk, R. L. Batista, T. F. Maciel, A. L. de Almeida, and M. Klymash, "A survey of converging solutions for heterogeneous mobile networks," in Proc. IEEE Int. Conf. Wireless Commun., vol. 21, no. 6, pp. 54–62, Dec. 2014

[77] E. Z. Tragos and V. Angelakis, "Cognitive radio inspired m2m communications," in Proc. IEEE Int. Symp. Wireless Pers. Multimedia Commun. (WPMC), 2013, pp. 1–5

[78] B. Xu, L. Da Xu, H. Cai, C. Xie, J. Hu, and F. Bu, "Ubiquitous data accessing method in IoT-based information system for emergency medical services," IEEE Trans. Ind. Informat., vol. 10, no. 2, pp. 1578–1586, May 2014

[79] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the internet of things: A survey," IEEE Commun. Surv. Tuts., vol. 16, no. 1, pp. 414–454, Jan. 2014

[80] G. Fortino, A. Guerrieri, W. Russo, and C. Savaglio, "Integration of agent-based and Cloud Computing for the smart objects-oriented IoT," in Proc. IEEE Int. Conf. Comput. Supported Coop. Work Design (CSCWD), 2014, pp. 493–498

[81] S. Nastic, S. Sehic, D.-H. Le, H.-L. Truong, and S. Dustdar, "Provisioning software-defined IoT cloud systems," in Proc. IEEE Int. Conf. Future Internet Things Cloud, 2014, pp. 288–295

[82] C. Turcu and C. Turcu, "The social Internet of Things and the RFIDbased robots," in Proc. Int. Congr. Ultra-Modern Telecommun. Control Syst. Workshops, 2012, pp. 77–83

[83] A. M. Ortiz, D. Hussein, S. Park, S. N. Han, and N. Crespi, "The cluster between internet of things and social networks: Review and research challenges," IEEE J. Internet Things, vol. 1, no. 3, pp. 206–215, Jun. 2014

[84] L. Atzori, A. Iera, G. Morabito, and M. Nitti, "The social internet of things (siot) when social networks meet the internet of things: Concept, architecture and network characterization," Comput. Netw., vol. 56, no. 16, pp. 3594–3608, 2012

[85] International Telecommunication Union. (2011). Internet of Things Global Standards Initiative (IoT-GSI) http://www.itu.int/en/ITU-T/gsi/iot/Pages/default.asp

[86] N. Kumar, S. Misra, J. Rodrigues, and M. Obaidat, "Coalition games for spatio-temporal big data in internet of vehicles environment: A comparative analysis," IEEE J. Internet Things, vol. 2, no. 4, pp. 310–320, Aug. 2015

[87] M. Gerla, E.-K. Lee, G. Pau, and U. Lee, "Internet of vehicles: From intelligent grid to autonomous cars and vehicular clouds," in Proc. IEEE World Forum Internet Things (WF-IoT), 2014, pp. 241–246

[88] K. M. Alam, M. Saini, and A. El Saddik, "Towards social internet of vehicles: Concept, architecture and applications," IEEE Access, vol. 3, no., pp. 343–357, Jan. 2015

[89] Y. Leng and L. Zhao, "Novel design of intelligent internet-of vehicles management system based on cloud-computing and Internet-of-Things," in Proc. IEEE Int. Conf. Electron. Mech. Eng. Inf. Technol., 2011, pp. 3190–3193

[90] X. Wu et al., "Vehicular communications using DSRC: Challenges, enhancements, and evolution," IEEE J. Sel. Areas Commun., vol. 31, no. 9, pp. 399–408, Sep. 2013

[91] K. Y. Ho and R. B. Ho, "A Bayesian game-theoretic approach for MAC protocol to alleviate beacon collision under IEEE 802.11 p WAVE vehicular network," in Proc. Int. Conf. Ubiq. Future Netw., 2014, pp. 489–494

[92] J. B. Kenney, "Dedicated short-range communications (DSRC) standards in the United States," Proc. IEEE, vol. 99, no. 7, pp. 1162–1182, Jul. 2011

[93] S. Cespedes, N. Lu, and X. Shen, "VIP-WAVE: On the feasibility of IP communications in 802.11 p vehicular networks" IEEE Trans. Intell. Transp. Syst., vol. 14, no. 1, pp. 82–97, Mar. 2013

[94] L. Goratti, S. Savazzi, A. Parichehreh, and U. Spagnolini, "Distributed load balancing for future 5G systems on-board high-speed trains," in Proc. IEEE Int. Conf. 5G Ubiq. Connect. (5GU), 2014, pp. 140–145

[95] P. Bonato, "Clinical applications of wearable technology," in Proc. IEEE Annu. Int. Conf. Eng. Med. Biol. Soc., 2009, pp. 6580–6583

[96] J. J. Rutherford, "Wearable technology," IEEE Eng. Med. Biol. Mag., vol. 29, no. 3, pp. 19–24, May/Jun. 2010

[97] P. F. Binkley, "Predicting the potential of wearable technology" IEEE Eng. Med. Biol. Mag., vol. 22, no. 3, pp. 23–27, May/Jun. 2003

[98] V. Oleshchuk and R. Fensli, "Remote patient monitoring within a future 5G infrastructure," Wireless Pers. Commun., vol. 57, no. 3, pp. 431–439 2011

[99] L. E. Dunne et al., "The social comfort of wearable technology and gestural interaction" in Proc. IEEE Ann. Int. Conf. Eng. Med. Biol. Soc., 2014, pp. 4159–4162

[100] Z. ZhiYing, Q. Lingzhen, and J. Yu, "Study on application of grid computing technology in financial industry," in Proc. IEEE Int. Forum Inf. Technol. Appl.,2009, vol. 2, pp. 344–346

[101] Accenture, "Rise of fintech," White Paper, 2014

[102] M. Erol-Kantarci and H. T. Mouftah, "Energy-efficient information and communication infrastructures in the smart grid: A survey on interactions and open issues," IEEE Commun. Surv. Tuts., vol. 17, no. 1, pp. 179–197, First Quart. 2015

[103] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid the new and improved power grid: A survey," IEEE Commun. Surv. Tuts., vol. 14, no. 4, pp. 944–980, Fourth Quart. 2012

[104] A. Roy, H. Kim, N. Saxena, and R. Kandoori, "LTE multicast communication for demand response in smart grids," in Proc. IEEE Int. Conf. Adv. Netw. Telecommun. Syst. (ANTS), 2014, pp. 1–6

[105] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," IEEE Commun. Surv. Tuts., vol. 17, no. 1, pp. 152–178, First Quart. 2015

[106] K. E. Skouby and P. Lynggaard, "Smart home and smart city solutions enabled by 5G, IoT, AAI and CoT services," in Proc. Int. Conf. Contemp. Comput. Informat., 2014, pp. 874–878

[107] V. O. Tikhvinskiy, S. V. Terentiev, and V. P. Visochin, LTE/LTE Advanced Mobile Communication Networks: 4G Technologies, Applications and Architecture, Moscow: Media, 2014

[108] Network Access & Identity. Trusted Computing Group. Available from: http://www.trustedcomputinggroup.org/solutions/endto end_trust/, 19 June 2015

[109] Trusted Network. Kaspersky Internet Security, 2005. Available from: http://support.kaspersky.com/6423

[110] ICT-317669-METIS/D6.6. Final report on the METIS 5G system concept and technology roadmap, Project METIS Deliverable D6.6, 2015

[111] V. O. Tikhvinskiy and G. Bochechka. Perspectives and quality of service requirements in 5G networks, Journal of Telecommunications and Information Technology 1: 23–26, 2015

[112] W. Yin. No-Edge LTE, Now and the Future 5G World Summit, 2014. Available from: http://ws.lteconference.com/

[113] P. Yongwan. 5G Vision and Requirements of 5G Forum, Korea, 2014

[114] V. O. Tikhvinskiy, G. S. Bochechka, and A. V. Minov. LTE network monetization based on M2M services, Electrosvyaz 6: 12–17, 2014.

[115] B. Sam. Delivering New Revenue Opportunities with Smart Media Network. 5G World Summit 2014, Amsterdam, 2014

[116] Dr. Fei Hu, "Opportunities in 5G Networks", Section II, 6. QoS in 5G Networks

[117] Y. Chen, et al., Fundamental Trade-offs on Green Wireless Networks, IEEE Communications Magazine, Vol.49, No.6, 2011, pp. 30-37

[118] A. Antonopoulos, et al., Energy-Efficient Infrastructure Sharing in Multi-Operator Mobile Networks, IEEE Communications Magazine, Vol.53, No.5, 2015, pp. 242-249

[119] A. Khan, et al., Network Sharing in the Next Mobile Network: TCO Reduction, Management Flexibility, and Operational Independence, IEEE Communications Magazine, Vol.49, No.10, 2011, pp. 134-142

[120] J. Xu, J. Duan, R. Zhang, Cost-Aware Green Cellular Networks with Energy and Communication Cooperation,

IEEE Communications Magazine, Vol.53, No.5, 2015, pp. 257-263.

[121] L. Gkatzikis, I. Koutsopoulos, T. Salonidis, The Role of Aggregators in Smart Grid Demand Response Markets, IEEE Journal on Selected Areas in Communications, Vol.31, No.7, 2013, pp. 1247-1257

[122] H. Q. Ngo, E. G. Larsson, T. L. Marzetta, Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems, IEEE Transactions on Communications, Vol.61, No.4, 2013, pp. 1436-1449

[123] E. Larsson, et al., Massive MIMO for Next Generation Wireless Systems, IEEE Communications Magazine, Vol.52, No.2, 2014, pp. 186-195

[124] A. Prasad, et al., Energy-Efficient InterFrequency Small Cell Discovery Techniques for LTE-Advanced Heterogeneous Network Deployments, IEEE Communications Magazine, Vol.51, No.5, 2013, pp. 72-81

[125] N. Bhushan, et al., Network Densification: The Dominant Theme for Wireless Evolution into 5G, IEEE Communications Magazine, Vol.52, No.2, 2014, pp. 82-89

[126] G. Wu, et al., Recent Advances in Energy Efficient Networks and their Application in 5G Systems, IEEE Wireless Communications, Vol.22, No.2, 2015, pp.145-151

[127] D. Wu, et al., Energy-Efficient Resource Sharing for Mobile Device-to-Device Multimedia Communications, IEEE
Transactions on Vehicular Technology, Vol.63, No.5, 2014, pp. 2093-2103

[128] D. Wu, et al., The Role of Mobility for D2D Communications in LTE-Advanced Networks: Energy vs. Bandwidth Efficiency, IEEE Wireless Communications, Vol.21, No.2, 2014, pp. 66-71

[129] C. Bae, W.E. Stark, End-to-End Energy Bandwidth Tradeoff in Multihop Wireless Networks, IEEE Transactions on Information Theory, Vol.55, No.9, 2009, pp. 4051-4066

[130] B. Smida, et al., A Spectrum-Efficient Multicarrier CDMA Array-Receiver with Diversity-Based Enhanced Time and Frequency Synchronization, IEEE Transactions on Wireless Communications, Vol.6, No.6, 2007, pp. 2315-2327