UNIVERSIDAD CARLOS III DE MADRID

SCHOOL OF ENGINEERING

BACHELOR'S THESIS

TECHNO-ECONOMIC ANALYSIS OF A 5G NETWORK IN SPAIN

BACHELOR'S DEGREE IN TELECOMMUNICATION TECHNOLOGIES

AUTHOR: **JUAN RIOL MARTÍN**

ACADEMIC TUTOR: **RAQUEL PÉREZ LEAL**

PROFESSIONAL TUTOR: **JULIO NAVÍO MARCO**

JULY OF 2017

Acknowledgements

This thesis could not have been completed without the priceless help of some people who I would like to mention specifically for everything they have done.

Firstly, I would like to express my sincere gratitude to my advisors/tutors for the continuous support on my thesis, for his patience, motivation, and immense knowledge. Thanks to my academic tutor Raquel Pérez Leal for her exhaustive tracing and corrections. Thanks to my professional tutor Julio Navío Marco, who gave me an overall vision, and helped me to focus on what was important for this project. Thank you for giving me the opportunity to work on a topic I like this much. My sincere gratitude to both of them for their enthusiasm, support and interest in being so involved in this project.

My sincere thanks also go to Sergio Hidalgo, Damián Rinconada and Sergio Mayoral, not only friends but also excellent professionals with whom I have approached this project from different perspectives. Also thanks to Victor Eugenio Martín Ugena, brilliant student, who has always helped and supported me, and has done it with this project as well.

Last but not least, I would like to thank mi family and girlfriend Marta. They have always supported me throughout writing this thesis. Thank you for your patience, understanding and support not only in the writing of this thesis but in every aspect and challenge I address in my life.

Abstract

Information society and mobile society are two concepts that are both linked and undeniable. The first one refers to the necessity of high amount of information to develop most aspects of our lives, while the second one is related to the importance of mobile devices to get, analyse and use that information. In other words, every mobile device (that embraces not only mobile phones but also many other gadgets) has become a tool that shall interact with information.

In order to fulfil those needs, technology has evolved, resulting into faster, more secure and more reliable networks. Needless to say, mobile networks are playing an indispensable role, as long as the society is evolving to a more and more mobile one, as above mentioned. Furthermore, new applications that had not been even imagined years ago must be fulfilled as well (i.e. smart cities).

There are many industries that carry the weight of this progress. Companies of various sectors of our economy must develop each piece of the puzzle to ensure that the jigsaw is solved. Another important player should not be forgotten. The regulatory institutions and frameworks must coordinate all this investigations and progress in order to assure the universality, integrity and reachability of itself.

The purpose of this document is to consider what the mobile communications needs of today's society are, what they will be on a short, mid and long run, and how can they be solved. To face this task, the two main actors above mentioned will be taken into account. From the regulatory perspective, the proposals and law measures (i.e. IMT-2020 and new frequency allocations) must be considered, as well as the technical requirements for 5G generation, whether to be considered the subsequent evolution of LTE network or a new network, or even both. From the mobile companies' point of view, a dense analysis on technical solutions to reach the above mentioned requirements will be followed by an economic analysis to discuss the profitability of the deployment of a 5G network.

It must be understood that this study contemplates several scenarios, due to the different possibilities in terms of the spectrum policies and demand evolution in the forthcoming years. To this end, the several scenarios combined with the different cases of use must be taken into account, as well as many other KPIs. The coherent combination and analysis of all this parameters will reveal the requirements' feasibility amongst varying scenarios.

<u>Key-words</u>: 5G, costs, IMT-2020, LTE, mobile communications, spectrum policies, standardizations, technical rate, techno-economic analysis,

Index of contents

List of figures

List of tables

List of acronyms

3

3GPP: 3rd Generation Partnership Project, 12

A

AENOR: Asociación Española de Normalización y Certificación, 12
AF: Application Function, 50
ARIB: Association of Radio Industries and Businesses, 17
AMF: (Core) Access and Mobility Management Function, 50
ATIS: Alliance for Telecommunications Industry Solutions , 17
AUSF: Authentication Server Function, 50

B

BGP: Bit Gate Protocol, 44
B(T)S: Base (Transfering) Station, 17

C

C-RAN: Cloud/centralized RAN, 39
CAGR: Compound Annual Growth Rate, 83
CAPEX: Capital Expenditures, 40
CEPT: Conférence Européenne des administrations des Postes et des Télécommunications, 12
CN: Core Network, 42
CNMC: Comisión Nacional del4 Mercado de la Competencia, 92
CNAF: Cuadro Nacional de Atribución de Frecuencias, 12

D

D-RAN: Distributed RAN, 39
DC: Deployment Costs, 73
DN: Data Network, 450
DSM: Digital Single Market, 29

E

E2E: End to Ends, 42
EC: European Commission, 24
EC (2): Enhancing Costs, 73 (from chapter 4 onwards)
EDGE: Enhanced Data rates for GSM Evolution, 2
EHF: Extremely High Frequency, 20

K

L

M

N

O

P

Q

R

S

T

U

V

W

1. Introduction

The first mobile generation based on analogic systems was released in the 80's. Since then, the uprising necessities have triggered every single breakthrough in this (and many other) field. Needless to say, mobile communications have evolved parallel and complementary along with other types of communications, such as Internet or ISDN. This dynamic has led to an undeniable reality: many daily activities, demarches or needs have been transferred to the net, making communications a vital pillar in everyone's life.

At this point, there is a loop that explains the meaning and necessity of improvements in our communications systems: the more activities are driven into the net, the better, more reliable and efficient it must be. Therefore, the better the net is, the more activities are susceptible to be digitalized and driven into the net. To this purpose, communications must be improved at least at the same level as nets themselves so as to guarantee the access not only among users but also among things (i.e. devices). Nowadays, this trend seems to be the direction society will follow in a long run, and therefore, as the loop goes on, so has to do the progress.

Mobile communications have become an important (maybe the most one) pillar, what seems quite logical. The first devices with communication capacity where fixed ones. The limitation of movement made them inefficient. Hence, mobility was an important milestone to achieve within the first generations of mobile communications. Once achieved, it was irrefutable to think that the convenience and easement accomplished would be essential in future generations.

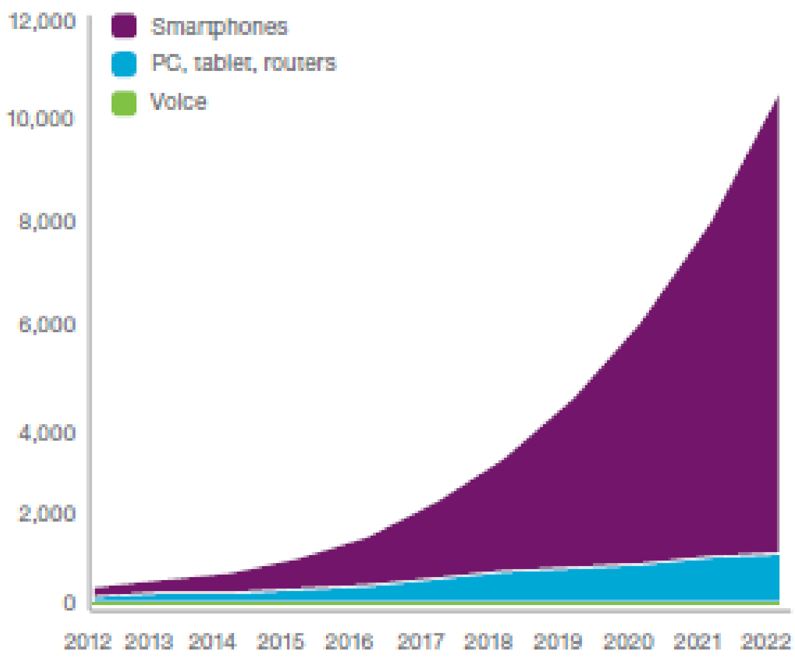


Figure 1. Mobile traffic per month in Western Europe (PB).Source: Ericsson [1]

As shown in Figure 1, it is evident that all the traffic due to smartphones represents the overwhelming majority of it, and will even grow in percentage and number in the future years. As explained above, it is crucial to have networks that ensure that all the subscribers of these technologies receive the proper service.

As for mobile communications, it is important to understand the importance of the history of generations which have been evolving until today. This investigation line has created the technology (both hardware and software) to satisfy the challenges alongside the timeline.

As firstly explained, the first generations (0G and 1G) were based on analogic systems and circuit switching. Due to the difficulties related to interoperability and scalability, the next generation achieved the milestone of becoming digital. The next system triggered the revolution of mobile communications: it was named GSM (Global System Telecommunications) and guaranteed important services such as ciphered calls, full-duplex communications… The rapid demand growth manifested that new solutions were needed to bear the new challenges of communications. This 2G system evolved into GPRS (Global Packet Radio Access) and EDGE (Enhanced Data rates for GSM Evolution), adding packet switching techniques.

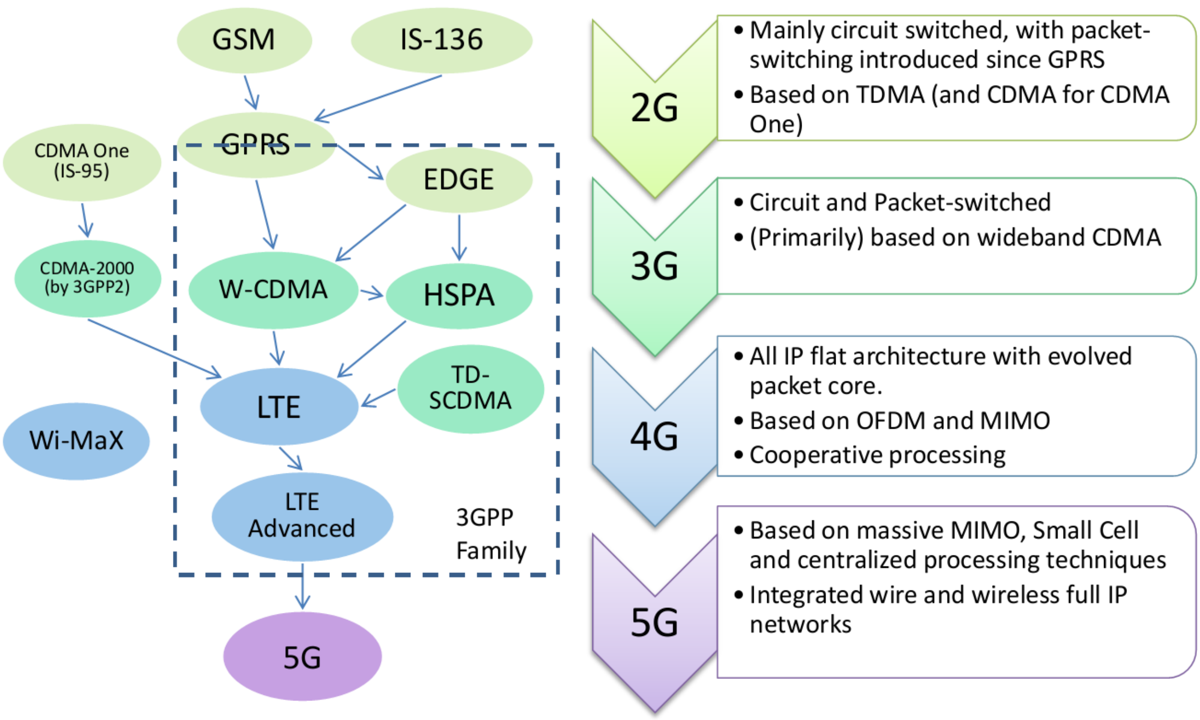


Figure 2. Evolution of mobile generations, Source: ajTDE [2]

Next step was to achieve significant bit rates and broadband internet connection using the packet domain to connect to the net and the circuit domain for voice calls. This was made possible by UMTS (Universal Mobile Telephone Service), which belongs to 3G.

The most recent breakthrough has been the consecution of efficiency and compatibility targets as well as higher binary rates to meet the users' demands. This has become a reality thanks to 4G requirements and its systems, mainly LTE (Long Term Evolution) and also WiMAX. The next step to make will be analysed in this document: 5G.

As it can be seen in figure 2 generations are modelled as evolutions one from another. This shows the importance of standardisation bodies in this study area, which carry the evolutions upon their responsibility. Most of the telecommunication systems throughout history have come as an aggregate of new technologies plus enhancement of existing ones.

Parallel, another investigation line focused on the need of new and more frequency bands to enhance the capability of networks. Needless to say, more frequencies result in more capability. Here lies one of the most limiting factors of mobile communications. It is quite obvious that the physical medium that mobile communications use to transmit is the radio-electric spectrum. This spectrum holds two important drawbacks to be considered:

1. **Limited**: the spectrum has a finite number of band frequencies. Furthermore, each band has some pros and cons that imply strategy and planification to make the most of each band: propagation, modulation techniques, power… are some factors to consider when choosing one frequency or another.

2. **Shared**: there are other services which use the radio-electric spectrum as its transmission medium. Therefore, they all share their common medium and harmonization is deeply needed in order to avoid interferences amongst them. Furthermore, the next implication is that each system cannot dispose from the desired bands.

As a conclusion, it can be summed that there are two main fields that collaborate narrowly to keep up to this unstoppable growth. On the one hand, the technology improvements and evolution have to enhance everyday our technology, enabling it to hold the increase of demand and its requirements. Moreover, it has some constrains, the main of which is the availability of the radio-electric spectrum. On the other hand, the bodies in charge of frequency allocations must try to harmonize the spectrum so as to provide the smartest and most efficient access to each one of the services which need this physical medium.

1.1 Motivation

The motivation behind this work is vast, as for the nature, content and procedure of itself. It can be best divided into two types, which are described in this section.

1.1.1 Professional motivation

As a plausible telecommunications engineer, mobile communications cope one of the main work areas amongst our field. Furthermore, it implies the application of knowledge acquired through all these years both in theoretical and practical concepts. It also provides the student of a unique tessitura as having to analyse the problem from its all different points of view, which are not few. Many agents are involved in this issue, from service providers to regulatory bodies, including users, intermediates, governments and standardization bodies. Each one of this agents holds its own interests, which may be opposite from others' ones (e.g. competence between service providers, spectrum policies to benefit one or other sector).

The problem can be approached from all those perspectives, and with different purposes: sizing the future networks, determining the profitability/viability of the deployment, discussing the consequences and implications of political decisions in terms of band allocations… All this combinations imply, first, a deep analysis of the Key Performance Indicators (KPIs) and their implications to the agents, and secondly the use of almost every concept learned at college, hence professional motivation is incredibly stimulated.

1.1.2 Personal motivation

Two main reasons explain the personal motivation that lies in this work:

Firstly, what is being driven under study is the immediate and mid-term future. This means that what will be discussed here will be proven in two to five years when this future fully arrives at last. It is quite motivating to analyse and reason what the directions enhancements in communications should follow are, and why. More specifically, 5G is setting high expectations and has postulated itself as a true revolution in mobile communications. Therefore, analysing whether these expectations will be able to be filled, how to achieve them, and why will they be necessary, seems quite interesting.

Secondly, this problem constitutes a dense typical engineering problem, in which all of the answers are not optimum and constitute a compromise among factors such as quality, costs… Furthermore, as above mentioned, all these sub-optimum answers may be analysed from different perspectives regarding to all the agents mentioned. Therefore, not only the solutions but also the classification of themselves seems demanding and a source of motivation.

1.2 Objective

The main goal for this project is to make a deep analysis of the techno-economic viability and possibilities of a future 5G network. More in detail, the aim will be to determine the technical features and requirements that 5G networks shall have in terms of architecture and functionality in order to fulfil the different uprising applications (e.g. virtual reality, Internet of Things, etc.). To analyse it, methodology will be based on two types of conditioning factors:

1. **Cases of use:** these cases of uses will be defined by the specifications announced by regulators. Each case of use refers to a desired application of 5G networks. These applications will be diverse and will have different requirements (low delay, spectrum efficiency or mobility among others).

2. **Deployment scenarios:** scenarios are used to provide different possibilities when we don't know yet how something will be like in the future. They represent a generalisation of a possible future option. In mobile communications, scenarios are often related to demography: they represent different demographic options such as big cities, rural environments, etc. For this analysis, scenarios will be based on deployment options. This means that they will contemplate different option when deploying future 5G networks.

Hence, the aim is to know what the 5G most important innovations and technologies are, determine which the best deployments scenarios-cases of use options are, and reason which of the mentioned technologies can best address each option.

It is **not** the aim of this project to propose a particular architecture or define elements such as the LTE's eNodeBs. It is more focused on proposing different solutions, both technical (i.e. understanding requirements of each case of use) and economical (i.e. reuse and enhancement of some parts of LTE networks vs. deploying new 5G networks) to address the different questions that 5G presents.

1.3 Project planning and budget forecast

This section will analyse the planning that the development of this project is intended to follow as well as the forecasted necessary budget to carry it out.

1.3.1 Project planning

Planning is disaggregated into weeks. This project will be developed from January to June 2017. The project will be therefore developed in a total of 24 weeks, **18 of which will be considered as activity weeks.** First, the tasks description is going to be disaggregated:

1. Phase 1- Problem addressing
    A. Information searching
    B. Documentation analysis
    C. Problem understanding & approach
2. Phase 2- Project's outline design
    D. Objectives definition
    E. Document structuring
3. Phase 3- Study Development Part I: Theory (Chapters 1-3)
    A. Document writing
4. Phase 4- Study Development Part II: Analysis(Chapters 4-6)
    G. Technical Analysis
    H. Economic analysis (Matlab software programming included)
    I. Result analysis & contrasting and conclusions
    J. Final revision

Next table displays the Gantt diagram that the project will follow. Note that red columns represent weeks in which the project development will be paused, while green cells represent a task development within that week. Note also that activities are identified by the corresponding letter on the aforementioned list.

Table 1. Project planning: Gantt Diagram

	January				February			March					April				May				June		
Weeks[1]	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Phase 1- Problem addressing																							
A																							
B																							
C																							
Phase 2- Project's outline design																							
D																							
E																							
Phase 3- Study Development Part I: Theory (Chapters 1-3)																							
F																							
Phase 4- Study Development Part II: Analysis(Chapters 4-6)																							
G																							
H																							
I																							
J																							

1.3.2 Project budget

This section will calculate the total budget to carry out this study. First premise is to calculate the amount of time dedicated to this project. Eighteen weeks is the total duration of the project. In each week, approximately 4 hours per day

[1] Week numeration is the week number in the whole year.

in 5 labour days per week will be dedicated to the project. This makes a total of **360 hours**. There are some premises on the considered costs:

a. *Staff costs*. Assuming the total dedication of 360 hours to this project:
   - A junior telecommunications engineer's fee is 30€/hour, which makes a total of 10.800 € for junior engineer Juan Riol Martín.
   - A senior telecommunications engineer's fee is 70 €/hour. Assuming that the time dedicated to this project is 36 hours for each senior engineer, it makes a subtotal of 2.520€ for each senior engineer: Julio Navío Marco and Raquel Pérez Leal.
   - Indirect costs will be calculated as a 15% of staff costs.

b. *Material costs*
   - No office rental will be considered. Therefore, electricity and internet connexion will be considered as direct costs. As they are paid monthly regardless of the amount of work or holiday weeks, time period will be considered 6 months. Imputable percentage will be considered 20%.
   - Amortization of a personal computer is considered to be 5 years.

Secondly, they will be displayed in next table. For the indirect cost, the percentage imputed will be written in the corresponding row.

Table 2. Project budget

		Concept	Amount	Subtotal
Staff Costs	Direct	*Senior engineering fee: Raquel Pérez*	70 €/hour	**2.520 €**
		Senior engineering fee: Julio Navío	70 €/hour	**2.520 €**
		Junior engineering fee: Juan Riol	30 €/hour	**10.800 €**
	Indirect	*Indirect staff costs*	15% of staff costs	**2.376 €**
	Subtotal	Total staff costs		**18.216 €**
Material Costs	Direct	*Personal computer (amortized in 5 years)*	1.500 €	**150 €**
		Office supplies	30 €	**30 €**
		Electricity	45 €/month	**54 €**
		Internet connexion	35 €/month	**42 €**
	Subtotal	*Direct material costs*		**276 €**
Subtotal	All	*Entire subtotal: staff + material costs*		**18.492 €**
VAT (21%)		*VAT over total costs*	*21 %*	**3.883,32 €**
TOTAL		*Subtotal + VAT*		**22.339,02 €**

1.4 Introduction to 5G

First of all, a clarification must be done. There is a difference between a generation (5G) and a telecommunication system (LTE, UMTS, GPRS…). On the one hand, generation is a set of requirements in terms of bandwidth, bit rate, latency and many other parameters that a standardization body (typically ITU, see section 2.1) specifies. On the other hand, a telecommunication system is an aggregate of SW and HW technologies that fulfil the proposed requirements. This aggregate consists of standards and protocols as well as hardware devices and equipment capable of ensuring the desired requirements.

Therefore, every manufacturer has to develop equipment compatible with one or more telecommunication system (depending on the usage). On another role, network developers must deploy the network plus the aggregate system (mentioned in the previous paragraph) to which devices will connect. As an example, a mobile phone nowadays must be compatible with LTE, Wi-Fi (IEEE 802.11 standard), Bluetooth (IEEE 802.15) or GPS (Global Positioning System) among other telecommunication systems. Each telecommunication system provides a different functionality and has its own network, many of which can be and are actually interconnected.

As shown in Figure 2, 5G is being modelled as a natural evolution of LTE. This is an important fact as it shall be compatible backwards in order to ensure that these two systems can be used simultaneously and complementarily. Furthermore, LTE-A (LTE Advanced), which is different from 5G, will play a capital role in this feasible future. Hence, coexistence, collaboration and compatibility between 5G and LTE-A are key concepts to understand what mobile communications will be like in the next years.

Another important fact by the date of this document is that 5G is still not well defined in terms of what it will be like. This means that many network elements are still unknown, and it is still uncertain what parts of the LTE network will be able to be reused and/or enhanced.

Nevertheless, it is known what the requirements for 5G will be. ITU-R (International Telecommunications Union) has released an important project called IMT-2020 (International Mobile Telecommunications for 2020) in which all these specs appear. To fully understand IMT-2020, please go to section 2.1.1.

Thus, the overview of 5G sets two development stages:

1. **Enhancement of LTE to LTE-A and LTE-A Pro** in order to fulfil some of the requirements raised by IMT-2020 and complementary features (latency, encryption…).

2. **Development of a new Radio Access Network (new RAN)** which is still undefined that must guarantee by 2020 the requirements above mentioned.

Every new generation in mobile communications comes as an answer not only to the enhancement of currently existing technical features but also to the changing types of demand. They come together with new necessities and tendencies that users ask for.

The fifth generation is no exception to this assertion. Although it will be deeply addressed in chapter 4, figure 3 reveals the demand evolution lines according to the ITU-R group (Radiofrequency group of ITU).

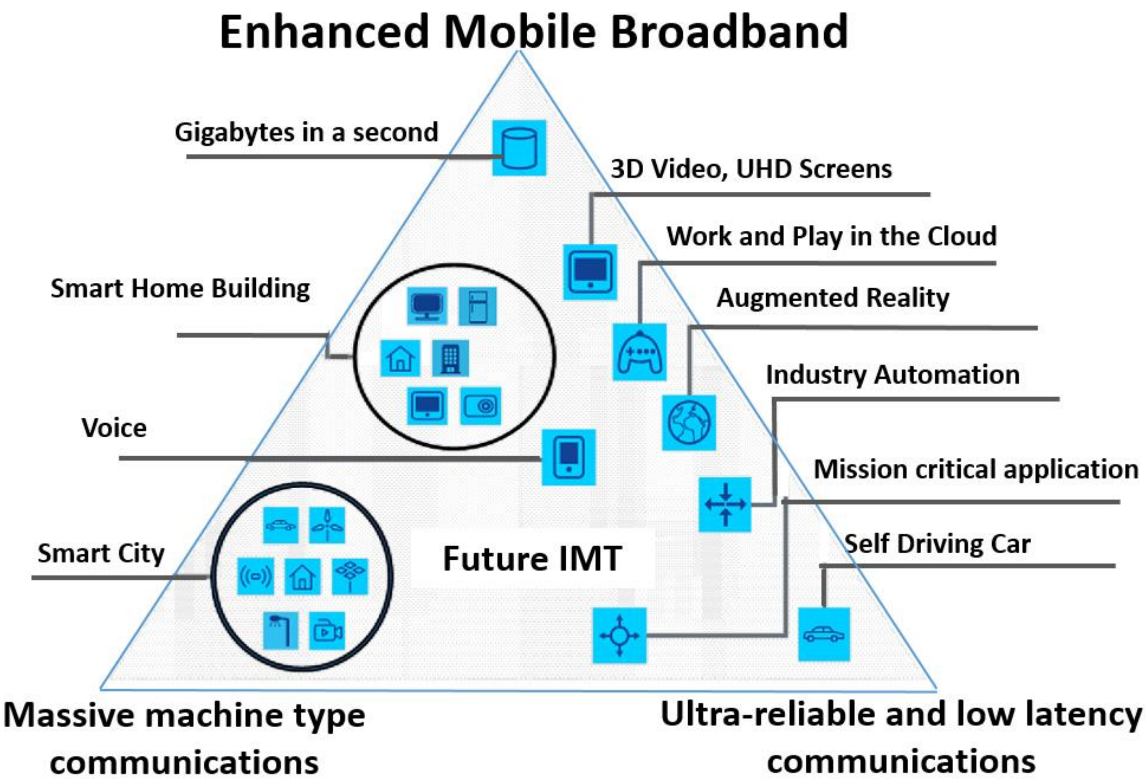


Figure 3: Groups of usage scenarios for IMT-2020. Source: ITU-R, 2015

As it can be seen, users and therefore other vertical industries (automobiles, energetic industry, videogames…) determine the questions, and communication industry must search for the answers. In figure 3, all those varied needs are grouped into sets that imply similar technical features or can be solved similarly.

In chapter 4, these groups of usage will be analysed thoroughly, as well as their intrinsic needs and features.

1.5 Document structure

This document has been thought to have a coherent structure so as to facilitate its comprehension and reading as well as to comply with the compulsory headlands. In order to accomplish this, the order and issues faced in the different chapters (including the introduction and project planning & budget, which have already been addressed) will be the following ones:

In chapter 2, the state of art will be deeply addressed. This chapter will be divided into the considered most important aspects among it: standardization and regulation. Inside standardization, it will be analysed the bodies in charge of this task. Besides, their collaboration and parallel interworking will also be dissected. On the other hand, regulation is extremely important (**regulatory framework is studied in this chapter**). In this case, spectrum regulation is the one to determine the paths to follow in terms of equipment development, software programming, analysis and planning, etc. due to the importance of the transmission media. Needless to say, standardization and regulation would not have any sense one without the other. They both create the rules which the other players involved follow. That's the reason why they are analysed together.

Chapter 3 is fully dedicated to treat the technical considerations for the deployment of a network that shall fulfil 5G requirements. Inside this chapter there is an important section referred to the premises, in which the network segments are explained as well as LTE architecture, whose legacy and enhancement will be crucial to achieve 5G goals. This will facilitate the comprehension of the three remaining sections inside this chapter, which will study the innovations in each segment of the future 5G networks.

This will end the theoretical part of the document. These chapters will enable the author to develop a coherent analysis and especially to extract coherent conclusions once obtained the results.

Chapter 4 is the most important chapter along with chapter 5. In this chapter relays the weight of the thesis, hence it treats the techno-economic analysis indeed. It justifies the solution methodology, based in the proposal of scenarios and cases of use crosses. Technical and economic analyses are carried out parallel to be contrasted at the end of the chapter.

In chapter 5, the results obtained from chapter 4 are gathered with real demand forecasts. By doing this, the optimum deployment lines are selected. **Chapter 5 comprises both the socio-economic environment an impact**.

To conclude, chapter 6 talks about the conclusions on methodology and results (result contrast) and feasible future work upon this topic that will come alongside the future breakthroughs (both in technology and standardization).

2. State of art

In this chapter, the two most influential aspects that determine the development of mobile communications and in this case 5G are studied under scrutiny: the spectrum and technology states of art. As it was explained in the previous chapter, these two aspects would not have any sense one without the other. Therefore, there will be many crossed references, as a consequence of the narrow collaboration between standardization bodies and spectrum regulators. In each of the sections, there will be a paragraph dedicated to the EU's situation.

Another important working line is focused on technology development. It can be considered as another part of the state of art. Nevertheless, it will be fully explained in next chapter (deployment of a 5G network) due to organizational issues.

2.1 Standardization and technology state of art

Technological progress would not exist without standardization. Standardization consists in creating the indispensable and mandatory rules that any kind of innovation must comply with. This process harmonizes separated researches in order to create a common direction to follow. Furthermore, it economizes and coordinates efforts, as long as the different agents have some constraints that make them focus their jobs.

Standardization must be differentiated from regulation, last of which is in charge of legislative aspects (i.e. political laws, disposals, etc.). Regulation is carried out by public entities (i.e. governments, supranational institutions…). Both regulators and standardization bodies elaborate either or both types of documents, may they be technical or legislative:

- **Mandatory rules**: this regulation/standardization type is compulsory. It is often (but not only) carried out by those institutions belonging to a government or legislative power. These rules often define standards in any field, may them be technical (such as Wi-Fi from IEEE) or quality ones (like ISO ones). In any of the situations, they are very concise rules that leave no room to doubt when adjusting to them.

- **Recommendations**: this regulation type does not set any predefined rule to follow, but it does set recommendations, advices or proposed requirements that are very often taken into account. Depending on the reputation of the body, its recommendations are more or less taken into account. The content of this regulation/standardization type documents is usually vaguer, and it sets some priorities, investigation lines and proposed milestones that other agents should achieve.

It seems quite obvious that, as the rest of agents have to do, also these bodies must work complementarily, not only to achieve higher levels of progress but also to facilitate the rest of agents' work. It must be understood that their task is completely necessary and indispensable. Once understood their importance, next step is to see who they are indeed:

Table 3. Regulators classification

	NATIONAL (SPAIN)	INTERNATIONAL
RECOMMENDATIONS		ITU CEPT 3GPP, 5GAMERICAS
MANDATORY RULES	CNAF AENOR	ISO IEEE WARC European Commission-ETSI

Table 1 shows the main agents classified upon region of effect and commitment in the field of communications[2]. Those underlined are the ones that take a heavier responsibility in mobile communications and will be further analysed in next paragraphs.

2.1.1 ITU: IMT-2020 & IMT-Advanced

The International Telecommunications Union (ITU) is the most important standardization body all around the world in the field of telecommunications. It is formed by 193 member countries and over 700 private entities and academic institutions [3]. It divides itself into three areas of activity which are:

• **ITU-T**: the telecommunications standardization sector is in charge not only of making standards in form of recommendations, but also of revising and updating existing ones.

• **ITU-R:** the radio-communications sector coordinates this vast and growing range of radiocommunication services, as well as the international management of the radio-frequency spectrum and satellite orbits. An increasing number of players needs to make use of these limited resources, and participate in ITU-R conferences and study group activities – where important work is done on mobile broadband communications and broadcasting technologies [4]

• **ITU-D**: the development sector is oriented towards different projects to achieve social responsibilities ITU has acquired such as the digital breach or the digital agenda that countries worldwide have as an objective.

[2] Acronyms of every regulator are explained in the glossary

Although each group has its own hierarchy and working groups, it is not rare to see them collaborating to face huge projects that can be dissected into smaller tasks, each of them can be solved separately. This is the procedure that is being followed with the project that ITU launched in relation to 5G: IMT-2020.

IMT-2020 is the acronym for International Mobile Telecommunications by 2020. Its main purpose is to set the roadmap of mobile telecommunications by 2020, following their working structure based on recommendations (standards from ITU-T) and planning on future frequency allocation possibilities. It is important to understand their user oriented vision, what makes them look into the trends and needs among users to define what they will need in the future years.

To that purpose, ITU-R has created the 'Working Party 5D' (WP 5D) [3], which is specialized in defining and addressing what the milestones should be. This group is focused in this IMT-2020 project, and has detailed the following timeline to face all the phases of the project (see figure **6**).

The figures pretended to reach by IMT-2020 are shown in figure 4.

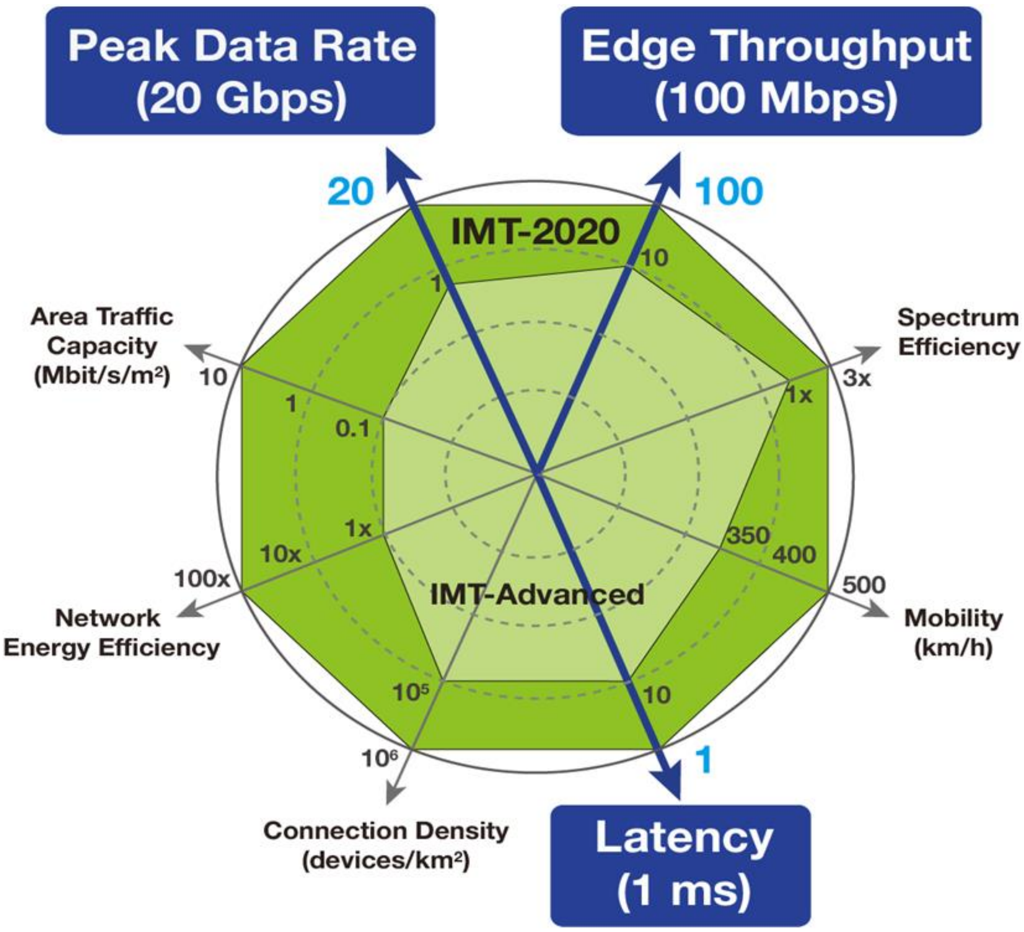


Figure 4. IMT 2020 requirements. Source: ITU-R,2015[5]

[3] This working party collaborates with other ITU groups as well as other regulators and companies.

As it can be deduced from figure 4, there are 8 KPI's in IMT-2020. These KPI's were defined so as to focus on the critical aspects of 5G networks cases of use and applications. These KPI's are[4]:

• **Peak Data Rate (Gbps):** it is defined as the maximum speed a user could enjoy in optimum conditions: being in the centre of the cell, having the bandwidth fully dedicated, no noise or interference…

• **Edge Throughput = user experienced data rate (Gbps):** this makes reference to the speed a user would sense in non-optimal conditions, like the edge of the cell. In a cellular system, edges between cells have the worst conditions in terms of speed and coverage. It seems quite obvious to understand that, given no best conditions, edge throughput will be much smaller compared to the peak data rate.

• **Spectrum Efficiency:** it is defined as a quality factor related to how the band is exploited, considering re-use of frequencies and other techniques. It is represented as a percentage because it is quite difficult to determine the absolute figure, but it shall be viable to determine whether future networks can be 3 times more efficient compared to LTE.

• **Mobility (km/h):** mobility is the maximum speed at which a device can roam from one radio node to another without losing service or a defined QoS (Quality of Service). The importance of mobility in uninterrupted calls or real time applications is indispensable.

• **Latency (ms):** this refers to the time that a packet with information takes to go from transmitter to receiver. This time is set by the RAN. This KPI is hugely important in real time and/or critical cases of use. Latency can be triggered by many factors coming from any process within the data transmission, but the aim is to reduce it to milliseconds.

• **Connection Density (devices/km^2):** it is defined as the maximum number of devices per unit area. Needless to say, this depends on the cells' size, as well as the spectrum efficiency. As said in the previous chapter, 5G is highly oriented to IoT, therefore it's quite important to know how many devices will be able to be served.

• **Network Energy Efficiency:** energetic efficiency is another KPI, as one of the milestones of 5G is to reduce OPEX (Operating Expenditure), many of which relay on energy costs. There also exists energy efficiency from the devices' side, which is defined similarly.

[4] Some numbers differ between figures 4 & 5. When necessary, operate to see that equivalences are the same despite being displayed in different units.

• **Area Traffic Capability (Mbps/m²):** this is another measure of the network capabilities. In this case, it is referred to speed per unit of area instead of device in the same area unit. This is other important parameter, often taken into account to determine the number of antennas and the cells' size.

These milestones are quite ambitious, but in order to fully understand the current state of art, figure 5 displays together the actual situations with LTE systems (that belong to 4G networks). This picture is commonly known as "The 5G Flower".

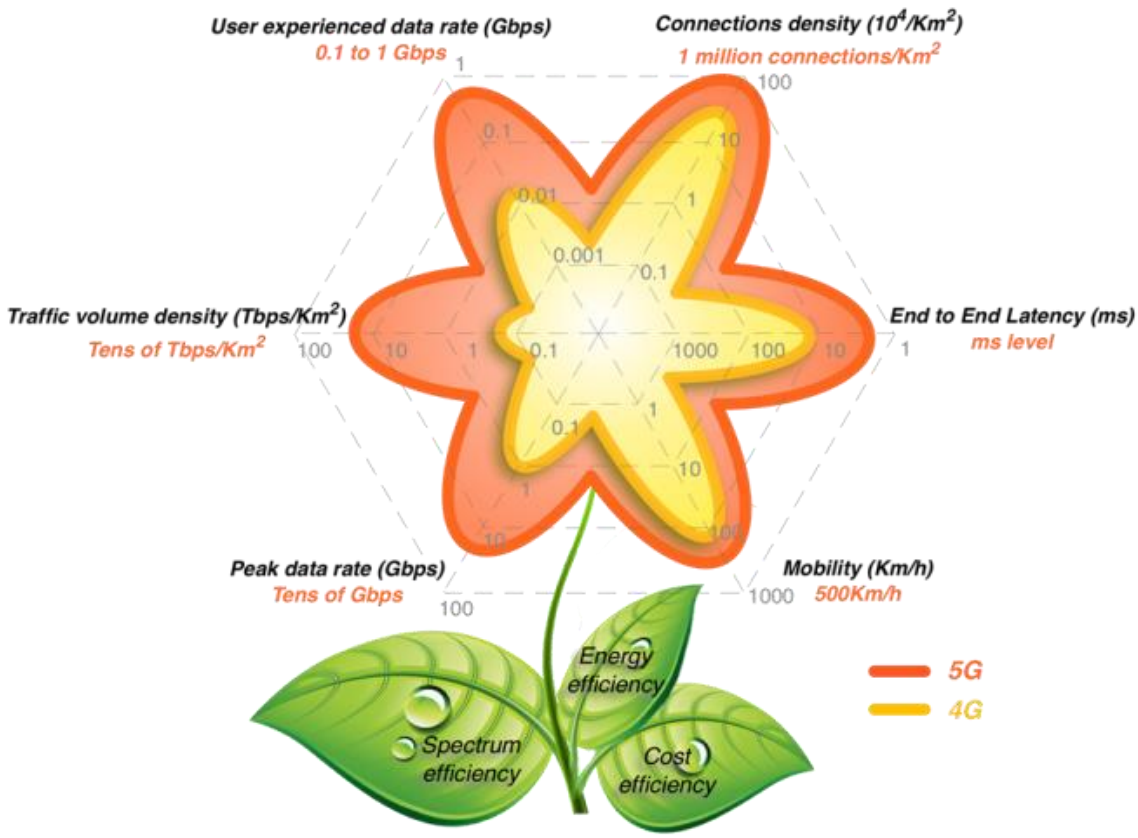


Figure 5. The 5G Flower. Source: ITU-R,2015

Some considerations should be done upon figure 5.

• 4G (the fourth generation) is mainly formed by LTE systems. Nevertheless, there are other telecommunication systems such as WiMAX that are not indicated in figure 2 for importance reasons. LTE-A belongs to 4G as well, therefore the best figures (drawn with the orange line) are achieved by LTE-A systems.

• Figures shown refer to the best of the use cases. This means that they won't be achievable in every use case, and each use case favours the achievability of different specs in each situation. (I.e. user experience data rate and latency will be higher in sparser environments while mobility will be higher in dense deployments). To see the expected figures for each case of use, please refer to chapter 4.

As it can be deduced from figure 5, this project consists of three milestones: two of which are set to 2020 plus a one long term one:

1. **IMT Advanced:** this first stage consists on fulfilling part of the requirements or filling them partially. There is no mandatory date to accomplish this phase. Nevertheless, it is expected to be finished between 2018 and 2019 (depending on each country).

2. **IMT-2020**: consists on achieving the requirements above mentioned completely. Its mandatory date to be finished is year 2020 for every country.

3. **IMT Beyond 2020:** this phase is still quite diffuse and not well defined yet. Its milestones are, as a vague approach, the ends of the flower's petals in figure 5.

This gives evidence that deployment of 5G networks must be done gradually in order to achieve these milestones gradually as well. Furthermore, it must not be forgotten that enhancement of LTE & LTE-A is essential as for the collaboration between these two generations. In figure 6, the planned roadmap by ITU can be seen. It shall be noticed that spectrum topics are also relevant, reason for the celebration of the WRCs.

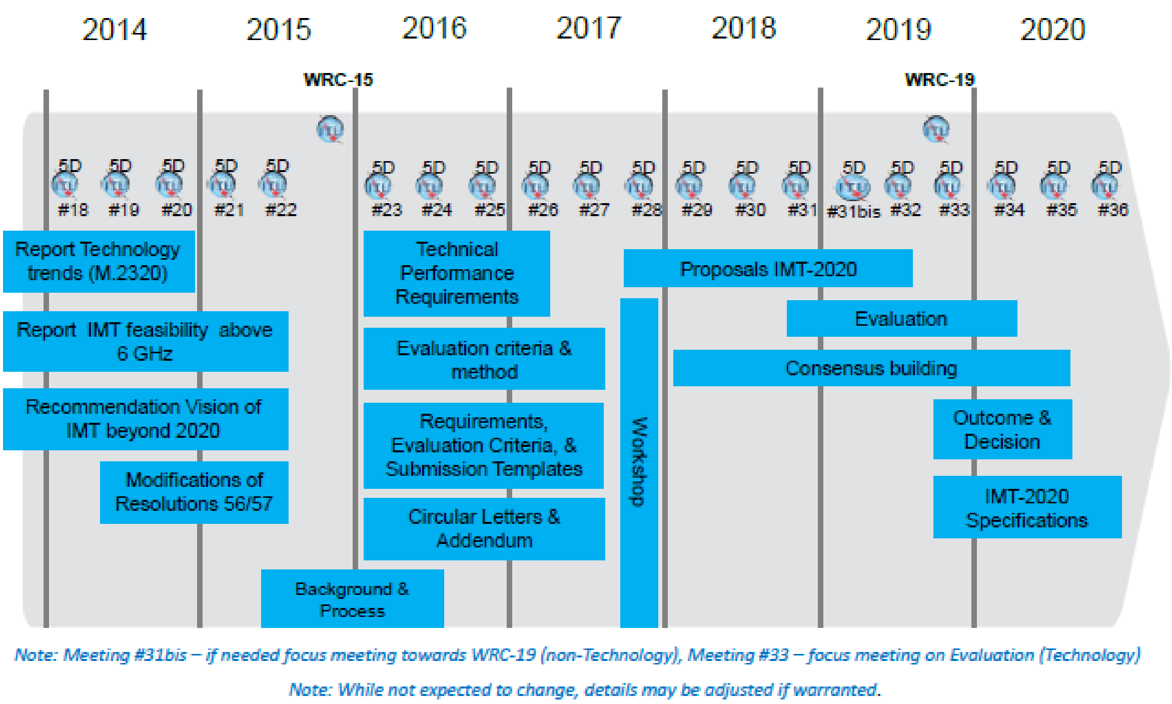


Figure 6. Detailed Timeline & Process for IMT-2020 in ITU-R. Source: ITU-R[5]

[5] In http://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/imt-2020/Documents/5D-Meeting-Schedule.pdf the date and description of each meeting (from #23 to #36) is detailed

2.1.2 3GPP: Releases 14 & 15

As in the previous section, the first important idea is to explain what 3GPP is. The 3rd Generation Partnership Project (3GPP) is collaboration project between groups of telecommunications associations, known as the Organizational Partners[6]. Although it was originally thought to develop a viable 3G system, it has been setting the next milestones in terms of technical standardisation due to the success with 3G.

The project covers all the aspects of cellular telecommunications, both SW and HW, hence, its aim is to define the entire system specifications. It is divided into three Technical Specification Groups (TSGs), which are:

- **Radio Access Networks (RAN)**: is responsible for the definition of the functions, requirements and interfaces of the UTRAN/E-UTRAN (Enhanced- UMTS Terrestrial Access Network) network in its two modes, FDD & TDD (Frequency & Time Division Duplex). More precisely: radio performance, physical layer, layer 2 and layer 3 RR (Radio Resources) specification in UTRAN/E-UTRAN; specification of the access network interfaces definition of the O&M requirements in UTRAN/E-UTRAN and conformance testing for UE and BS

- **Services & Systems Aspects (SA)**: is responsible for the overall architecture and service capabilities of systems based on 3GPP specifications and, as such, has a responsibility for cross TSG co-ordination.

- **Core Network & Terminals (CT)**: is responsible for specifying terminal interfaces (logical and physical), capabilities (such as execution environments) and the Core network part of 3GPP systems. More specifically: User Equipment - Core network layer 3 radio protocols (Call Control, Session Management, Mobility Management), signalling between the core network nodes, interconnection with external networks, Wireless LAN - UMTS interworking and descriptions of IP Multimedia Subsystem, mapping of QoS, etc.

Each group is as well divided into Working Groups, which address specific issues within those 3 general topics: security, QoS, frame architecture… Working groups usually have to collaborate with each other to make the standards.

3GPP follows this methodology: the TSGs meet every three months (four times per year). Together, they set common objectives, review achieved ones and revise the proposals and documents written by the different Working Groups. During those three months, each Working Group is committed to a specific task amongst each of the three areas of investigation. They write the different documents upon the selected issue and hand them to the TSGs in the next meeting.

[6] These Organizational Partners are ARIB, ATIS, CCSA, ETSI, TSDSI, TTA and TTC. See glossary.

With this working methodology as their basis, they structure all these documents provided by the different working and technical specification groups in releases. Therefore, a release is defined as a set of documents (in order of hundreds) that contains all the necessary technical standards in each layer and side of the system to define the actual telecommunication system[7]. As an example, Release 8 defined LTE and Release 10 defined LTE-A. The last release that has been completed is Release 13, which faces the topics of Security Enhancements, WiMAX and LTE/UMTS Interoperability, Dual-Cell HSDPA with MIMO, Dual-Cell HSUPA and LTE HeNB[8].

Hence, there are three stages in each release standardization process in which both groups and associated companies participate through their membership (this is called contribution driven). These stages are:

• **<u>Stage 1</u>**: specifications define the service requirements from the users' point of view.

• **<u>Stage 2</u>**: specifications define architecture to support the service requirements.

• **<u>Stage 3</u>**: specifications define an implementation of the architecture by specifying protocols in detail.

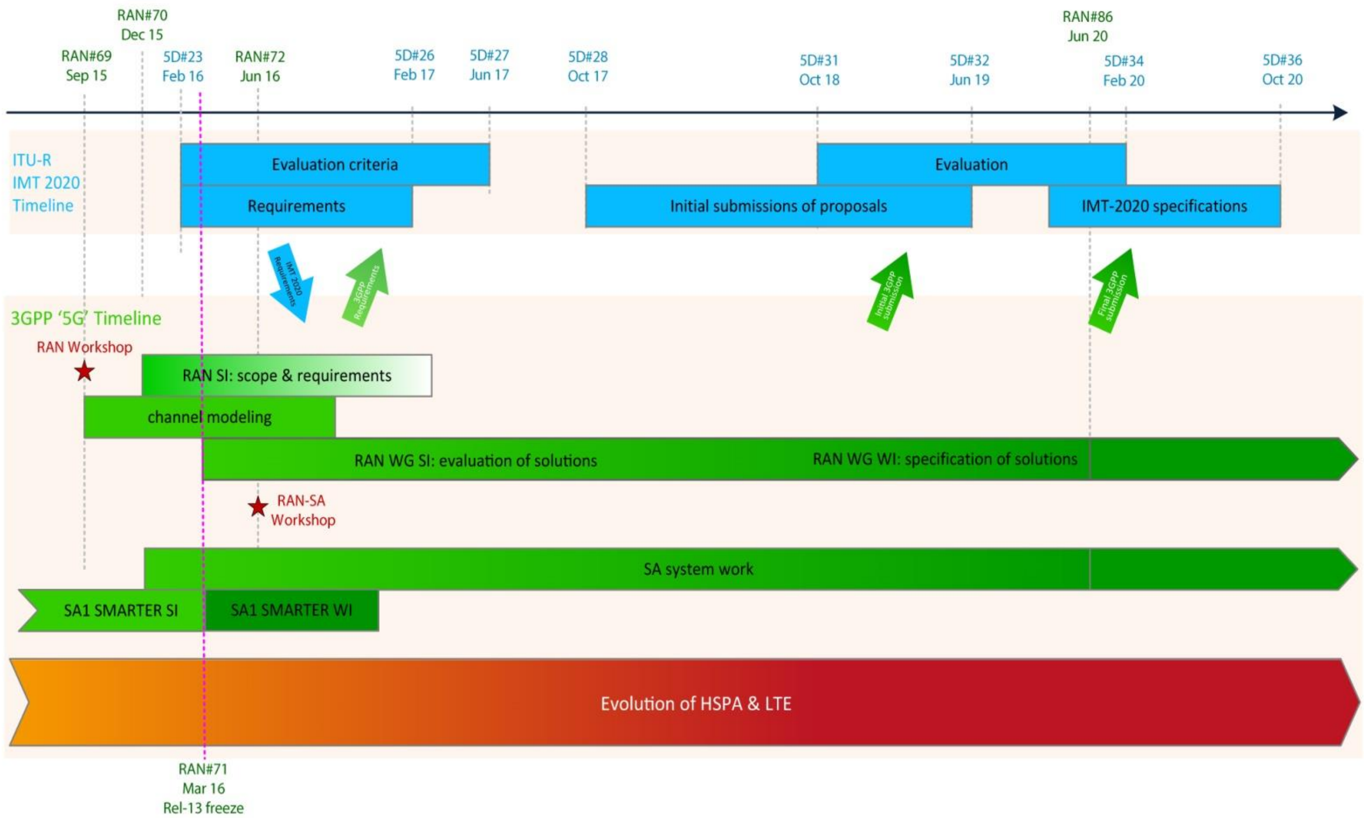


Figure 7. Collaborative Timeline of 3GPP and ITU for 5G roadmap. Reconstructed from 3GPP. Source Inside5G [7]

---

[7] Although 3GPP defines and elaborates these future standards, is ETSI's task to standardize them, as 3GPP is not a standardization body.

[8] All the information has been obtained from 3GPP website. Source [6]

As it has been mentioned throughout this and previous sections, the collaboration between all the standardization bodies, but most importantly between ITU and 3GPP is undeniably fundamental. In figure 7, the aggregate timelines of these organisations are shown together, detailing the relations between them: 3GPP is committed to submitting a candidate technology to IMT-2020 for evaluation. It will be evaluated upon ITU requirements. Parallel, evolution of LTE and HSPA will be going on. The evolution of these existing systems is an important part of the releases that will be delivered, as they constitute an important part of the goals as well.

<u>The importance of Releases 14 and 15</u>

Proceeding with the order above mentioned, releases 14 and 15 (which have been already planned and are being written by this date) will be defining the standards for 5G. Rel-14 and Rel-15 represent the first phase of work in 3GPP towards 5G standards with Rel-14 focused on the study items towards 5G and Rel-15 on the first phase of normative specifications for 5G. More specifically, the details of these realises can be seen in table 4:

Table 4. Features and objectives of 5G releases.

	Release 14	Release 15
Start date	17/9/2014	1/6/2016
End date[9]	9/6/2016	14/9/2018
Main objective	Study items	Normative specs.
Issues addressed	Energy Efficiency	Delivering the first set of 5G standards
	Location Services (LCS),	Enhancement on LTE-A Pro and use for 5G requirements
	Mission Critical Data over LTE,	
	Video over LTE, Flexible Mobile Service Steering (FMSS),	Specification of New Radio (NR) architecture and Evolved Packet Core
	(MBSP)[10] enhancement for TV service	
	massive Internet of Things	Forward compatibility between LTE-A and NR
	Cell Broadcast Service (CBS)	

[9] This date is expected, though it can change, especially for release 15.
[10] Multimedia Broadcast Supplement for Public Warning System

2.2 The spectrum problem

In this section, the other main conditioning factor for mobile communications will be studied: the radio waves electromagnetic spectrum. It is fundamental for mobile communications, as it is the physical medium information travels within. The definition of the spectrum is "the set of wavelengths belonging to all the electromagnetic radiations". And on the other hand, a radio wave is defined as "an electromagnetic wave having a wavelength between 1 millimetre and 30,000 meters, or a frequency between 10 kilohertz and 300,000 megahertz".

2.2.1 Problems of a limited and shared resource

As it was explained in chapter 1 there are two main constraints for the use of the spectrum, its natural limitation and its shared use among all kind of wireless and broadcasting services. Thus, problems will appear when planning to use particular bands of the spectrum.

From the definition given in the heading of this section, it can be easily assumed that the spectrum is limited between **10 KHz-3000 GHz**. This limitation sets a finite number of band allocations and therefore a finite number of services that can be provided[11]. Furthermore, each band has its limitations as well, as it will be seen in next section 2.2.2.

On the other hand, it is really necessary to understand the importance of the spectrum sharing. This is one of the main constraints in the mobile communications sector. Many other services use the spectrum: TV, other radio networks (Wi-Fi, ISDN), satellites… Furthermore, there are other services such as military applications, medical services and radio, etc. that enjoy a privileged position in terms of band allocations, having the best ones[12].

The main drawback due to these constraints seems clear: the binary rate (the speed that users experience) is directly proportional to the available bandwidth. Although it depends on many factors such as type of modulation, carrier spacing, etc. there is no doubt that they hold a narrow relation. Therefore, if the bandwidth is limited, so will the binary rate be.

At this point, engineers must come up with different strategies and techniques to minimize the impact of the lack of bandwidth or even compensate it. This topic will be addressed in chapter 3, yet it can be told in advance that previous generations have also needed of these techniques to achieve with worst technology the actual results (that now relay insufficient).

[11] The number of services varies upon many factors such as energy and bandwidth efficiency.
[12] See section 2.2.2 to understand the meaning of the best bands.

2.2.2 Bands and their characteristics: regulatory models

As it has been introduced in previous sections, each band holds different properties which can be considered as pros or cons depending on the application. In this section, all the bands that are significantly important to mobile communications will be classified and analysed.

Hence, the first point is to address the band classification and which of them are susceptible of being used for mobile communications. In next figure (Fig. 8) (see next page), the classification of bands upon frequency and their uses is displayed. First of all, it must be clarified that this information belongs to the Japanese government and therefore it refers to the Japanese situation. Nevertheless, frequency allocations all around the world are similar in an overall approximation. There are three main conclusions that shall be taken out from this figure:

1. The higher the frequency, the greater the transmission capacity (in terms of information rate) and the smaller the propagation capacity (in terms of distance). Therefore there is a commitment between distance and information rate which is the main constraint when designing a system. Although it is not displayed in figure 8, there is another constraint, which relays in the complexity of technology. The higher the frequency is, the more difficult it is to build the technology capable of operating in those bands. In the same sense, the battery life becomes shorter when working on higher frequencies as modulation techniques become more complex.

2. There are bands that are uninteresting or inaccessible for use. Specifically, under 300 MHz (VHF and below), other services occupy the bands, where governments, military and rescue services have absolute priority. In this sense, these services have the best bands named in previous section, the ones with the best commitment between propagation and rate information features. The upper boundary in this sense is limited by the technology capacity, as it was explained in previous point.

3. Geographical harmonization holds great importance as well as efficiency in terms of frequency. It is obvious that two waves won't interfere as long as they do not share geographical space and frequency. Hence, geographical planning of services is as important as band allocations. This planning is even more important to harmonize local services with regional, inter-regional, national, global… As an example, GPS satellites are planned to have similar frequencies globally[13].

[13] GPS carrier frequencies are 1575 for civilian use MHz and 1227 MHz for military use.

There are two specific cases of allocation within the different bands that should be considered. These two opportunities often facilitate spectrum harmonization, especially for local and personal use of itself.

a. White spaces

Related to the third conclusion, there is a specific case that is often given: white spaces. These are frequencies allocated to a broadcasting service that are not used in a specific local region. The different regulatory bodies assign different frequencies for specific uses, and in most cases, license the rights to broadcast over these frequencies. This frequency allocation process creates a band plan, which for technical reasons assigns white space between used radio bands or channels to avoid interference. In this case, while the frequencies are unused, they have been specifically assigned for a purpose, such as a guard band. Most commonly however, these white spaces exist naturally between used channels, since assigning nearby transmissions to immediately adjacent channels will cause destructive interference to both. In addition to white space assigned for technical reasons, there is also unused radio spectrum which has either never been used, or is becoming free as a result of technical changes (such is the case of digital television). Hence, these spaces can be used by other services.

b. Unlicensed spectrum

These bands are the ones that have not been officially allocated by the actual regulatory bodies. Hence, anyone can use them but there is no guarantee of avoiding interferences. The best known example is the use of Wi-Fi in the 5 GHz band, which is currently in use, supported by a backhaul connection in the licensed band of 2.4 GHz. The meaning of the use of the 5 GHz is to have a higher data rate in a small area, but the lack of QoS makes necessary the support of the backhaul connexion.

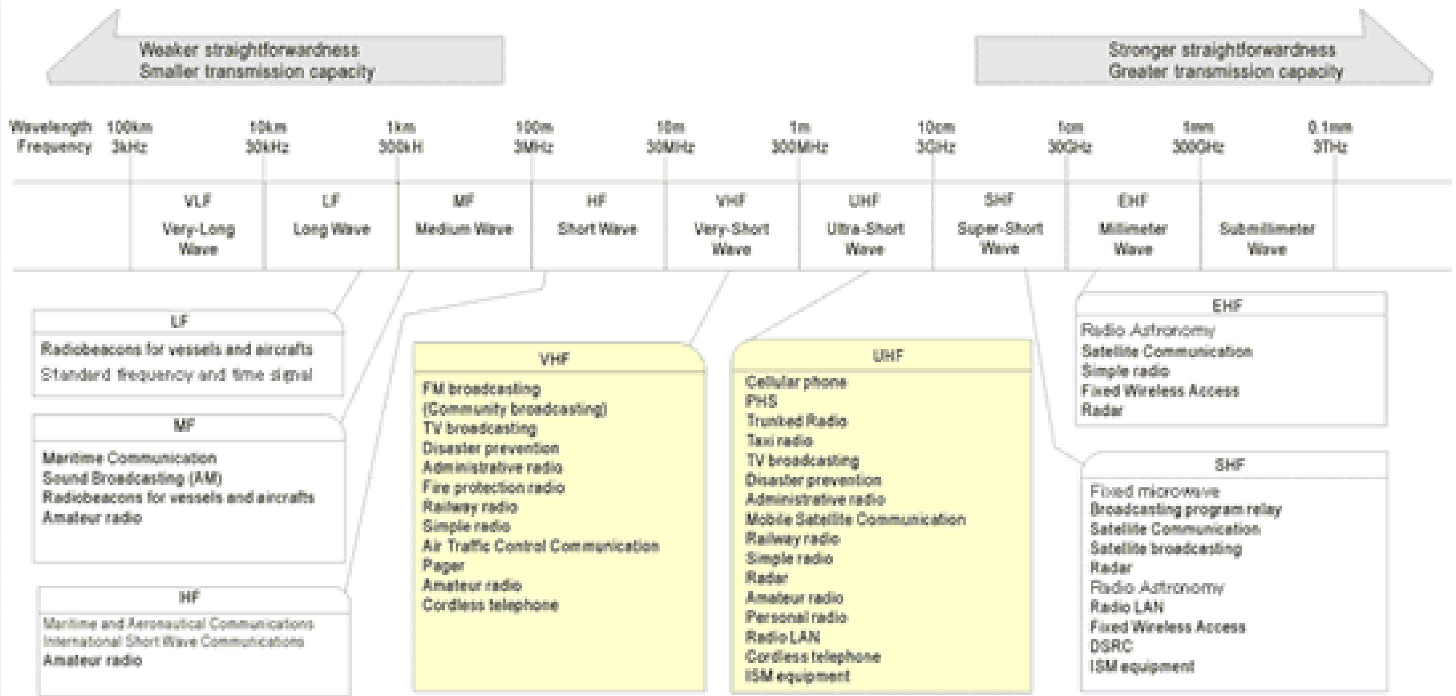


Figure 8. Band classification. Source: MIC [10]

Next step is to limit the bands for mobile communications. As it can be seen in table 5, UHF (300-3000 MHz) is the band that is mainly used for cellular phone mobile communications[14]. The upcoming investigation line is the use of the SHF and EHF bands for this type of communications. Recalling what was concluded in the first point of this section, these two bands will have short scope but high information rate capacity, which will result into higher speeds in small areas (cells).

At this point, with the bands already limited, there is another issue that must be given an answer. It is quite clear that regulatory bodies allocate each band to the different services based on their criteria, which represents the common interest of a nation or community. But, how are the frequencies within a band distributed among different operators with the same or different purpose? In other words: in this case there are different services in UHF (see figure 8) and therefore different providers (Vodafone, Orange and other providers) that have interest in particular frequencies, as well as other providers of the different services above displayed; hence a fair procedure is needed to address this problems.

Spectrum management models

Far from a unique solution, there are three answers to that question, which are the three existing models for the management of the spectrum. These are:

a. Command and control

This is the traditional model, in which a national regulatory body allocates the different bands, specifying the requirements they wish to strike, such as the type of technology to deploy, the levels of QoS or coverage, license period, etc. This is a rigid model, as long as the regulatory body determines whether a new technology can access or not the market by allocating bands to it or not, yet it provides a very good harmonization of the spectrum, avoiding interferences in most cases.

b. Property rights model

In this model, the correspondent regulatory body initially specifies the licenses' rights of use as well as the first allocation, yet it is the market forces the ones to determine the band re-allocations (and new allocations) to each service as well as the types of technology used for each service. It is opposite to the command and control model as for the lack of responsibility of the regulatory body once made the first allocation, though it provides the spectrum with more flexibility. However, a big drawback of this model is the inexistence of a procedure to harmonize regions or countries, creating a big constraint to global companies which cannot coordinate their economies of scale due to the constant change in the different spectrums.

[14] Please remember that since GSM all (commercial) mobile systems are cellular based.

c. Commons theory

This model consists on letting every single user use the spectrum without any license or previous allocation. However, there are some constraints that can be set such as the maximum power allowed. With the application of this model, it is not possible to assure neither QoS levels nor interference avoidance. Nowadays, this model is applied in Wi-Fi and Bluetooth networks, amongst other local uses, such as amateur drones. This model cannot be applied in other cases of use due to the complexity and amount of band allocations in a bigger environment.

Traditionally, the command and control model has been used throughout history, in which licenses were handed out with a set duration and technology specified by the regulatory body. This model has resulted adequate, due to the slow variations in technology and the consequent investments' payback periods, until the 00's. Nowadays, it has been proved to be inefficient in a globalised world with such a high innovation speed because it slows the different technological breakthroughs.

Therefore, the model has migrated to property rights model, combined with the commons theory applied to very specific cases (Wi-Fi, Bluetooth, remote controls…). With the migration to this model, the drawbacks above mentioned raised, and solutions to them were to be investigated. The main problem was harmonization in different regions. There are two important points to understand how this problem has been issued:

- **Bandplans:** these are roadmaps set by international regulatory bodies (ITU-R, WARC) that the overwhelming majority of the countries follow. These will be fully explained in next section.

- **Technology Neutrality Principle:** Technological neutrality is a widely accepted regulation principle in the EU framework for electronic communication services. It consists on licenses adjudication to give a service without specifying any mandatory technology to do so. This principle was set to facilitate the innovation and the gradual incorporation of new technology. Nevertheless, there are some articles that criticise the lack of discussion upon this principle which could have led to an empty and inconsistent principle[15].

[15] See "Technological Neutrality in the EC Regulatory Framework for Electronic Communications: A Good Principle Widely Misunderstood", Ulrich Kamecke and Torsten Körber

2.2.3 World Radio Conferences: digital dividends & current spectrum display

In this section, both the current situation and the upcoming possibilities of the spectrum will be analysed. The first premise is to remember that the regulatory body which carries the weight of the band allocations planning is ITU-R (Radio communications Group). As it was explained, ITU makes recommendations according to its working procedure. In Europe's case, there is another international body called CEPT (European Conference of Postal and Telecommunications Administrations), formed by 49 countries, whose responsibility is to elaborate the recommended band allocation plan as well as harmonization procedures between its members in coordination with ITU. Nevertheless, it is the national regulatory body (belonging to the actual government) the one to allocate the bands. For international harmonization reasons, most of the countries follow the scheduled planning proposed by ITU. In Spain's case, a body called CNAF (Cuadro Nacional de Atribución de Frecuencias) is in charge of that commitment.

Before deepening in this issue, there is an important detail to be mentioned in order to understand this section. ITU divides de countries around the world in three main regions, which are shown in figure 9. The importance of these regions lay on the fact that ITU recommendations may not only be global, but also be designed for a specific region in some cases in which some bands are not harmonized among regions. For each combination of region and frequency band there are some particular allowed services, which don't have to be (although they usually are) the same for its neighbour countries.

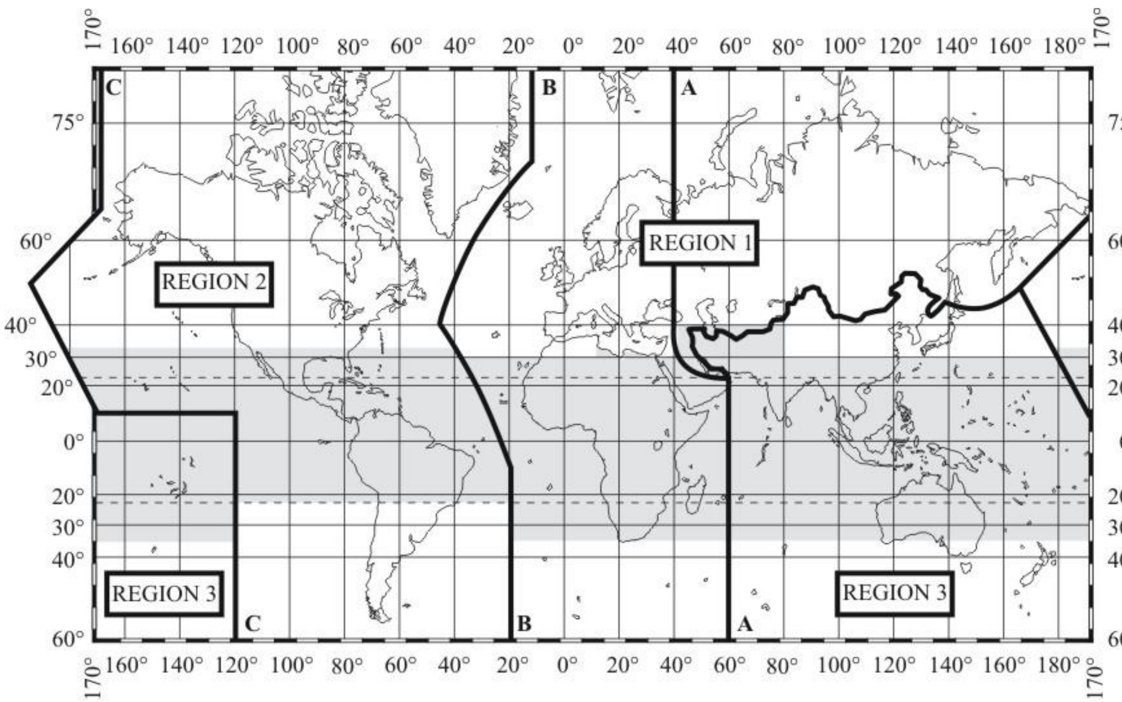


Figure 9.World map of ITU regions. Source [12]

Besides the recommendations mentioned, ITU carries every three or four years a special conference called World Radio (Administrative) Conferences: W(A)RC, whose aim is to review, and, if necessary, revise the Radio Regulations, the international treaty governing the use of the radio-frequency spectrum and the geostationary-satellite and non-geostationary-satellite orbits. Revisions are made on the basis of an agenda determined by the ITU Council[16], which takes into account recommendations made by previous world radiocommunication conferences [11]. Needless to say, these conferences are held to address not only mobile communications problems but also every issue related to the spectrum and the allocations in its bands. Before the first of these conferences referred to 5G was celebrated, the current situation and objectives, both results from previous conferences (WRC-07 and WRC-12) were the following ones:

1. **Assigned Spectrum for IMT**

Table 5. Assigned Spectrum for IMT

Application area	Band(s) in MHz
Global scope	450-470 & 2300-2400
Region 1	790-862 (Europe)
Region 2	698-806 (America)
Region 3	790-806 (Pacific Asia)
Others	698-790 in 9 countries R3 & 3400-3600

2. **First digital dividend**: a digital dividend is a re-assignation of a specific band that was dedicated to another service due to social and technology efficiency issues. The first digital dividend liberalised the 800 MHz band that was dedicated to analogical TV, concurring with the switchover from analogical to digital TV, the last of which is quite more efficient and therefore does not need such a wide band to broadcast.

3. **Research on new spectrum for mobile communications**: back in 2012, the need for more spectrum bands was seen as a requirement in the short-mid run. Besides a plausible second digital dividend, other solutions were needed. As lower bands were already in use, the scope was initially set on the first free upper bands.

4. **Research on spectrum for PPDR (Public Protection and Disaster Relief)**: this was and actually is an important issue to be addressed. One of the most important applications in communications is critical and health services in extreme situations such as natural disasters, wars, etc. Hence, spectrum must be dedicated to this application mandatorily.

---

[16] ITU Council is the governing body in the interval between Plenipotentiary Conferences. Its role is to consider broad telecommunication policy issues to ensure that the Union's activities, policies and strategies fully respond to today's dynamic, rapidly changing telecommunications environment.

There are two main WRC's that are important to 5G, whose starting point was the summation of the results above mentioned. These two conferences are:

a. WRC-15

This conference took place in Geneva (Switzerland) in November 2015. Many topics such as Radiolocation and maritime services were approached, but as for mobile services (which include IMT and mobile broadband), the outcome resulted as follows[17]:

• Increase in +60% in IMT bands. Total IMT spectrum: 1886 MHz

• Increase in +39% in globally harmonized spectrum (+318 MHz in more than 80% of countries). Total IMT harmonized spectrum: 1228 MHz

• Second digital dividend: this second dividend consists on the liberalisation of the 700 MHz band, what would result on a huge band in the 700-900 MHz band. Consensus was not reached due to the discrepancies between the different regions and its members, most of which have different uses for that particular band. In most Arabic and Asian countries (regions 1 and 3), 700 MHz band was not used at all. Moreover, fixed telecommunications infrastructure is still being developed, and would need that band for military and government services. For all these reasons and inconveniences alleged by many different countries, this second digital dividend has been postposed with 2020 as its deadline.

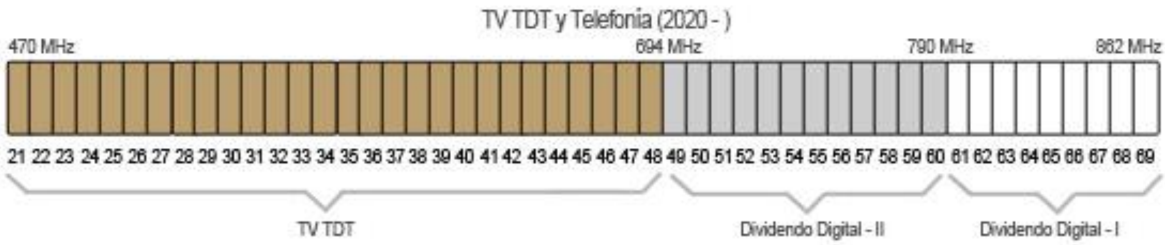


Figure 10. Spectrum forecast displayed with both digital dividends applied in Spain.

As it shall be seen in figure 10 and table 5, ITU recommendations differ from current Spanish spectrum organisation in the 400 MHz band. In figure 13, it can be seen the amount of spectrum that is not harmonized yet despite ITU recommendations. All the problems of this shared media (explained in section 2.2.1) come to the surface, even more if a global harmonisation is needed.

In next three figures, all of which come from the same source [13], the outcome of WRC-15 is best summarised in terms of spectrum assigned and harmonized for IMT.

[17] Reader must remember that these are recommendations by ITU, not mandatory rules.

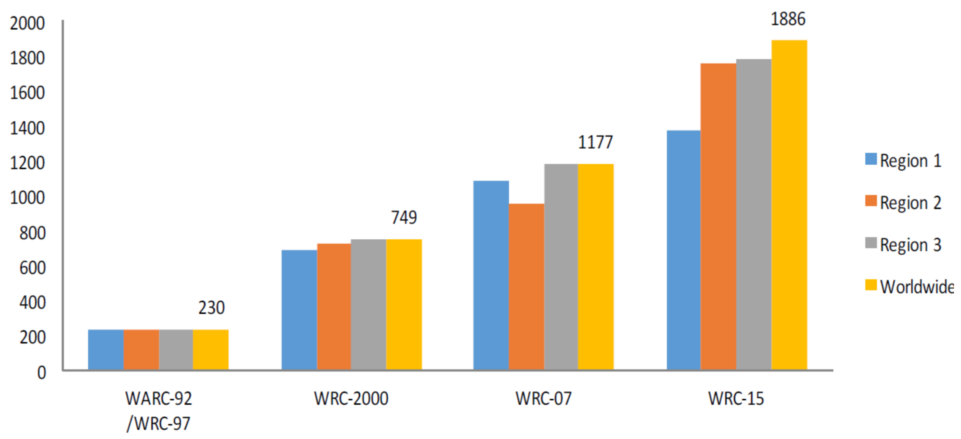


Figure 11. Total amount of spectrum identified for IMT (MHz)

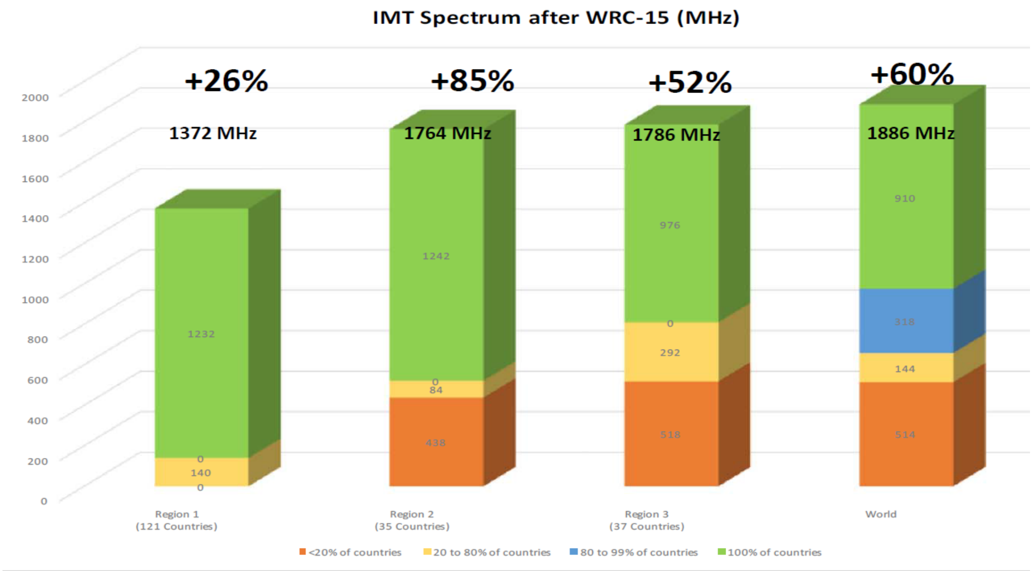


Figure 12. IMT Spectrum after WRC-15 (MHz)

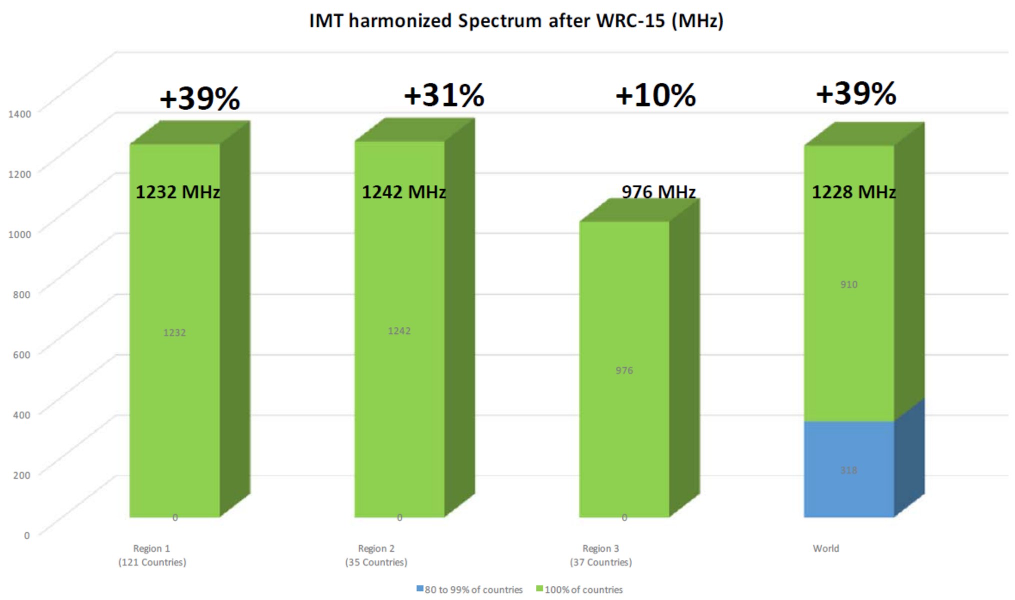


Figure 13. IMT harmonized Spectrum after WRC-15 (MHz)

These figures show what was concluded upper in this section: over 800 MHz, the spectrum is globally harmonised, while under 800 MHz, some countries (especially in Region 3) have other uses for that bands. Despite those differences, it is quite positive to analyse that both IMT spectrum and IMT harmonised spectrum are growing. However, the question that remains and will be solved in next chapters is the same: is this growth enough to hold the parallel demand growth?

It must not be forgotten that harmonisation's main purpose is to facilitate the interoperability and coordination between countries to develop the new services maximizing the efficiency and minimizing efforts. As an example, a service provider like Vodafone could operate in different countries with different frequencies in each of them, but undoubtedly it would be much easier if they only had to develop equipment and technology to operate in a particular band for every contry.

b. WRC-19

This conference has not been held yet. It will be held between October and November 2019. As for today, the first preparatory meeting and studies are being developed to give a direction on what this conference will have to treat about. This particular conference will be important due to its proximate date to 2020, the most important year for IMT roadmap. Therefore, the topics that are already known to be discussed in WRC-19 are the tendering of the bands over 6 GHz, the possible reassignments of bands below 6GHz and a plausible second digital dividend, as long as the different national regulators accept it.[18]

Another important topic for this conference is the use of 5 GHz Wi-Fi as a licensed service. Although it is not directly related to IMT, it does influence as long as these networks to which phones connect (and therefore they carry traffic into these networks instead of carrying it to the corresponding telecommunications service provider) are increasingly carrying more and more traffic[19]. Furthermore, the 5GHz spectrum could be an interesting option in the new frequencies research for IMT. Either it is assigned or not to IMT, it must be licensed to regulate interferences and QoS in order to pursuit the efficiency mantra.

Furthermore, the new transport systems will be brought to discussion as well. This topic is referred to autonomous and self-driven (driverless) vehicles, intelligent transportations, drones, etc. Although is not strictly spectrum-related, it is also important for 5G because it constitutes one of the possible cases of use for 5G (please refer to chapter 3).

[18] In terms of digital dividend, it would be a first digital dividend for the mentioned countries of Region 3, as they never had to liberalise the 800 band.
[19] Over the Top (OTT) applications are a good example of the traffic that is being driven to Wi-Fi networks.

2.2.4 Harmonization in the EU: Lamy's Report & Digital Single Market

As it was mentioned in previous section, there are two agents involved in EU's situation. Firstly, CEPT represents all European countries as well as other non-European neighbours with whom we share borders. In this sense, CEPT actively participates in WRCs representing the interests of these countries. Secondly, the European Commission holds the European legislative power. Hence, it disposes the mandatory assignments for each of the EU members, which are discussed and coordinated with each member's national regulator (in this case CNAF). The two most remarkable aspects of EU's harmonization milestones are:

a. Lamy's report

Pascal Lamy is the head of a High Level Group (HLG) of the EC. This group was created to study the possibilities of the UHF band in Europe, which included the terrestrial broadcasting TV service versus the mobile services not only for mobile communication but also for broadcasting. Since there wasn't a consensus within the HLG, Lamy delivered and signed a report on behalf on himself but not the rest of the HLG.

In this report, he has proposed a "2020-2025-2030" formula with the aim of enabling Europe to fulfil 'Digital Agenda for Europe' broadband targets in three steps, while giving broadcasting a clear path to invest and develop further [14]:

- The 700 MHz band (694-790 MHz) is currently used by terrestrial broadcasting networks and wireless microphones) should be dedicated to wireless broadband across Europe by **2020** (+/- two years);
- Regulatory security and stability for terrestrial broadcasters in the remaining UHF spectrum below 700 MHz to be safeguarded until **2030. This involves national, EU and international measures.**
- A **review by 2025** to assess technology and market developments. In order to take into account the evolving change in consumer demand as well as new technologies, such as converged networks or large-scale roll out of optic fibre**, a stock-taking exercise of UHF spectrum use should be performed by 2025.** It would give Europe the opportunity to re-assess where we stand and avoid any freeze of regulation compared to the rapid advance in technology and consumer behaviour.

This gives evidence that EC's positioning towards the second digital dividend will be favourable, and therefore TV broadcasting will have to reassign its channels and increase efficiency in order to maintain the channels that are currently available.

b. Digital Single Market

The DSM is one of the main objectives in the European Digital Agenda. Although DSM embraces many application fields[20], there are various key aspects that are defined as some of the many missions that it carries out. These particular aspects are spectrum management responsibilities; therefore they affect the future of 5G. These are:

• The definition of common rules for every member. They must be mandatory when applying their own spectrum policies so that they all guarantee a common harmonization basis.

• The power of the EC to apply extraordinary measures within members' spectrum policies in order to ensure that aforementioned harmonization (i.e. availability, assignment times, licenses continuity, etc.).

• A particular consultation procedure with which the EC can discuss its member's national policies in this field.

• Simplification of deployment and supply conditions for service providers in order to facilitate the incorporation of new competitors and therefore potentiate free market and competition.

The last point to address in this chapter is to see what the spectrum current situation in Spain is, as for the main title of this document. This particular display is the consequence and combination of the policies applied by each of the regulatory bodies that have been analysed in this chapter. So, lastly, in next table, the specific bands that each service provider has in Spain can be seen. It should not be forgotten that a provider is given more or less bands based on its subscribers. This figure dates of 2015, date in which the last modification of the current applied attribution (CNAF 2013) was published.

Xataka movil	800 MHz (Banda 20) 4G	900 MHz (Banda 8) 2G y 3G	1.800 MHz (Banda 3) 2G y 4G	2.100 MHz (Banda 1) 3G	2.600 MHz (Banda 7) 4G
movistar	10 MHz FDD	14.8 MHz FDD	19.8 MHz FDD	15 MHz FDD 5 MHz TDD	20 MHz FDD 10 MHz TDD Madrid
vodafone	10 MHz FDD	10 MHz FDD	19.8 MHz FDD	15 MHz FDD 5 MHz TDD	20 MHz FDD 20 MHz TDD
orange	10 MHz FDD	10 MHz FDD	19.8 MHz FDD	15 MHz FDD 5 MHz TDD	20 MHz FDD 10 MHz TDD 10 MHz TDD regional
yoigo	-	-	15 MHz FDD	15 MHz FDD 5 MHz TDD	-

* MHz pareados en FDD.
**Regional en Orange disponible en toda España excepto Castilla La Mancha, País Vasco, Asturias, Galicia, Madrid y Melilla.

Figure 14. Bands of each service provider in Spain. Source: Xataka-movil [15]

[20] To see all the application fields: https://ec.europa.eu/commission/priorities/digital-single-market_en

3. <u>Deployment of a 5G network</u>

So far, the analysis has been focused on what the fifth generation overall requirements are and how the spectrum looks nowadays, as well as all its possibilities (i.e. the application of the second digital dividend). In the present chapter, the technical features, procedures and technologies necessary to deploy a network that shall meet the 5G requirements are discussed. Therefore, this chapter will go further the mentioned goals for 5G networks: it will deepen into the technical issues to address, may they be enhancement of current technologies or development of new ones.

First of all, the corresponding premises and current technologies will be explained. This will allow us to better understand the subsequent innovations in each layer of the network as well as its motivation in each of the cases. Next, every layer of the net will be analysed alongside its plausible innovations.

Once again, it must be recalled that by the first trimester of 2017 there is not much known from a technological point of view, assuming that 3GPP hasn't finished Release 15, and Release 16 is still being planned and organised. Therefore, it will henceforth be necessary to make assumptions (that will be sufficiently justified) when no other options can be contemplated (especially for the analysis driven in chapters 4 and 5).

3.1 Premises

Firstly, it is totally indispensable to review a set of premises without which the rest of the chapter would not be understood. These premises refer to basic technological concepts of mobile networks. They will be useful to scope out the technological state of art that was vaguely mentioned in the previous chapter.

This section will be mainly focused on analysing the technological premises of 5G networks, which obviously belong to 4G networks. Their importance here is obvious because (as mentioned in previous chapters) the enhancement of these networks constitute one of the two pillars of 5G. 4G networks have been mainly developed upon two different systems: LTE[21] (and its next enhanced version LTE-A) and Wi-Max. This second system won't apparently evolve in this direction. Hence, the system to be analysed is LTE.

3.1.1 Segments of a network's architecture

Step zero is to assume that the architecture is common to all cellular-type mobile communications networks. There are three high level segments in all these networks (from GSM to LTE going through UMTS, LTE…). These three segments/planes are:

[21] LTE is sometimes not considered 4G because it is slightly under the 4G requirements. LTE-A is.

• **<u>Users' Segment</u>:** this side of the architecture embraces all the devices, named user equipment (UE), which can be connected to the network. To connect to a mobile communications network, these devices must be able to incorporate technologies capable of accessing the network via radio waves (i.e. antennas). Typically, mobile phones have been the only type of devices that can connect to networks. However, other kinds of devices will be able to connect via radio to the network, especially with the upcoming development of IoT.

• **<u>Radio Access Network Segment</u>:** this side is the intermediary between the users' segment and the core network. Devices only connect to this segment via radio, and the modules of this segment connect to the core network via radio or/and cable. This segment consists on the base stations to which devices connect. They are in charge of giving coverage to the subscribers. The operations carried out by the RAN are the most basic and simplest ones: admission and connection control, signalling, basic configuration (i.e. resource allocation in calls), etc.

• **<u>Core Network Segment</u>:** in this segment relays the entire network's intelligence. All the operations and decisions are made by the modules of the core network and driven to the radio access network, where they are transmitted to the users. The core network's gateway can connect to other networks with which traffic is exchanged.

Each of the three segments and modules within them are connected by interfaces. Each interface is designed to support a specific type of data: traffic, signalling, or both. Traffic in each interface must comply with a standard and structure that is determined by the communication protocols of the system[22] to ensure network's interoperability. The high level architecture of a network is displayed in next figure. As mentioned, there is no physical connection between the Users Equipment (UE) and the Core Network (although there is logical connection).

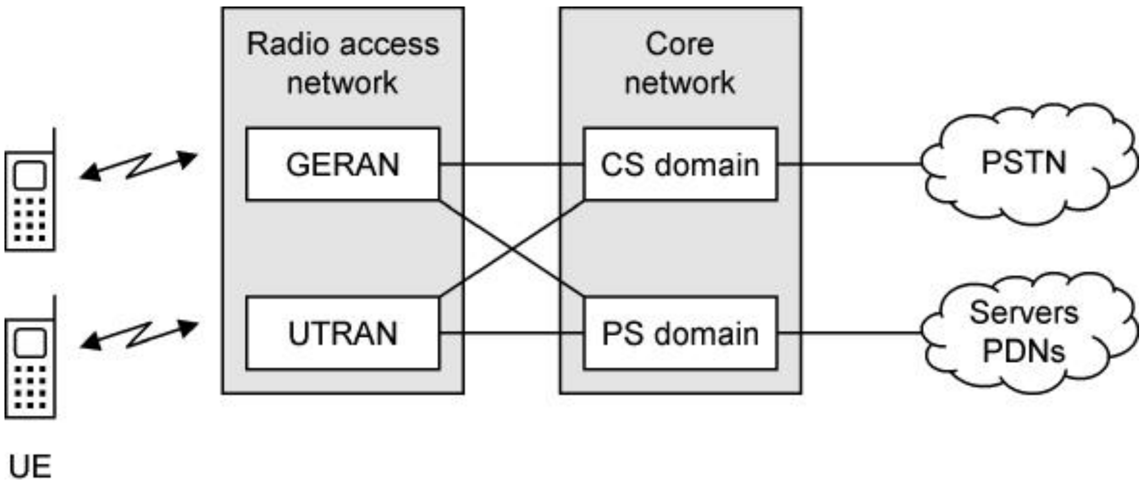


Figure 15. High-level architecture of the network. Reconstructed from [16]

[22] Some communication protocols examples are IP, TCP, Voice over LTE (VOLTE), etc.

3.1.2 LTE architecture and components

In the following figure, the architecture of 4G (including all the entities and interfaces inside each part of the networks) are displayed. Afterwards, the tasks of each entity will be explained.

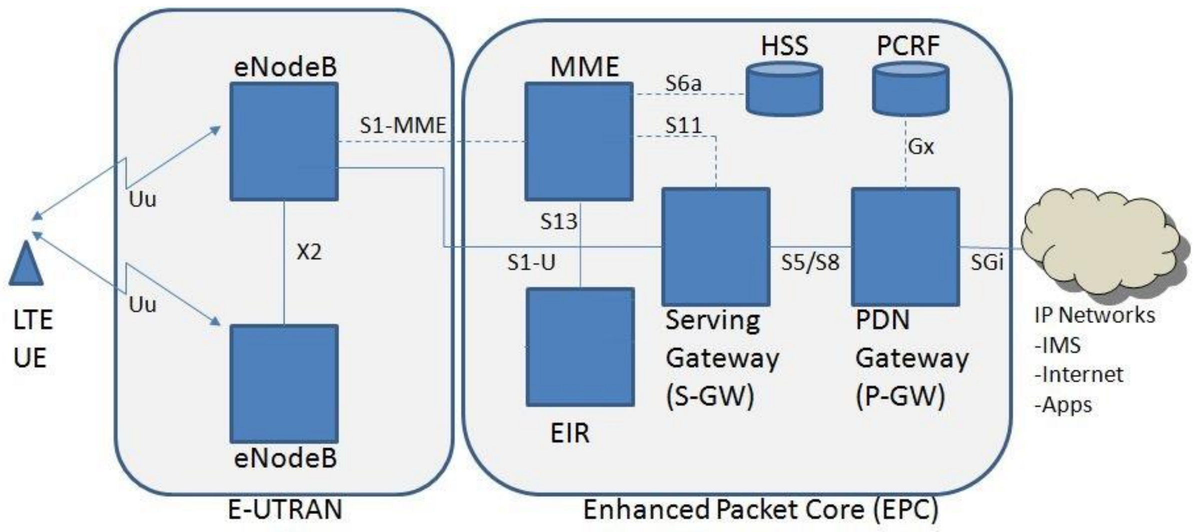


Figure 16. LTE architecture. Reconstructed from [17]

As it may be seen from figure 16, the three segments explained appear in it. In this figure, interfaces are painted as continuous or discontinuous lines between modules. Continuous lines represent the data information interfaces; while discontinuous ones represent control information interfaces. It must be pointed out that UEs can simultaneously connect to other networks, may they be 3GPP (GSM, UMTS) or non 3GPP (Wi-Fi, WiMAX). Assuming there is no module to be explained in the Users' Segment, the two segments studied will be the following ones:

a. **E-UTRAN:** its name means Evolved UMTS Terrestrial Radio Access Network, and it is called so because it represents the evolution of the network that was used in the third generation (UMTS). This is LTE's access network via radio. The evolution of UTRAN was meant to make it as simple and lean as possible. Hence, the modules were simplified to just one: the eNodeB.

   - eNodeB: the evolved Node B gathers the tasks that were carried by both the obsolete nodeBs and the RNCs (Radio Network Controllers) in the previous generation[23]. Devices can be (and usually are) connected to more than one eNodeB in order to get a better service, and eNodeBs are connected between them by the X2 interface to exchange all the necessary mobility information. They are in charge of managing the radio resources, choosing the most

[23] Almost every RNC's function is transferred to the eNodeB. However, some functionalities are transferred to the core

adequate MME (refer to next paragraph) and routing traffic from/to the corresponding gateway.

b. **<u>EPC</u>:** its name means Evolved Packet Core because it is as well the evolution of the Core Network of UMTS, and it constitutes the Core Network Plane in LTE. The main difference with its ancestor is that it only courses packet traffic: there is not circuit switching. Hence, voice over LTE (VoLTE) is not always supported and therefore voice calls are coursed through UMTS & GSM (please remember what was explained at the beginning of this section). The modules that conform the EPC are the ones that follow:

- Mobility & Management Equipment (MME): as it's connected in both segments by signalling interfaces, it may sound logical that it is in charge of signalling tasks (belonging to the control plane): controlling security while accessing the network, selecting the best gateway (GW) and coursing the signalling traffic of the UEs.

- Equipment Identity Register (EIR): it is connected only to the MME because this module helps the MME in the development of its tasks. It may be defined as a devices' database. It stores the information of each device, what may be useful to locate stolen devices to block them. As of the nature of a database, it is connected by a data traffic interface.

- Serving Gateway (S-GW): it receives the data traffic from the RAN (user plane) and the signalling traffic from de MME to control the up bound traffic alongside its routing and QoS. The information it remits to the other gateway is data traffic, but packets sent there also contain signalling information within.

- Home Subscriber Service (HSS): this is another sort of database that the MME uses to develop its tasks. In this case, interface S6a that links them is a control-type one because the type of data stored in the HSS is the users' mobility information, which is obtained by signalling protocols. The HSS is especially important when users transfer from one cell to another and handover must be eliminated or reduced as much as possible.

- Policy Charging Rules & Fares (PCRF): as it may be deduced from its name, PCRF controls the billing and authorization. In other words, it is in charge of calculating how much does a user pay for the traffic coursed. PRCF is especially important in roaming situations[24], in which it also controls and reports information on QoS.

- Packet Data Gateway (P-GW): this gateway is the only connection between LTE networks and the packet networks and IMS network. Therefore it carries out

[24] With the end of the roaming in the EU set to 2017, PRCFs will have to be reconfigured on commandment of the EC.

typical gateways' tasks: assigning the UE's IP, controlling QoS, filtering and discarding (when necessary) traffic…

In next figure, a wider perspective on LTE networks can be taken into insight. It can be seen that LTE constitutes a part of a bigger framework that makes possible mobile communications. As it was mentioned, 3G networks are still important due to its use in voice calls, as well as a backhaul connection. Furthermore, femtocells are also important for some indoor small scenarios. In next section, it will be widely discussed the importance of femtocells.

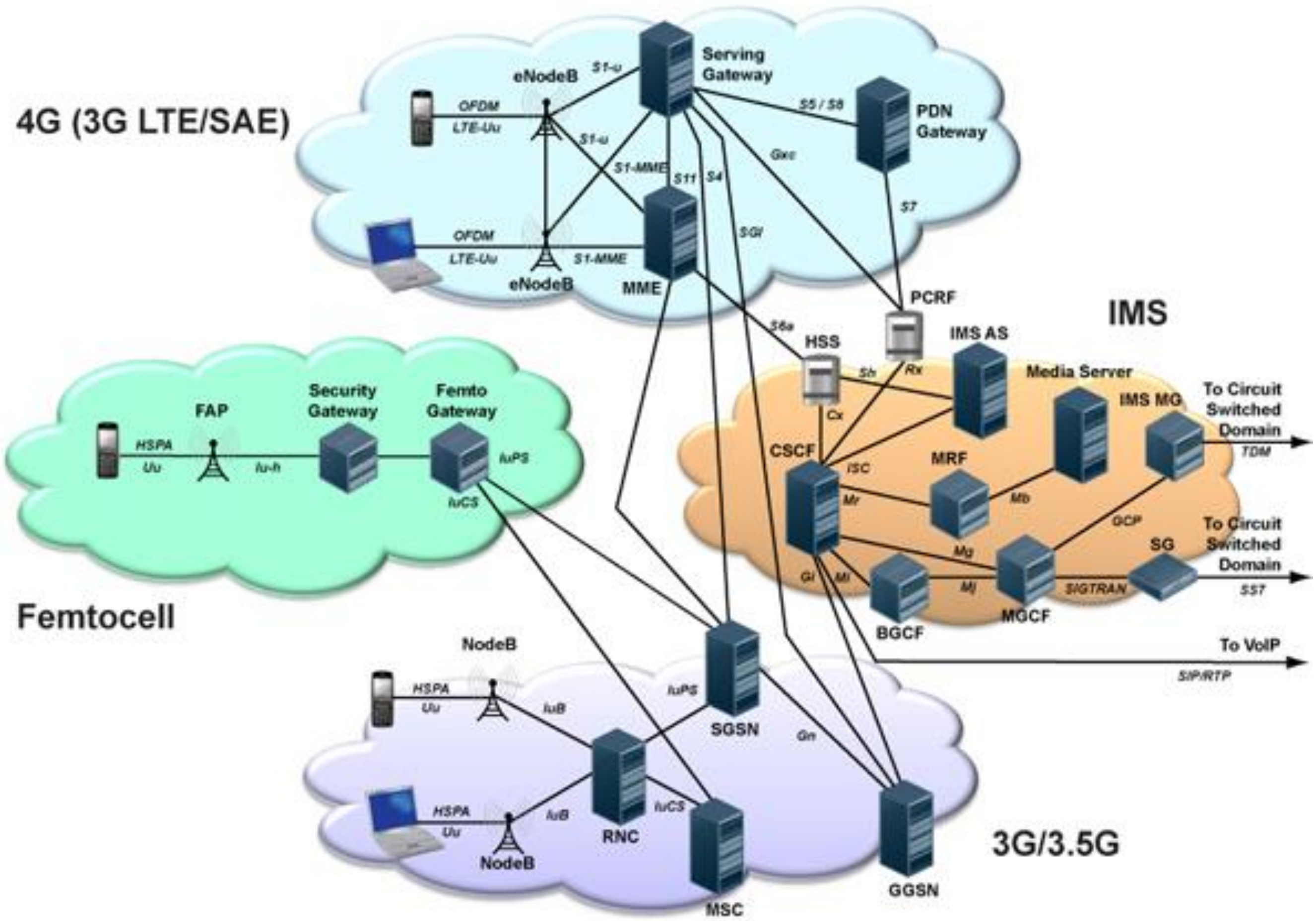


Figure 17. Connections of LTE networks. Reconstructed from [17]

Henceforth, what will be analysed is a set of technical possibilities that are currently being developed. All the figures, diagrams, and options will come from reputed sources, (such as companies, manufacturers and standardization bodies) that have proposed their own solutions to this challenge.

On the other hand, there are some research lines and/or new elements that are common to each of the actors. In every case, it will be explained whether it is being discussed a common element or, on the other hand, a proposed solution by a concrete actor that is still not widely accepted.

3.2 Innovations in RAN

The radio access is perhaps the most important topic to discuss in the development of 5G. There are many directions that might be followed, many of which will not be possible to achieve without achieving a set of own requirements. Here are the main innovations that are being studied related to the 5G RAN:

a. **New Radio**: An important term that must be introduced at this point is **New Radio** (NR). This concept is related to the new technology that must ensure communication in higher frequencies (see section 2.2). Its importance is tremendous as for the necessity of the higher frequencies to achieve the lower latency and multi-service KPIs. At this point Sony Ericsson , one of the companies at the head of 5G investigations, defined 5G in the IEEE Summit in February 2017 like "5G = LTE evolution + NR ".

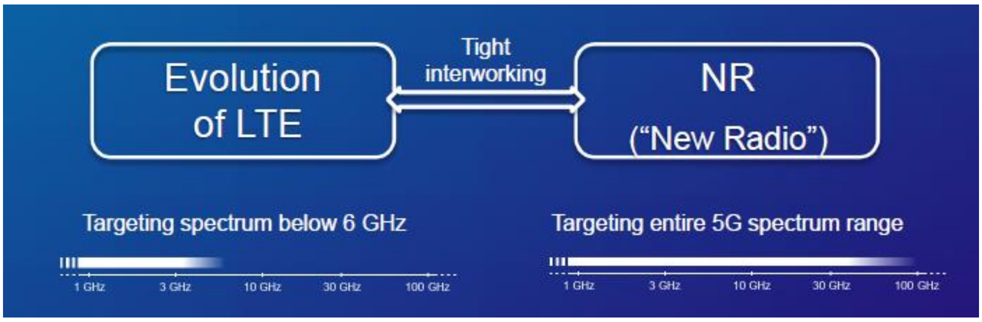


Figure 18. Development lines in RAN. Source: Ericsson

b. **Femtocells**: cells are the actual basic units for accessing the network via radio. Although it varies depending on the case of use, the general trend seems to be making these cells smaller. This has the obvious drawback of limiting the coverage area, but sets a good bunch of pros: less interference, higher rate and the possibility of connecting to more than one antenna (multi-antenna techniques). Internet and phone providers are starting to deploy femtocells that are incorporated in new Internet routers[25]. Femtocells are designed to work in small places and environments: homes, small offices or businesses… In this sense, internet providers (who are also mobile operators) have invested money in an important asset whose main advantage is the cost reduction: not only in the deployment of more RRHs (Remote Radio Heads), but also in the

[25] New femtocells (that incorporate cellular coverage and Wi-Fi access) are sustained by convergent contracts in which the user decides to hire mobile and internet service with the same provider.

routers'(antennas') maintenance and power costs, which are necessarily paid by the customers. Furthermore, there is another big advantage: the saving of frequencies. As they include Wi-Fi connection technology, part of the mobile traffic is coursed via Wi-Fi (see paragraph below). Therefore, the NR will be less congested and more frequencies will be available.

c. **LTE/Wi-Fi interworking:** in relation to point b, the fair coexistence and compatibility of LTE and WLAN (Wireless LAN, another name to refer to Wi-Fi), especially when working in similar frequencies, is quite important. Before approaching this issue, it must be introduced the concept of LTE License-Assisted Access (LTE/LAA), which is the attempt to exploit unlicensed spectrum bands (5GHz) only for LTE. As explained in previous paragraph, Wi-Fi may work either at 2.4 or 5 GHz, last of which frequencies belongs to this unlicensed spectrum. Hence, interference becomes a major issue to deal with when working in this kind of bands, and technological solutions must avoid interferences. There are two main procedures to achieve so: LTE-WLAN Aggregation (LWA) and LTE WLAN Radio Level with IPsec Tunnel (LWIP) (both defined in 3GPP's Release 13).

- **LWA:** In LWA, a mobile handset supporting both LTE and Wi-Fi may be configured by the network to utilize both links simultaneously. It can be deployed without hardware changes to the network infrastructure equipment and mobile devices. For a user, LWA offers seamless usage of both LTE and Wi-Fi networks and substantially increased performance. For a cellular operator, LWA simplifies Wi-Fi deployment, improves system utilization and reduces network operation and management costs. LWA can be deployed in collocated manner, where the eNB and the Wi-Fi access point are integrated into the same physical device (e.g. new femtocells) or in non-collocated manner, where the eNB and the Wi-Fi AP or access point are connected via a standardized interface. The latter deployment option is particularly suitable for the case when Wi-Fi needs to cover large areas and/or Wi-Fi services are provided by a 3rd party (e.g. a university campus, airport), rather than a cellular operator. Either way, part of the traffic is coursed through the Wi-Fi. Thus, OPEX costs are reduced, as explained above.

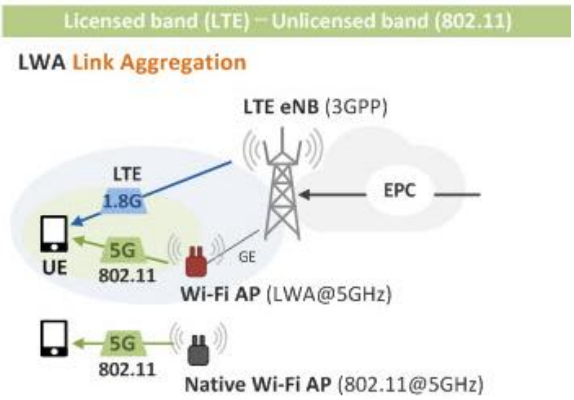


Figure 19. LWA scheme. Source: [18]

In figure 19, it can be seen the link aggregation between the two frequencies' paths. All the traffic ends up in the EPC. Hence, synchronism and redundancy are easily controlled. Moreover, as shown in the second case, the Wi-Fi AP (access point) may also be configured to receive and route both LTE and Wi-Fi traffic.

• **LWIP:** LWIP is also controlled by the eNB(based on the reports send by the WLAN system), but now the WLAN is hidden from the EPC. It therefore enables the use of legacy WLAN infrastructure and reduces the time to market for operators. Moreover, as a virtual tunnel is set up between UE and the eNB, traffic supported is

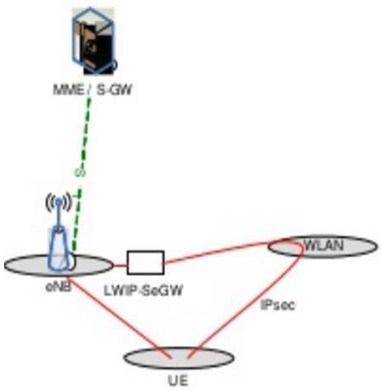


Figure 20. Scheme of LWIP. Source: Release 13, 3GPP

Table 6. Comparative between LWIP ¬ LWA. Source: Release 13, 3GPP

	eNB control	WLAN measurements	Offload granularity	WLAN traffic direction	Feedback/flow control	Fast WLAN authentication	WLAN infrastructure impact	New network nodes
LWA	Yes	Yes	Split bearer	DL only	Yes	Yes[2]	Yes[4]	WT
LWIP	Yes	Yes	Bearer[1]	DL + UL	No	No[3]	No	LWIP-SeGW

1. When a bearer is configured to use IPsec, LTE DRB configuration remains, however eNB is not expected to send packets on LTE and IPsec simultaneously, as LWIP does not support re-ordering
2. After connecting to WLAN, LWA UE only performs 4-way handshake (if network uses the eNB based authentication)
3. After connecting to WLAN, LWIP UE performs WLAN native 802.1x EAP/AKA authentication, IP address acquisition and IPsec tunnel establishment
4. Impact due to eNB based authentication mechanism, if used by network. Optional UE feedback mechanisms (as opposed to network feedback) allow to limit WLAN infrastructure impact of LWA

bidirectional, unlike LWA, where the traffic is only supported in its downlink. LWIP's biggest drawback is the static reservation and use of the virtual circuit (i.e. the IP tunnel), what is reflected in table 2 in the lack of LWIP's fast authentication.

d. **Flexible-RAN:** there are a couple of concepts that must be explained to understand this term. First, the **Distributed RAN** (D-RAN) is the next step of the current deployed RAN. This term refers to the scheme in which RHHs are distributed and work apart when processing information to later communicate with other RRHs and the corresponding eNBs and perform multi-antenna and MIMO (Multiple Input Multiple Output) techniques in communication. This RAN holds several disadvantages, some of which are the inefficiency and redundancy in signalling and error control, and intercellular interference due to the high amount of RRHs that must be deployed.

On the other hand, **Centralized-RAN** (C-RAN), often referred to as **Cloud-RAN**, is a proposed architecture for future cellular networks, defined as a "Clean, Centralized processing, Collaborative radio, and a real-time Cloud Radio Access Network"[26]. It consists on a centralized BBU (Base Band Unit) pool to which high number (around 1000) RRHs connect. Assuming that RRHs connect to the BBU via fiber, spectrum bands are liberated, whilst in D-RAN are used to connect different RRHs. Therefore, interference is highly reduced and operations are centralised in order to reduce delay.

[26] This was the definition given by the China Mobile Research Institute

Hence, it may be deduced that a flexible RAN must come along with a C-RAN. Eliminating interference constraint and reducing eNBs' CAPEX provides mobile operators of a flexibility they could not have before. Flexibility's importance relays in the so called 'tide effect', which explains that, because users are mobile, the traffic of each BTS fluctuates (called 'tide effect'), and as a result, the average utilization rate of individual BTS (Base Transmitting Stations) is pretty low. However, under D-RAN scheme, these processing resources cannot be shared with other BTS. Therefore, all BTS are designed to handle the maximum traffic, not average traffic, resulting in a waste of processing resources and power at idle times. Hence, flexibility will allow operators to design more efficient BTS/RHHs and move them if necessary without having to change the entire deployment.

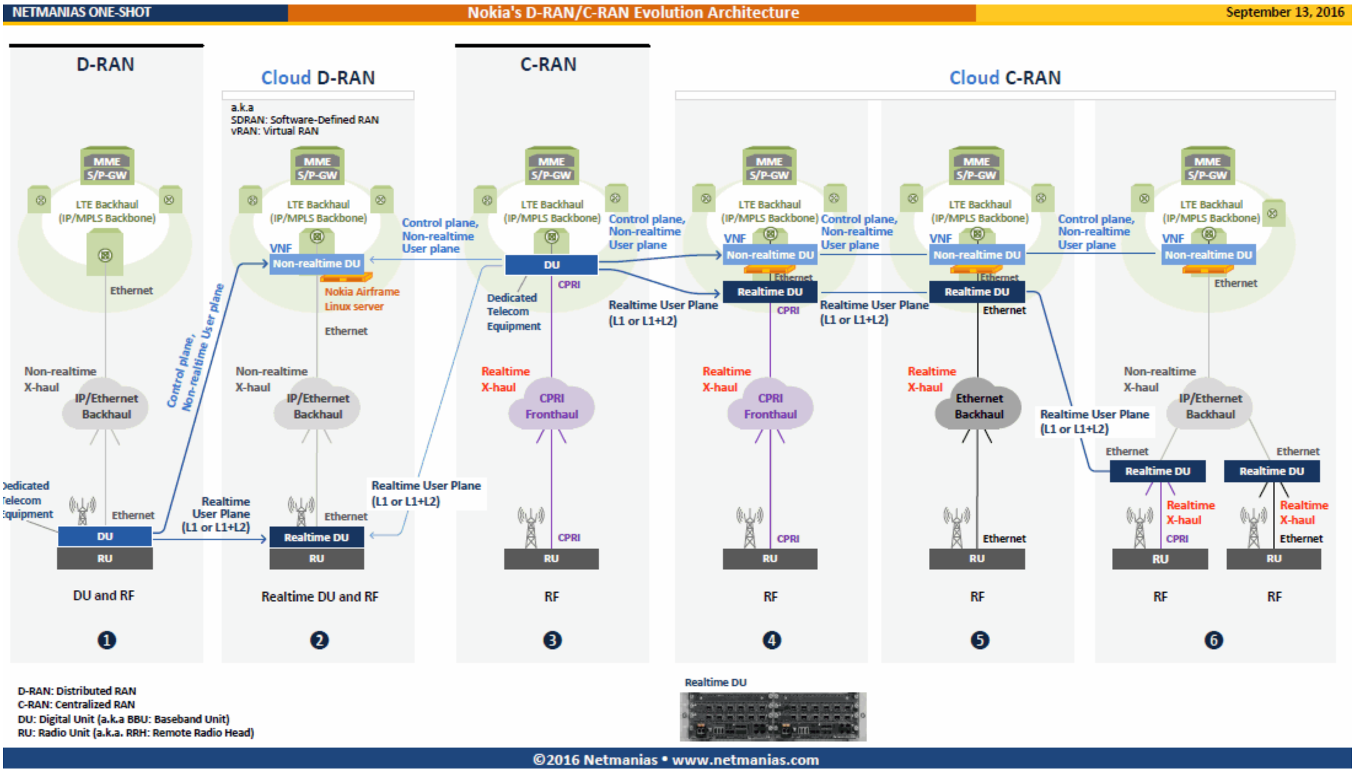


Figure 21. Nokia's D-RAN/C-RAN evolution of Architecture. Source: shown on figure

As shown in figure 21, these trends pretend to be the elimination of dedicated BTS in favour of a centralised BTS connected (or actually inside) to the EPC. In the last steps (4, 5 and 6), optimization follows the path of splitting the BTS tasks into real-time and non-real-time ones, so as to classify and better process traffic.

e. **Massive MIMO:** several issues implying multi-antenna interworking, LTE/Wi-Fi interworking, and higher speeds have been addressed. These envisaged goals will not be achieved without a basic concept in RF: massive Multiple Input-Multiple Output transmissions. The adoption of massive MIMO systems will necessitate more efficient interference management schemes. Massive MIMO is pretended to be achieved through **Non-coherent Large Scale MIMO**, as for the nature of the future communication channels, where channel status information is useless to estimate, [19]. Due to the new types of frequency bands (like micro and millimetre waves) that will be used, small cells will become even smaller and denser than in current setups. The smaller cells are, the more of them will be approachable to connect simultaneously (MIMO) and get a higher speed.

3.3 Innovations in Core

As for the other side of the network, there was (and still remains) doubt between two possible development lines; shall the existing core evolve and morph into a new core, or would a new different core coexist and interwork with the currently deployed one? Furthermore, new protocols were meant to be developed in order to efficiently course and process the new types of traffic that 5G will trigger. The main new features that 5G's core will have are the following ones:

a. **First Step- a basic EPC:** the first issue to be addressed is the backwards compatibility with eNBs. Hence, a frozen EPC at 3GPP Rel. 11 stage enabling the interconnection with real LTE eNBs and off-the-shelf smartphones is essential to ensure that every user can enjoy 5G accessing the core through the previous radio access, [20]. This idea answers to the dogma of considering the enhancement of LTE as a fundamental part of 5G. The envisaged figures represented in figure 5 will not be achieved by only improving the radio speed and bandwidth. Optimizing the core operations will significantly lower the delay. Hence, with the same radio access, higher speeds will be managed, without losing compatibility with out-dated devices.

b. **Common Converged Core:** another important trend to be managed is a basic set of common functionalities installed in a common core, which is meant to be as small and lean and possible. This core is pretended to be "open access", allowing the service providers to use a common physical infrastructure provided by a network provider, making the deployment of parallel physical networks unnecessary, [21]. Many of these functionalities will be set up as software (see section 3.3 in which all the issues about softwarization will be addressed). This common core is meant to be the evolution of the LTE-A core, adding the necessary functionalities (i.e. new software and hardware modules) and enhancing those LTE modules that must be updated to fulfil 5G requirements. Another obvious basic feature is the core's transversality in order to process all kinds of traffic coming from very different types of devices, and different service providers.

c. **Network Slicing:** without any doubt, network slicing is one of the most ambitious goals for the 5G core. By definition, network slicing consists on dividing those functionalities that can be separated into independent slices. In this sense, by doing so, scalability of each of the functionalities becomes independent as well, and therefore it also becomes ensured. Furthermore, unnecessary functionalities (e.g. coursing an unnecessary kind of traffic depending on the geographical area) can be eliminated without affecting the rest of the deployment as long as the proper slice of the network is removed. Slicing will allow operators to split a single physical network into multiple virtual networks.

Since the concept of network slicing was initially proposed to be adopted by the 5G core network (CN), NGMN (Next Generation Mobile Networks) uses the term "end-to-end (E2E) network slicing" to refer to the overall system design concept, including both CN and RAN aspects. In that context, network slices must fulfil a set of requirements such as the need for sharing and efficiently reusing resources (including radio spectrum, infrastructure, and transport network); differentiation of traffic per slice; visibility of slices; protection mechanisms among slices (a.k.a. slice isolation); and support for slice-specific management. The support for E2E network slicing appears as one of the key requirements in 3GPP, it is still under discussion how exactly network slicing would impact the RAN design, on both the access network and user equipment (UE) segments, although these concepts are currently under investigation[27].

Figure 22. Network slicing diagram. Source: Ericsson via Inside5G [22]

As shown in figure 22, network slicing will depend on the cases of use (see section 4.2) defined for 5G. It must be recalled that 5G pretends to be integrated within and alongside many vertical industries. Therefore, traffic assortment will define the amount of slices that might be differentiated for 5G, all of which will have some different needs to be met.

[27] Extracted from [21]

d. <u>Control-User Plane Split</u>: this is another feature that will come along with network slicing. First of all, it is essential to differentiate both of them. The user plane (sometimes known as the data plane) carries the network user traffic (i.e. traffic a user generates, either it is sent or received). The control plane carries signalling traffic. Control packets originate from or are destined for a router. Control traffic is important to many communication aspects such as error control, header information or channel status information.

Traditionally (until LTE, where they are partially separated), they both have been implemented together in the core. However, network slicing will provide the opportunity to separate them in order to accomplish the goals above mentioned. But, regardless of the slicing, this planes' separation holds a strong significance. Separation allows control- and user-plane resources to be scaled independently, and it supports migration to cloud-based deployments. By separating user- and control-plane resources, the planes may also be established in different locations. As an example, LTE MME scalability, which depends on the number of users, is quite different from routers' scalability, which is also affected by the demanded bandwidth for each application and service.

The apparent line to be followed will consist on deploying a central node to carry out all the control operations in order to make management and operation less complex; and distributed closer-to-user nodes to carry out data plane functionalities to improve speed figures.[28]

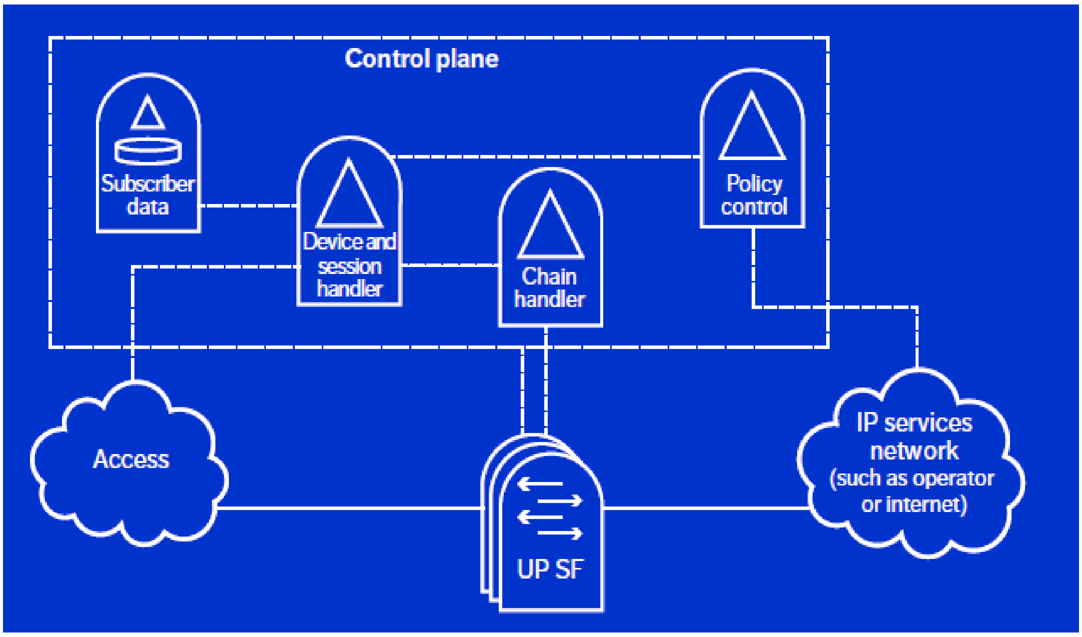


Figure 23. Ericcson's 5G control-plane architecture. Source: [23]

[28] Although the line seems to be clear at a high level approach, service providers differ on what are the actual functionalities that can (or can't) be decentralized.

Figure 23 above shows control-plane architecture thought by Ericsson. Regardless of the rest of the image, the control plane carries out part of the total functionalities that the packet core (EPC in LTE) must take care of. May it be noticed that policy control is currently in charge of PRCF, and Subscriber Data is in charge of HSS and EIR, for example. As it can be seen, control plane takes all the functionalities that can be centralized, whilst data/user plane, named as UP SF (User Plane Short Functionalities) in the figure, keeps itself (its nodes) decentralized and close to each of the accesses points.

e. **Flexibility Features**: it may sound redundant because flexibility itself is an applicable dogma in the whole 5G network: flexibility does not hold less importance in the core than it holds in the RAN. It comes as a consequence of the previous explained innovations, especially because of control-user plane split and network slicing. Being a unique packet-based network is another important trigger to achieve flexibility: all the kinds of traffic and therefore the interfaces are packetized (i.e. adapted to course packet traffic). Packet switching provides of many flexible mechanisms (Dijkstra's algorithm, BGP, OSPF)[29] to route the traffic through different paths.

f. **LTE-5G Protocol Stack:** this point serves as a good introduction of next section 3.4. Throughout this entire section 3.3, many concepts have approached the idea of making the core (and the entire network) more flexible, partly by implementing some modules as software. Besides, all the networks are deployed accordingly to the ISO-OSI or TCP/IP Protocol Stack, either in which the multiple layers can be easily differentiated: physical, data link, network… Gathering these two ideas, 5G protocol stack concept arises in order to standardize the future deployments.

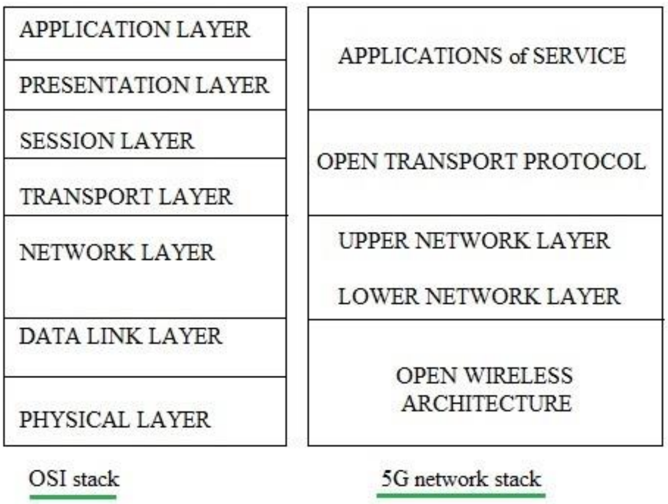


Figure 24.OSI & 5G network stack. Source: rfwireless-world

[29] BGP: Bit Gate Protocol & OSPF: Open Shortest Path First

In this sense, the top levels hold great impact as for the importance of software in 5G networks. As seen in figure24, the correspondence between OSI and 5G stacks gives evidence of the importance of standardization. It shall be remarked that both wireless architecture and transport protocol layers are named as **open**, referring to the idea discussed in paragraph b about a common converged core. Moreover, the lowest layers (physical and data link) give the idea of an open wireless radio (NR).

3.4 New technologies: NFV & SDN

The two previous sections dissected 5G networks into its two main segments to analyse them separately. The two techniques that will be discussed in this section are applied end-to-end (E2E), may them be configured in one segment or another. They are intimately related to the last points mentioned before related to the need of softwarized networks. But not only do these technologies come as consequence of leaner and more flexible networks. They also come as a response to the economic challenge of reducing costs. Needless to say that software holds economic advantages against hardware; both CAPEX and OPEX (especially the first ones) costs are minimized. Therefore, service providers will seek to market software-based networks to reduce costs and hopefully increment their revenues.

3.4.1 The importance of The Cloud

There is no doubt that cloud technologies (i.e. storage and computing) are one of the main triggers not only of 5G networks but also of many other networks and technologies. Traditionally, the cloud used to be understood as the part of any communications network where information was stored and accessed (i.e. datacentres full of servers that are connected to the Internet, through which information amongst users travels).

This term has evolved due to the new trends that have upraised. As for today, The Cloud is defined as the data on servers that can be accessed by the Internet. As traffic increasing in the figures shown in figure 1, the trend has become to store data in the net: in the Cloud. So users keep their information not only in their physical devices, but also in datacentres full of servers that are connected to the Internet. Hence, information is available everywhere, in any device, as long as you can access the net.

Next step that is being developed (and is currently in use) is cloud computing. After offering information storage, operators realised that not only data but also functions could be held in the cloud. Cloud computing could be defined as the provider's HW and SW services on the Internet[30] (the "cloud"). Cloud providers replace in-house operations and are invaluable for companies, no matter their size or applications.

[30] Internet is the most known example of cloud. However, there are other private clouds that work for specific companies, ministries, etc. 5G will comprise both the public cloud (Internet) and private clouds.

Cloud servers can be configured to handle tiny or huge amounts of traffic and expand or contract as needed. Cloud computing comprises software, infrastructure and platform services[31]. There are two main advantages that cloud computing features:

• **Self Service**: services and functions are accessed upon request and everywhere.

• **Scalability**: servers can be quickly configured to process more data or to handle a larger, temporary workload such as Web traffic over the holidays.

3.4.2 SDN: Software Defined Networks

Born as a research topic in universities and colleges, SDN was invented due to the student's need to change the software in the network devices each time they wanted to try a new approach, protocol or functionality. They came up with the idea of making the behaviour of the network devices programmable, and allowing them to be controlled by a central element. This lead to a formalization of the principle elements that define SDN today:

• Separation of control and forwarding functions

• Centralization of control

• Ability to program the behaviour of the network using well-defined interfaces

The next area of success for SDN was in cloud data centres. As the size and scope of these data centres expanded it became clear that a better way was needed to connect and control the explosion of virtual machines. The principles of SDN soon showed promise in improving how data centres could be controlled.

As for 5G networks, SDN will have to be used alongside NFV to achieve the flexibility and context-awareness features that the network is meant to have. Nevertheless, SDN will be specifically used to achieve the three points above stated. To ensure scalability, these virtualized functions (see next section) could be managed by a pool of SDN controllers or orchestration platforms. The architectural organization and physical locations of these SDN controllers/orchestrators should be optimized to allow high cost efficiency and high performance.

OpenFlow: Why is a standard needed?

OpenFlow (OF) is a communications protocol that gives access to the forwarding plane of a network switch or router over the network. It enables network controllers to determine the path of network packets across a network of switches. The controllers

[31] Three main platform services are SaaS (Software as a Service), IaaS & PaaS (Infrastructure and Platform as a Service) and FaaS (Function as a Service)

are distinct from the switches. OF is considered one of the first SDN standards[32]. It originally defined the communication protocol in SDN environments that enables the SDN Controller to directly interact with the forwarding plane of network devices such as switches and routers, both physical and virtual (hypervisor-based), so it can better adapt to changing business requirements, [24]. To work in an OF environment, any device that wants to communicate to an SDN Controller must support the OF protocol. Through this interface, the SDN Controller pushes down changes to the switch/router flow-table allowing network administrators to partition traffic, control flows for optimal performance, and start testing new configurations and applications.

As SDN started to gain more prominence it became clear that standardization was needed. The Open Networking Forum (ONF) was organized for the purpose of formalizing one approach for controllers talking to network elements, and that approach is OF. OF defines both a model for how traffic is organized into flows, and how those flows can be controlled as needed. This was a big step forward in realizing the benefits of SDN. This argumentation gains even more importance if readers recall section 3.3. The concept of an open core which aims to be partly (or even fully) software-based needs a standardized protocol stack. OF is just a part of that stack, as for the wider sense of the stack itself which compels more layers.

3.4.3 NFV: Network Functions Virtualization

NFV was created by a consortium of service providers (AT&T, China Mobile, BT, etc.). Service providers attempted to speed up deployment of new network services in order to advance their revenue and growth plans, and they found that hardware-based appliances limited their ability to achieve these goals. They looked to standard IT virtualization technologies and found NFV helped accelerate service innovation and provisioning. With this, several providers banded together and created the European Telecommunications Standards Institute (ETSI) NFV division, since ETSI was (and still is) the main European standardization body. The creation of ETSI NFV resulted in the foundation of NFV’s basic requirements and architecture, [25].

NFV is used to implement certain network functionalities in software rather than in hardware. Unlike net virtualization, where the entire network is virtualized, operators can choose which functions to virtualize, all of which are implemented in servers, routers, switches and the rest of routing elements. Therefore, NFV reduces networks’ dependence on hardware devices (what increases power efficiency and reduces costs), enhancing its scalability and customization.

In next two figures 25 and 26, it can be seen a typical 5G scenario and the proposed solution using complementarily SDN and NFV. Firstly, by using NFV, each enterprise’s router function is virtualized leaving only an access point in each of the enterprises,

[32] Other protocols aim the same open-concept software protocols, such as Open Daylight

whose only function is to give users access and measure traffic (no routing). Secondly, by using SDN, the control and data plane are separated. Whilst without SDN, both data and control traffic packets would have to be forwarded by the (virtualized) centralized router, with SDN only the control pane traffic (i.e. the routing functionalities) has to go through the router. Data plane traffic (information) is forwarded by an optimized data plane machine whose software is enhanced and specialized in forwarding data packets.

As stated in sections 3.2 and 3.3, the development line to follow is the centralization of both the radio and the core network. Henceforth, scalability this scenario is optimal because every function that may be scaled is located in the central office and depends on software enhancements or changes.

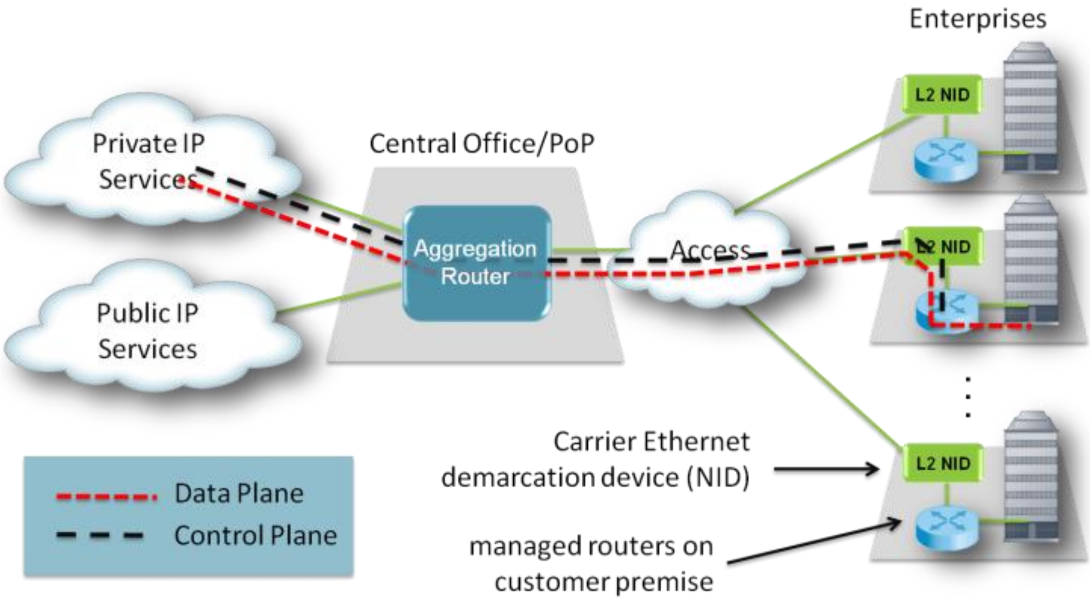


Figure 25. Network deployment before SDN & NFV. Source: [25]

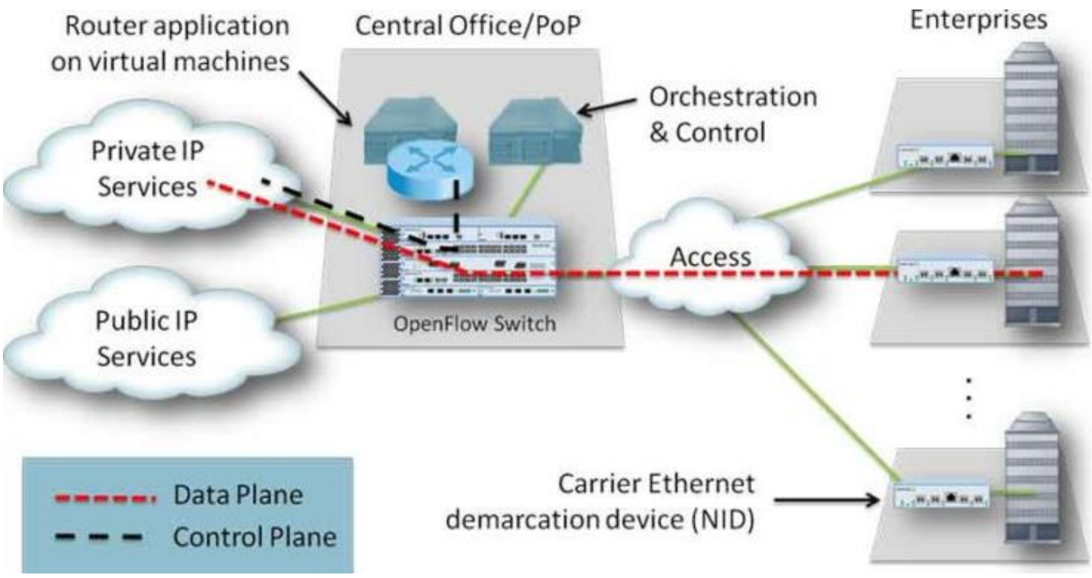


Figure 26. Network deployment after SDN & NFV. Source: [25]

3.5 Network architecture

Once studied each of the main innovations within each of the segment and layers in future 5G networks, it's now time to take a view at the overall networks architecture. Needless to say that there are only proposed architectures by some of the most important operators, standardization bodies, etc. Therefore, these proposed architectures gather all the features stated, but are not worldwide accepted or deployed.

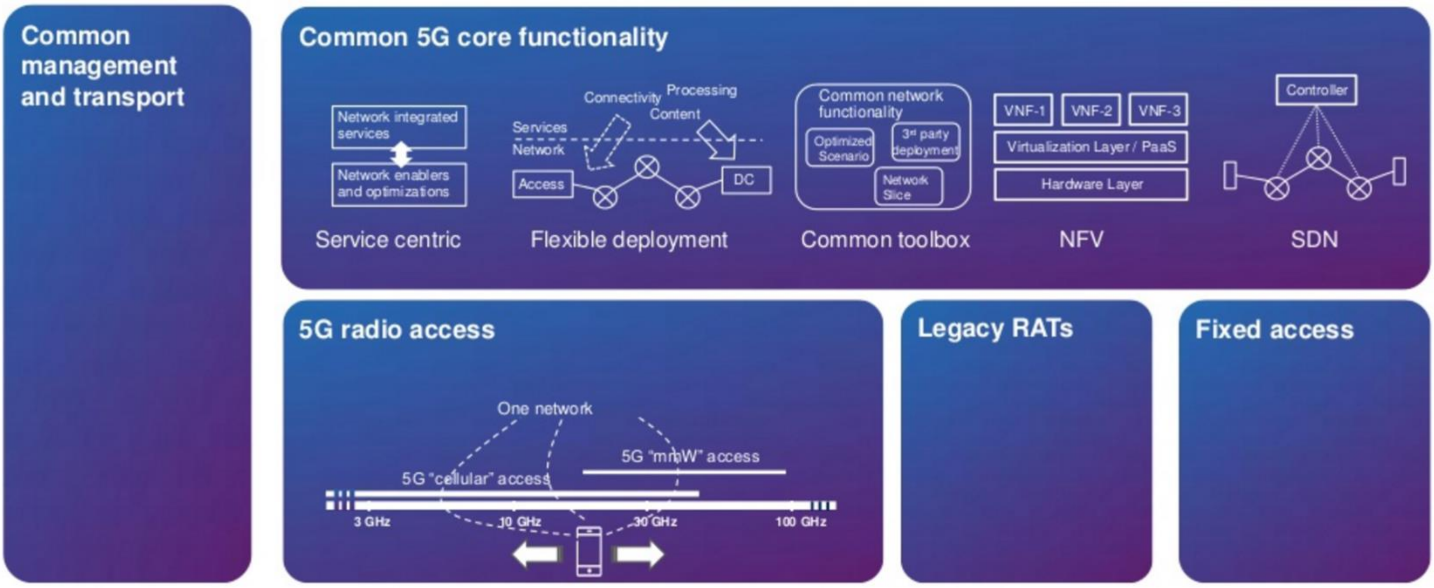


Figure 27. High level architecture proposed by Ericsson

Figure 27 shows Ericsson's the proposed architecture. Some aspects to be pointed out:

• Both management-transport and core functionalities are defined as "common", as explained throughout this entire section.

• RAN includes not only 5G radio access (NR), but also legacy (enhanced) radio access technologies from LTE and fixed access, such as femtocells and WLAN ones.

• NFV and SDN are included within the core, as the affect almost every module belonging to the core.

• Centralized Core and flexibility, both mantras mentioned all around this section, are included in the core features.

In next figure (Fig. 28), the proposed architecture by BT along with 3GPP can de dissected from another perspective. In this case, the different modules that constitute the architecture are shown. This architecture includes and enhances LTE modules (see figure 16) and includes new modules (that may be virtualized and implemented on software) that provide further functionalities.

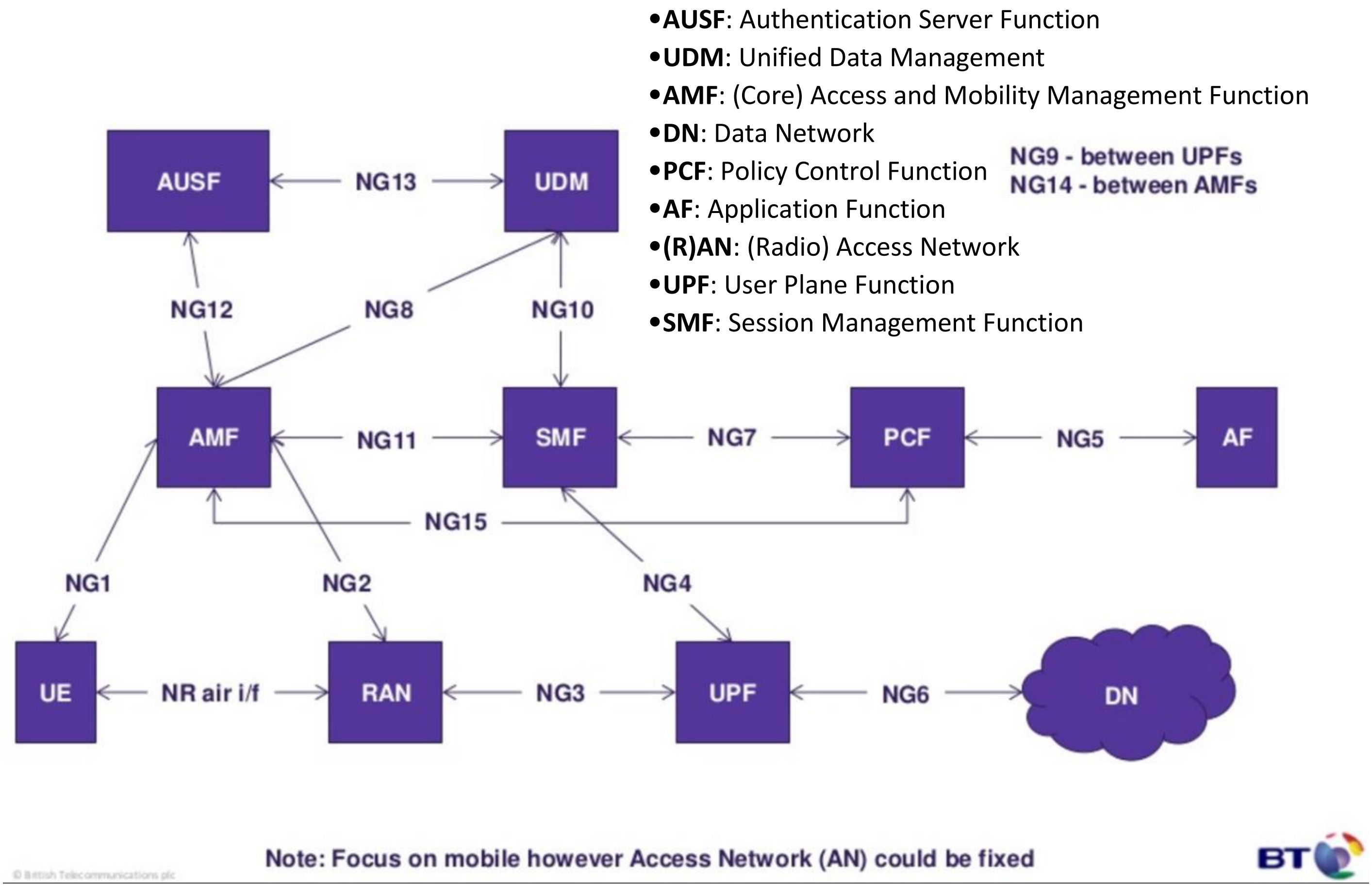


Figure 28. Proposed architecture by BT & 3GPP. Source: [26]

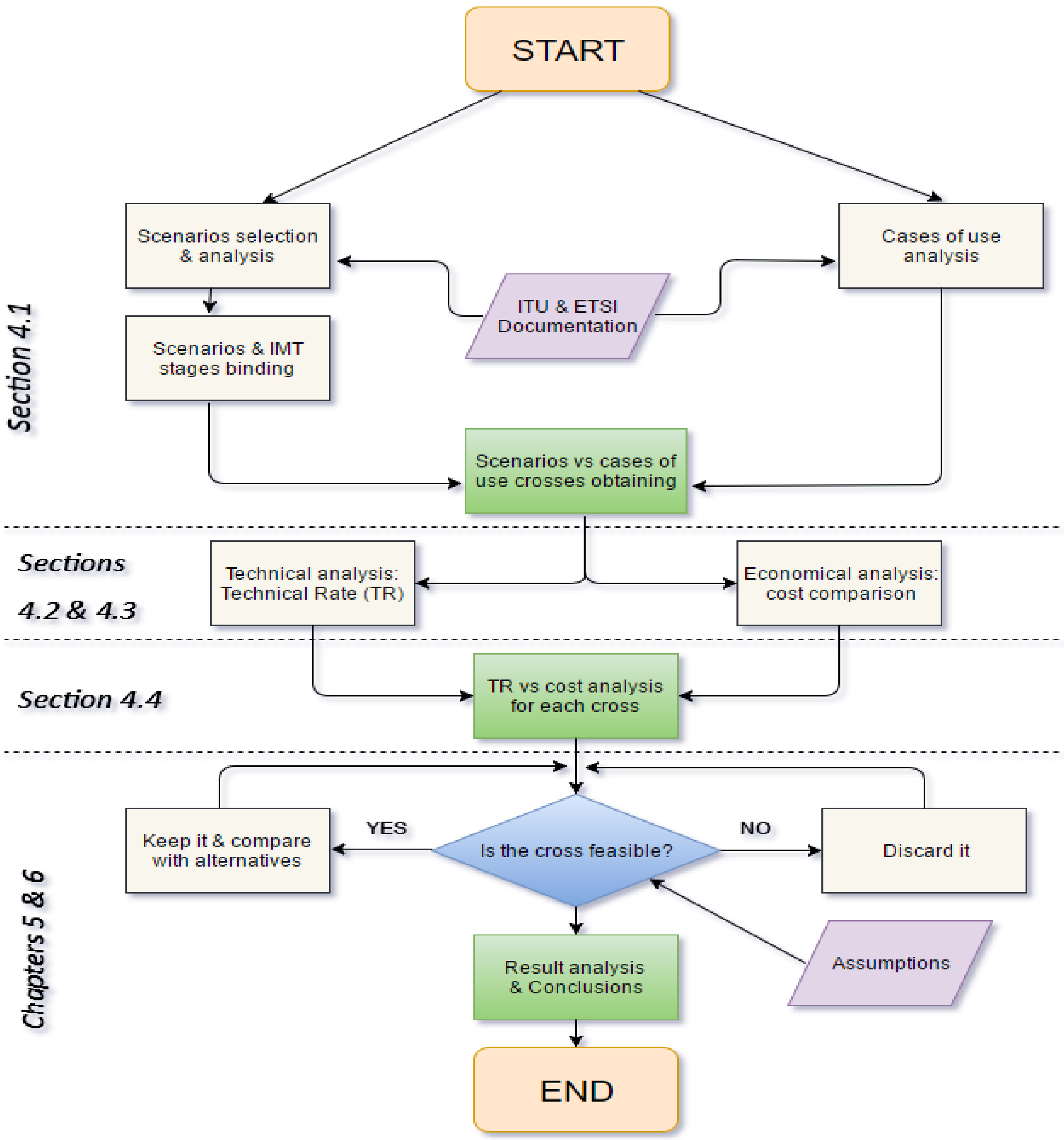
Some aspects to be pointed out about this modular architecture:

• Planes are split. UPF is separated from DN, which is drawn in a cloud shape as for being in the cloud and implemented using both SDN and NFV.

• All LTE modules have its corresponding module in this network serving its same function (MME-AMF, HSS-UDM & AUSF, S-GW-UPF, P-GW-SMF, and PCRF-PCF)

4. Techno-economic analysis

So far, the document has introduced the relevant theory for 5G. These previous chapters have introduced the main issues and concepts that are indispensable to both understand and elaborate this techno-economic analysis rigorously. Henceforth, all the concepts listed will be used to make an exhaustive analysis, approaching it by two sides: technical and economic. From the technical side, it will be discussed the need, pros and cons of the application of both the new and enhanced technologies introduced in chapter 3. As for the economic perspective, it will be focused on the cost analysis on deploying the aforementioned technologies.



Use cases/scenarios	Scenario A	Scenario B	Scenario C
Use case 1	X	X	X
Use case 2	X	X	X
Use case 3	X	X	X

(Information in each cell depends on the part of the analysis)

Figure 29. Flow diagram and example matrix of the whole analysis process

Figure 29 shows the methodology flow diagram. This methodology starts with the proposal of diverse deployment scenarios that represent the different network deployment and configuration possibilities. Next step is to set a chronological order amongst these scenarios. This order will allow creating a relation between a scenario and a particular set of milestones to achieve, and therefore create a binding between a time period and an achievable goal.

Parallel, cases of use that will be common to all the scenarios will also be proposed. A case of use is defined as the standardization of a typical trend many users are involved in. For example, one case of use will be the traffic carried out by Internet of Things. Therefore, the analysis will be dissected by crossing each of the scenarios with each of the cases of use (these situations will henceforth be named **crosses**) and analysing them separately at a first instance.

At this point, there are two parallel analyses to carry out, both of which will be contrasted as the last step of the process:

• On the one hand, the technical analysis will study the different technical features of each of the crosses and compare both the ones that belong to the same case of use and the ones that belong to the same scenario. This will be done using a particular unit called Technical Rate (TR), which will be explained in section 4.2

• On the other hand, the economical analysis will study the different costs that each of the crosses would incur in.

Therefore, the analysis will be completed by gathering the economical and technical analyses to discuss the commitment between the technical features and the costs to face for each of the crossings. Shall it be noticed that its last steps will be fulfilled in chapters 5 and 6 (Result Analysis and Conclusions respectively). Furthermore, all the ITU and ETSI information explained in previous chapters will be used to elaborate the actual analysis together with some assumptions that will be made (sufficiently justified by consulting external sources and/or by personal criteria).

From a top approach, this methodology could be seen as a matrix to fulfil like in fig. 29, whose rows and columns are cases of use and scenarios respectively; the content of each cell will depend on the type of analysis (technical or economical) discussed.

The main reason to justify this solution is that this document deals with technologies to be deployed in 2020. When making predictions, scenarios and cases of use are the best way to contemplate all the options (still vague) in order to get a wider perspective on the topic. Furthermore, all the players involved in this game have defined quite similar cases of use and scenarios. Therefore, with (more or less) agreed conditioners and triggers, it will be easier to classify these solutions and assess them.

4.1 Methodology: scenarios & cases of use

As explained in this chapter's introduction, scenarios and cases of use will be the selected methodology. The aim of this section is to justify the selected cases of use and scenarios that will constitute the "matrix" to analyse.

4.1.1 Scenarios selection & analysis

A scenario is a deployment option for the future mobile network. As studied in chapter 3, a network is considered as the interworking of two deployment modules: the radio access network (RAN) and the core network. As it was explained in the previous chapter, a new core and a new radio access network are intended to be deployed. Furthermore, as the user and control plane are separated (and their traffic can be coursed through one or another radio), the number of possible scenarios increases up to 12 options, which are[33]:

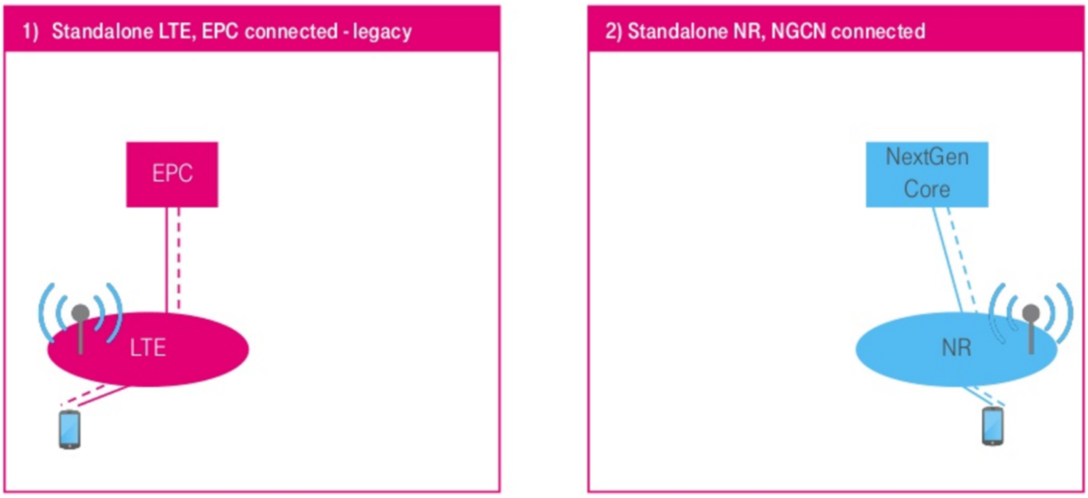


Figure 30. Scenarios 1 & 2. Source: [27]

These two scenarios are standalone options shown in Fig. 30. Hence, in both cases, both segments of the network belong either to 4G or 5G. Scenario 1 represent legacy: it represents how 4G mobile networks look nowadays and how will they look before the start of 5G deployment. Beyond representing the current situation, it also represents a quite feasible option for the early upcoming years.

Scenario 2 represents the other standalone option. This scenario is very unlikely to be seen, since 5G includes both the deployment of a new network and LTE enhancing. In other words, not using LTE networks for not so-demanding applications once it has been enhancing would be a waste in all senses.

[33] These scenarios are proposed by 3GPP in their documents RP-161266 in collaboration with DeutscheTelekom. Therefore, all come from source [27]

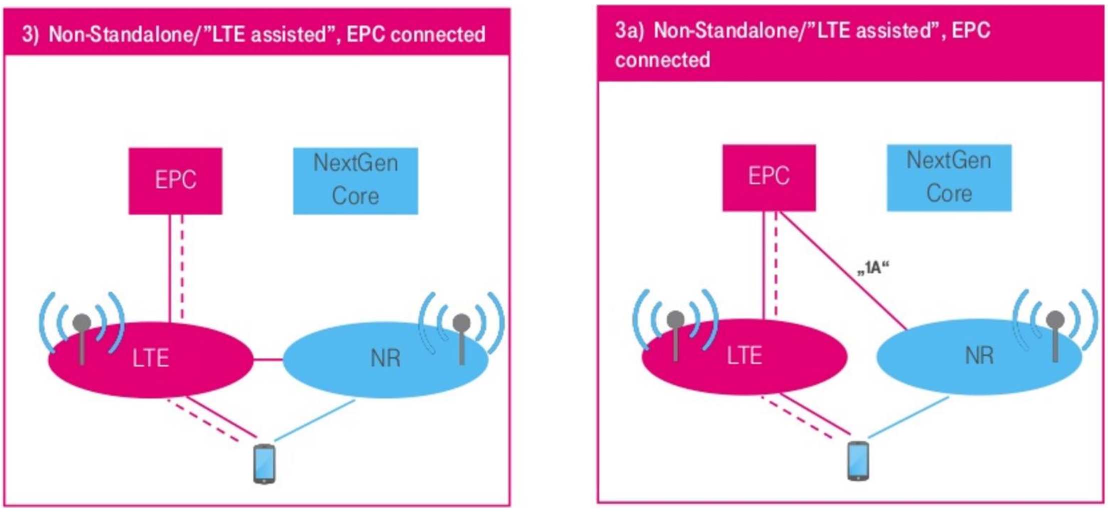


Figure 31. Scenarios 3 & 3a. Source: [27]

Scenarios 3 and 3a show transition scenarios (Fig. 31) including 4G and 5G elements considering the same deployment scheme. The difference between them is the RAN-Core connection. This is not trivial, since it would mean that NR would only have to interwork with other radio nodes or, on the contrary, it would have to be programmed to transfer information (coming only from the data plane) also to the core.

*Note that for this and the rest of figures, data plane information is represented with a solid line while control plane information is represented with a dashed lin*e.

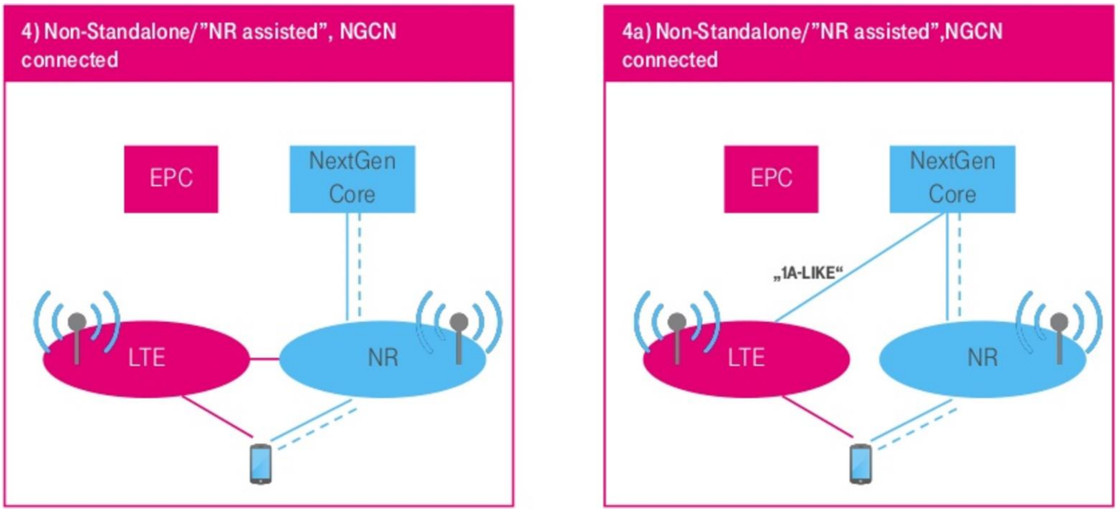


Figure 32. Scenarios 4 & 4a. Source: [27]

The scenarios represented in Fig. 32 may be considered as the goal to meet for 5G. In this case, traffic is coursed through the NR, using LTE when necessary but keeping control plane traffic information being coursed via NR. Hence, any UE shall be able to connect to LTE as well as NR. As for the core, NGCN will be used due to its benefits over EPC. However, an enhanced EPC shall be used to support the core as it may be able to cover same cases of use.

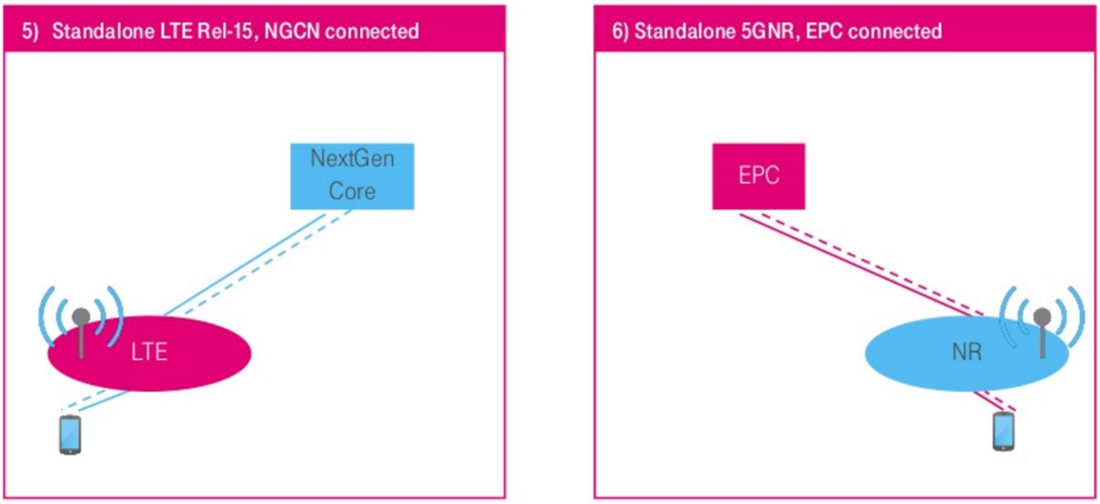


Figure 33. Scenarios 5 & 6. Source: [27]

The cross possibilities are represented in this scenarios. At first appearance, they do not seem to be very relevant inasmuch as they are not representative. They could be used in very specific cases of use but, once again, it would be rare to handle them.

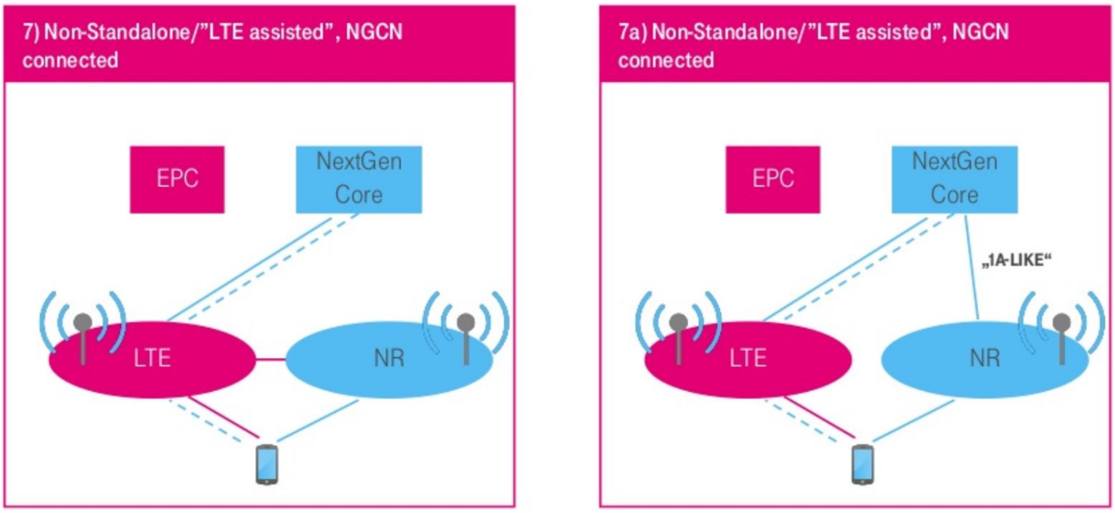


Figure 34. Scenarios 7 & 7a. Source: [27]

These scenarios represent a feasible intermediate point in the evolution towards 5G (Fig. 34). The difference between them is the same as for scenario 3. In fact, the difference between scenarios 7 and 3 is the core to which the radio access networks are connected. In this case, they are connected to the NGCN, what seems a plausible option as the intermediate step between the enhancement into LTE-A PRO network and the deployment of fully operating 5G networks. Note that EPC is represented since it has already been enhanced.

In this case, the control plane information is always coursed through the LTE RAN (eNodeBs). As a matter of fact, NR would only course data plane traffic. This scenario could perfectly use a developing and premature NR to provide a backhaul or support network in some areas.

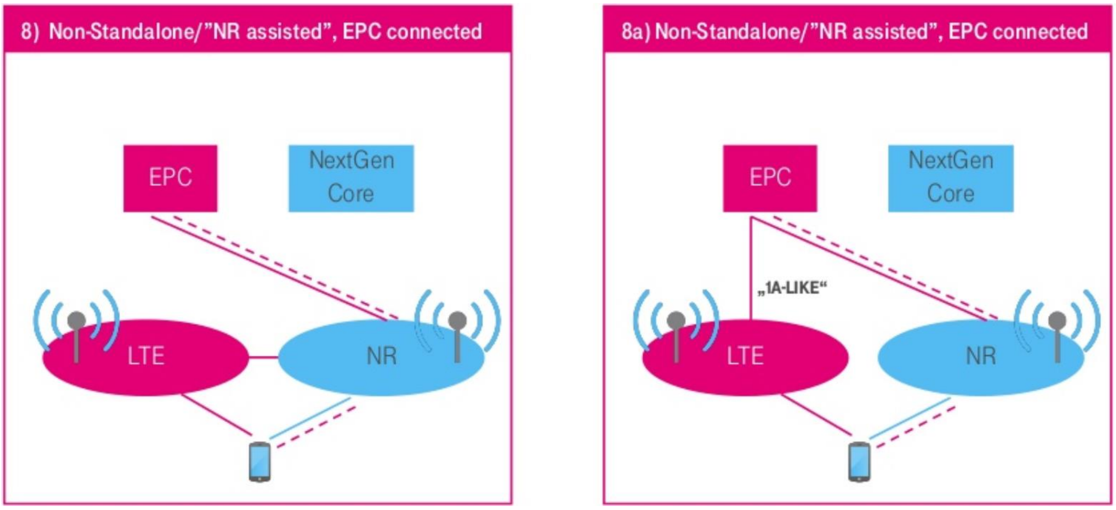


Figure 35. Scenarios 8 & 8a. Source: [27]

These scenarios represent another option for an intermediate point in the evolution towards 5G (fig. 35). In these cases, NGCN isn't connected to any RAN, so it can be considered non-deployed or under deployment.

Scenarios selection

There will be three chosen scenarios that will represent three phases: the initial point, intermediate stage and final stage for 5G networks. They must unconditionally be as much representative as possible. They will be renamed to facilitate the understating of the rest of the chapter as well as the remaining ones. These are:

• Scenario A: Starting point

This scenario will be the same that the explained scenario 1 from 3GPP. Its representativeness is obvious: it represents the deployment architecture of an LTE-A Pro network.

In other words, it is the enhanced 4G network that will be able to cover the less demanding cases of use and the network from which all the deployment options will be contemplated.

Therefore, both the core and the radio access network must be both considered sufficiently enhanced to be regarded LTE-A Pro network segments.

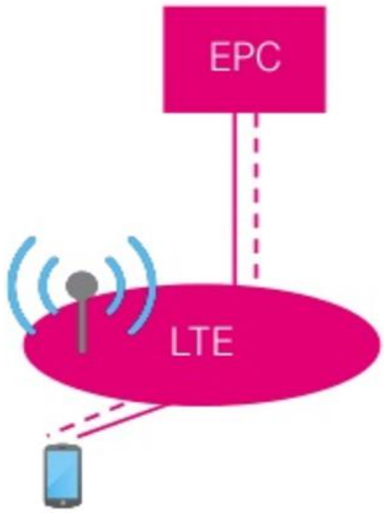


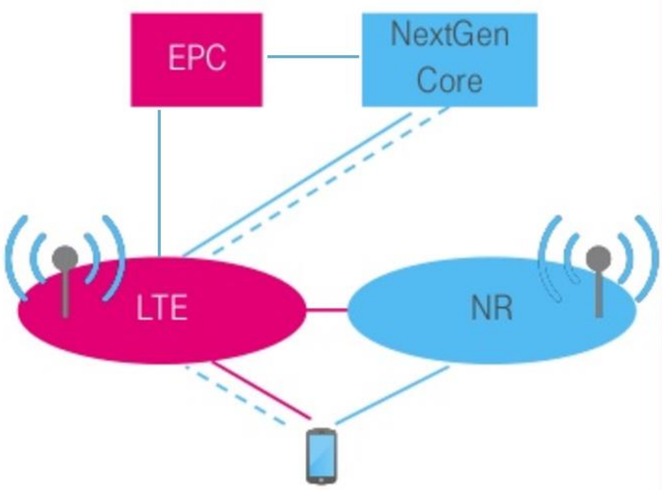
Figure 36. Scenario A: Starting point

• Scenario B: Intermediate Stage

There were several candidate scenarios that could be chosen for this intermediate stage: scenarios 7 (and 7a), 7 (and 3a) and 8 (and 8a). These three scenarios have the common feature that they have at least one of the network segments deployed while the other one is still being developed.

As seen in chapter 2, RAN standardization will depend on several milestones (e.g. WRC-19), ITU-3GPP coordination and spectrum regulatory changes upon national and supranational institutions. On the other hand, core network will be less sensitive to that process as for the flexibility granted by software technology such as NFV and SDN.

Therefore, scenario 7 will be chosen upon some assumptions. First, EPC will have already been deployed, so the core costs will be in concept of deploying the NGCN. Secondly, NR will be partially deployed[34]. Thirdly, LTE RAN will be able to connect to the two cores, and both cores will interwork to optimally deal with every case of use. With these assumptions, Scenario B looks like:



• NR can only connect to eNodesB

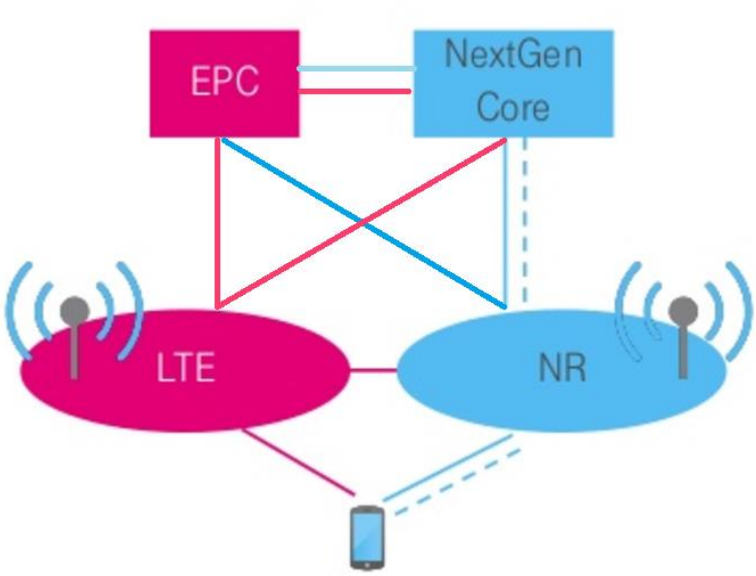
• EPC and NGCN interworking

Figure 37. Scenario B: Intermediate Stage

• Scenario C: Final Stage

This stage is clearly represented by scenario 4 (or 4a). It is based in the latest and best technologies working complementarily with the previous LTE enhanced ones. At this point, both NGCN and NR will be fully deployed. For this scenario it will also be assumed that RANs and Cores can interwork amongst them in benefit of network's efficiency and performance. As discussed in chapter 3, core will be software-based, what will facilitate interoperability between them and with the RANs. It looks like:

[34] NR deployment ratio (i.e. how much is NR deployed) will be explained in next sections.



• 4 connection possibilities

• Best performance through 5G modules

Figure 38. Scenario C: Final Stage

Scenarios discard

At this point, it has been justified the election of the three scenarios. However it has not been justified the discard of the remaining ones, although it was advanced in each description. Following the same representativeness criteria:

• Standalone scenarios 2, 5 and 6 were discarded because very few applications would be covered by a standalone 5G scenario and a core-RAN from different generations' combination.

• Scenario 8 deployment would mean that the NR would be deployed before the NGCN. As for the small feasibility of this option, it was discarded.

4.1.2 Scenarios & IMT Stages binding

It is now time to set a chronological order for these 3 scenarios (A, B & C). Figures 6 and 7 from section 2.1.1 showed MT-Advanced is expected to be concluded between 2018 and 2019. Hence, it can be assured that the first network architectures to fulfil IMT-Advanced requirements will consist on LTE-A Pro networks (Scenario A).On the other hand, IMT-2020 will be lastly fulfilled by scenario C networks, which include the new core and new radio belonging to future 5G networks.

The upcoming topics that will be discussed in WRC-19 are the tendering of the bands over 6 GHz, the possible reassignments of bands below 6GHz and the use of 5 GHz Wi-Fi as a licensed service, amongst others. Regarding the importance of the spectrum over 6 GHz for 5G NR that has been discussed all over the document, it seems quite logical to assume that NR deployment will depend on the resolution of WRC-19. Hence, it does not seem feasible to have NR deployed until 2020 (see section 2.2.3).

Lastly, scenario B must be placed anywhere between scenarios A and B. As explained in chapter 3, 5G Spectrum comprised both mm Waves and 700-900 MHz bands. Assuming that bands over 6GHz won't be standardized, only part of the NR will be ready to be deployed. Therefore, as for now, it will be assumed that scenario B will be ready to be deployed in 2019 while scenario A will be deployed by year 2018. This explanation is best shown in next figure 39.

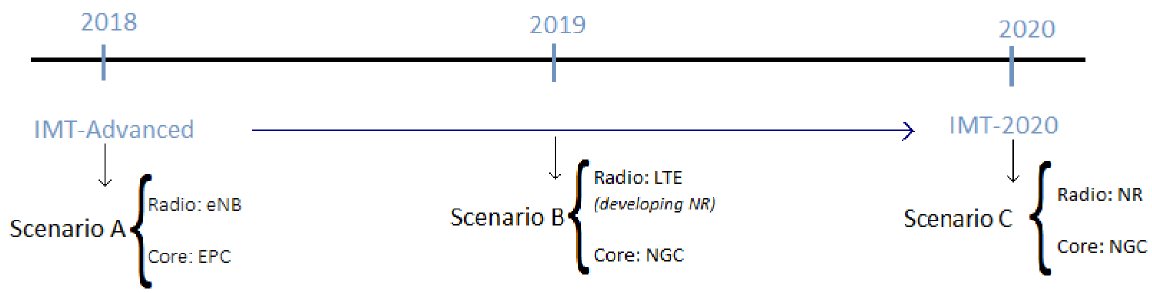


Figure 39. IMT planning alongside scenarios timeline

Although it is justified by the reasoning explained in the previous paragraph, the actual binding of a scenario with an IMT Stage (each of which is represented by a year) has **not** been directly taken from any source, since it has be done following this reasoning line. Hence, it is an assumption made upon self-research and reasoning. However, this strategy will simplify and enable a powerful advantage, since it will be possible from now on to give a chronological sense and order to any improvement stated in next sections.

As a summary, table 7 shows in a very lean way the composition of each scenario in terms of deployed architecture:

Table 7. Scenarios' architecture breakdown

	Scenario A-2018	Scenario B-2019	Scenario C-2020
Core	OLD(EPC)	NEW (NGC) Assisted by EPC	NEW (NGC) Assisted by EPC
RAN	OLD (LTE)	OLD (LTE) Assisted by developing NR	NEW(NR) Assisted by OLD (LTE)

With this binding, the scenarios' analysis is terminated. Retaking the simile at the beginning of the section, the columns of the matrix to be filled in have just been described. Now, the rows of this actual matrix (the cases of use) have to be defined as well. Once done this, the analysis itself will be ready to be carried out.

4.1.3 Cases of use analysis

Next step is to define the cases of use that will constitute the matrix from which the technical and economic analysis will depart from. Furthermore, it will be discussed which are the most important figures in IMT-2020 (refer to section 2.1.1) for each of the cases. This will make it easier to turn the spotlight on the correct keys in every case of use. Beneath, the same figure that was shown in chapter 1 is displayed again in order to get a good perspective of the cases of use:

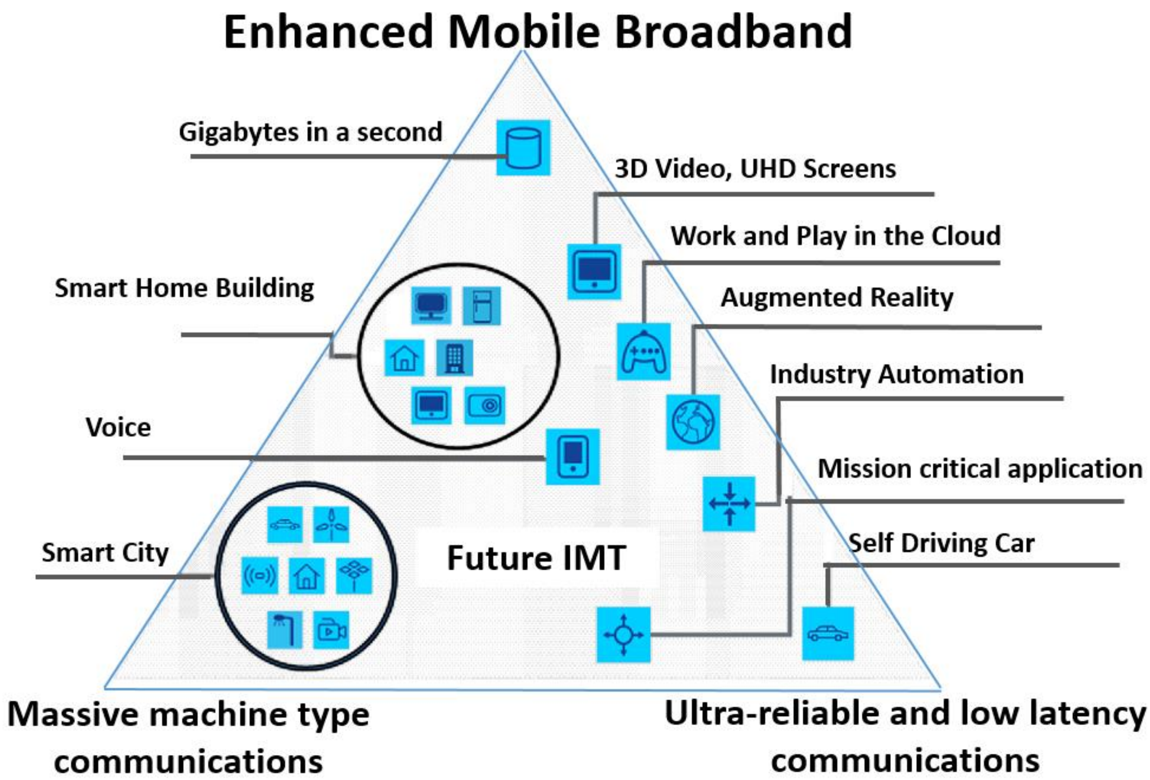


Figure 1 (listed in chapter 1): Groups of usage scenarios for IMT-2020. Source: ITU-R, 2015

Once again, these cases of use have been taken from various sources that have been listed throughout the document: ITU, Ericsson, Nokia, 3GPP, ETSI, etc.[35] Hence, they are considered to be sufficiently justified, as the main players inside the industry have commonly reached an agreement on which are the cases of use, yet they haven't reached it on how to address them. Although there are around 20 different cases of use, they can be well separated into 3 cases of use (that include the rest within them), which are:

[35] The cases of use were initially proposed by ITU-R. Since them, the rest of the main actors have used them in their respective documents.

a. Enhanced Mobile Broadband (EMBB)

This case of use shall be the first one to be achieved-implemented. It is actually the natural evolution of the mobile broadband, the traditional case of use towards which every previous generation has been oriented. It will address the challenge of giving consistent coverage in any kind of environment and structure. This will include from small cells in offices and homes (addressing the issue of interworking with WLAN deployments in different spectrum bands) to macro cells, may they be deployed in large dense cities or rural environments. Furthermore, EMBB will be diversified to adjust the new necessities and innovations that are going on: augmented and virtual reality, video streaming traffic in high definition…

Some specific cases of use within EMBB are [28]: enhanced outdoor wireless broadband, fixed wireless broadband deployments, augmented and virtual reality or extending mobile computing (in which the C-RAN will be key). The actual figures from IMT-2020 that most affect EMBB and are indispensable in its performance are:

- Edge Throughput: as for the applications that require more data to be processed and sent, edge throughput is undeniably essential to provide a continuous service in all of the specific cases of use within EMBB.

- **Peak Data Rate**: this factor will be important as well as the edge throughput (experienced data rate) in the same sense: the higher the peak data rate is, the higher will be the throughput for each of the users. As it was defined, peak data rate is defined in optimal conditions, while the experienced data rate takes several interferences into account. But in any case, both are proportional.

- Area Traffic Capability: this factor will be important, especially when considering the typical cells that may be deployed for example in shopping centres, malls, etc. More devices will have to be covered in each of the cells, independent of the size of the actual cell. Otherwise more cells will have to be deployed, incurring in more costs as well as more problems in terms of spectrum.

- Spectrum Efficiency: as it was mentioned while EMBB was being introduced, EMBB will have to address important topics such as shared spectrum, interworking with WLANs, etc. And it will have to do so because the optimal bandwidths (e.g. in offices or homes) will be in use by Wi-Fi or other kinds of networks (either in licensed or unlicensed spectrum). Furthermore, as for the mentioned higher bandwidth needed to enhance the mobile broadband, it can be easily deduced that the more devices connected (with more traffic per device), the more bandwidth is needed. Hence, as for the spectrum's limitation, the only way to address the problem is to make it more efficient.

b. <u>Massive Internet of Things</u> (MIoT)

This is the second case of use to be achieved. Although mobile communications have traditionally been envisaged as a human-to-human idea, it has been proved throughout these recent years that the future is being oriented towards machine-to-machine (M2M) communications, also called MTR (Machine Type Communications). In this sense, IoT makes reference to these machine things that can transmit and/or receive useful information without it being oriented to the user.

The information involved in MIoT is not usually very important, in the sense that there are small blocks of information that can suffer a certain delay without affecting the system's performance. MIoT information is usually sent by simple devices such as sensors, simple state machines, switches, etc. Furthermore, MIoT is intended to be able to operate both in licensed and unlicensed spectrum.

At this point, a temporal solution has been almost finished using LTE-A Pro features: Narrow-Band IoT (NB-IoT)[36]. This technology standard has been focusing on whether IoT-M2M communications could be feasible without the new core and RAN, at least at its applications where the traffic is lowest. Hence, it focuses specifically on indoor coverage with low cost devices, where it has proved to be a pretty good solution to the IoT issue. To have a vision, the figures achieved by NB-IoT are the following ones:

Table 8. NB-IoT specifications. Reconstructed from Release 13

	DL Peak	UL Peak	Latency	Antennas	Duplex	Rx. BW	Rx. Chains	Tx. Power
NB-IoT	250 kbps	250 kbps	1.6-10 s	1	Half Duplex	180 KhZ	SISO[37]	20/23 dBm

Some specific cases of use within MIoT are [28]: asset tracking, smart agriculture, cities and homes, energy monitoring or physical infrastructure. With all this said, the most relevant figures to MIoT are:

- **Connection density:** as it was mentioned, due to the huge amount of devices and their ability to operate in licensed and unlicensed spectrum, connection density is essential (even assuming the small bandwidth that devices need).

- **Energy efficiency:** (shown in figure 5-5G flower). It is the most important as the long battery (around 10 years) of each device will reduce CAPEX to make it viable.

- **Cost efficiency:** related to energy efficiency, it will be very important as well to make cheap devices, what will enable economies of scale. It shouldn't be hard as long as the underlying technology shall be the simplest.

[36] NB-IoT is a 3GPP technology, whose specification was frozen at Release 13
[37] SISO: Single Input Single Output

c. <u>Mission Critical Service</u> (MCS)

This case of use is where 5G will show its complete potential. Consequently, it could trigger new and tremendous growths in sales. The kind of service provided in this case of use is a set of new future mobile applications whose information must be sent and processed as quickly as possible (i.e. make the latency almost zero). Hence, this type of traffic would have the highest priority compared to the rest of the cases of use. Furthermore, this information must be totally reliable and must have lack of errors, since the applications for which it will be used hold such an importance that a mistake would be determinant.

Moreover, there are other important features that this important traffic must include. First, it must be always available to access (what can be easily deduced by stating the case of use as “critical”). Secondly, as a consequence of the importance and priority, it must hold a strong security to ensure its reliability.

There are some examples on this case of use [28], most of which are being developed now: autonomous vehicles[38], drones, industrial automation, telehealth, smart grid[39]… In each of the cases, the importance of the information is such that it may condition whether a life is lost or not (e.g. wrong information in a self-driven vehicle). Hence, the most important IMT parameters related to this case of use are:

- Latency: it may represent the most important aspect as for MCS. As for the real-time applications for which MCS is intended, low latency is crucial to reach the goals and avoid failure, errors or losses.

- Mobility: for many of the cases which imply moving while transmitting, mobility is crucial. Retrieving the vehicle or the drone examples, mobility must guarantee the flow of information without which MCS can’t work. Moreover, as speed intended in future cars or drones keeps growing, so does mobility have to do.

- Edge throughput: this factor is important as long as it must not represent any constraint in the case of use development. In other words, it represents a basic need without which MCS would not have any sense.

As it was mentioned, this case of use is mainly represented by applications which belong to emerging markets. Hence, investments and Research and Development are crucial to determine the viability of these applications.

[38] Autonomous vehicles represent a new revolution in which investors hold great hopes. A particular case of use(traffic) is being studied to improve it at the most: V2X (Vehicle-to-everything)

[39] Smart grid is an electrical grid which includes a variety of operational and energy measures including smart meters, smart appliances, renewable energy resources, and energy efficient resources. [29]

4.2 Technical analysis: Technical Rate evaluation

Once disaggregated the problem, the next task is to analyse the influence of every Key Performance Indicator (KPI) stated in section 2.1.1 (figures 4 and 5) in each of the cases of use, the most important of which have already been discussed. To do so, a matrix will display a number ranging from 1 to 3, where 1 represents the smallest influence and 3 the biggest one. Influence 0 is going to be omitted because it is supposed that any KPI will have influence in every case of use (even if it is significantly small). Please remember that each of the cells represents a cross: a combination of a scenario and a case of use.

Figure 40 shows the given values (by ITU) to every KPI within each of the cases of use:

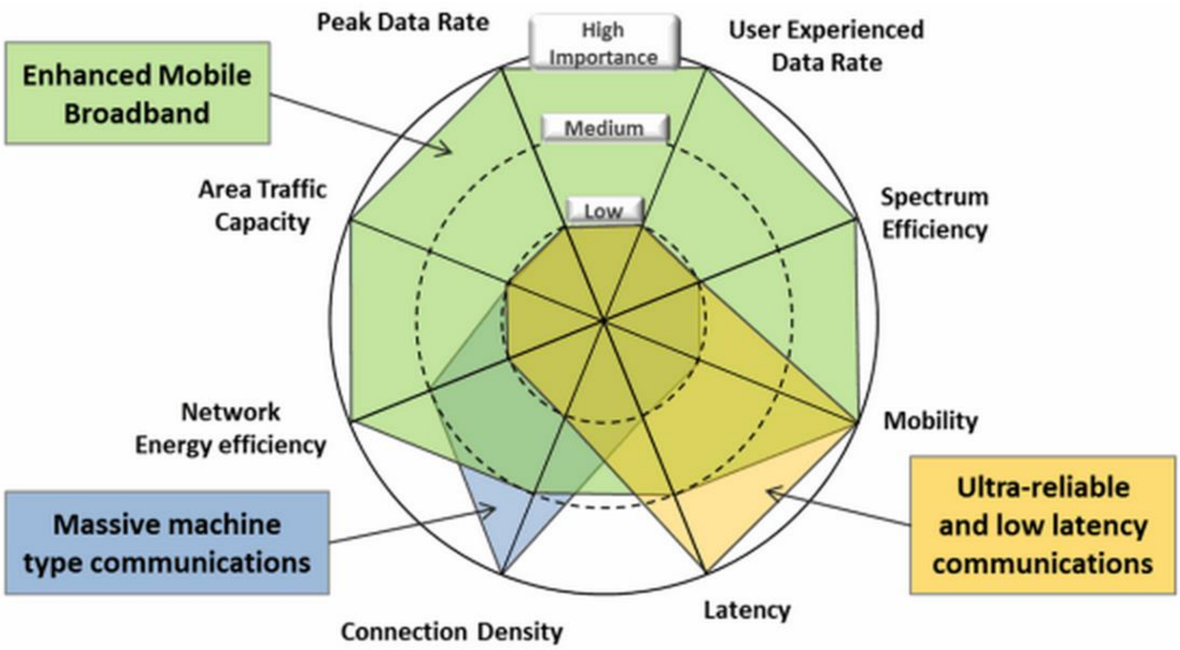


Figure 40. IMT requirements classified upon case of use. Source ITU-R, 2015

Therefore, it is easy to translate the three importance levels into its corresponding three values to fill in the matrix, according to Figure 5 (ITU). In next table, the mentioned matrix is displayed:

Table 9. Influence of KPIs in each case of use (named w_{KPI})

	Peak Data Rate	Edge Throughput	Spectrum Efficiency	Mobility	Latency	Connection Density	Network Efficiency	Area Traffic Capability	Average (out of 3)
EMBB	3	3	3	3	2	2	3	3	**2.75**
MIoT	1	1	1	1	1	3	1	1	**1.25**
MCS	1	1	1	3	3	1	1	1	**1.5**

1: Less important 3: Most important

An extra column was added at the right of the matrix, containing the average (i.e. the summation of the values of each column divided by 8). This average just gives a clue on the number of KPIs that are involved and how much they have to be improved. However, it does not exactly reflect whether a case of use is more demanding than other, even more if it is considered that each case of use comes from a previous aggregation of many others.

<u>What is the Technical Rate (TR)</u>?

The Technical Rate will be the measure through which the technical performance will be calculated. It actually is an invented measure that is obtained by weighting every KPI's importance in each case of use and multiplying it by the relative improvement of the actual KPI (and then making the summation). It gives the idea of the technical improvement in the transition between two scenarios (the current LTE and the selected amongst A, B, C) for a case of use. In other words:

$$TR_{Scenario\ i}^{Case\ of\ use\ j} = \frac{\sum_{l=1}^{8} W_i^{KPI\ l} \times Improvement_i^{KPI\ l}}{24} \quad ec(1) \qquad where:$$

(Please note that super index does not refer to exponentiation but to the actual KPI)

$- W_i^{KPI\ l}$ is the given value, ranging from 1 to 3, that reflects the importance of every KPI in each case of use.

$- Improvement_i^{KPI\ l}$ is the improvement rate (number of times that is better) each KPI compared between the two scenarios(LTE compared to A, B, C)

There are some features to be explained about this invented measure:

- It must be noticed that these technical rates don't have units. Firstly, it is quite logical as all the factors involved in the operations don't have them either. Secondly, the sense of the procedure is to compare different scenarios. Hence, as long as the units are the same for each scenario, it does not matter what the units are because they will be simplified in the operations.

- The sense of dividing by 24. As the weights rank from 1 to 3, to normalize the summation, the result will be divided by 8 x 3 = 24. This will result in more stable results. Nonetheless, as improvement rates are (quite) bigger than 1, the result will never be smaller than 1.

- TRs give us a clue on how good (comparatively) does a scenario work for a particular case of use. Nonetheless, it does not justify whether an actual scenario actually works for a case of use (i.e. each cases of use restrictiveness). That part of the analysis will be addressed throughout chapters 5 and 6 by making assumptions that will consider how restrictive a case of use is.

So, next step is to analyse the situations generated in each cross. The aim is to argument which is the technical rate of every case of use and scenario cross. It may sound obvious that any of the cases of use will be optimally achieved by developing scenario 4, which will have the newest RAN and core. Nonetheless, as mentioned at the end of section 4.1, achieving some of the cases of use without having to develop scenario 4 would result in a substantial cost saving.

The procedure to analyse each of the crosses will be as follows. First, it is going to be determined which is the ratio between the current situation (LTE-A Pro systems), IMT-Advanced and IMT-2020 (what was called $W_i^{KPI\,l}$). This will be done by extracting the corresponding data from figures 4 and 5. This will be also useful for the economic analysis. Next table shows this comparison[40]:

Table 10. Comparison between IMT-2020 & IMT-Advanced figures (Sources: figures 4 and 5)

(Units not displayed)	Peak Data Rate	Edge Throughput	Spectrum Efficiency	Mobility	Latency	Connection Density	Network Efficiency	Area Traffic Capability
Current LTE	0.3	6	1x	100	100	$2x10^4$	1x	0.1
Advanced	1	10	1x	350	10	10^5	1x	0.1
Advanced Improvement	**3.33x**	**1.66x**	**1x**	**3.5x**	**10x**	**5x**	**1x**	**1x**
IMT-2020	20	100	3x	500	1	10^6	100x	10
IMT-2020 Improvement	**66x**	**16.6x**	**3x**	**5x**	**100x**	**50x**	**100x**	**100x**
Advanced -2020 Improvement	**20x**	**10x**	**3x**	**1.4825x**	**10x**	**10x**	**100x**	**100x**

Next step is to calculate the actual technical rates of each of the scenarios, following the methodology stated in the explanation in the previous page. As for scenarios A and B, the following assumptions will be made:

- The change amongst scenarios doesn't change the influence of the KPIs with the cases of use. Therefore, the same formula can be used and only the Improvement will change.

- Scenario B is **not** exactly half the way between scenarios A and C, because it must be taken into account that NGC will be already deployed and NR will be under development. Hence, assuming that TR_{SA} and TR_{SC} will already be calculated, the task is to determine TR_{SB}. As there is no other way, it must be assumed that scenario 7 represents a slightly more advanced point that the mid-point between scenarios A and C. At this point, there will be two considered options:

[40] Units are not displayed, they may be seen in figure 4 (section 2.1.1)

a. Assuming that technical performance depends equally on the core and the network, it may be supposed that:

$$TR_{SB-O} = \frac{TR_{SA}+TR_{SC}}{2} + \frac{\left(\frac{TR_{SA}+TR_{SC}}{2}\right)}{2} = 3x\,\frac{TR_{SA}+TR_{SC}}{4} \qquad ec(2)$$

This equation represents that "three quarters" of the network will be deployed (enhanced core, enhanced RAN and NGCN). Under this optimistic assumption, this could be directly converted into the equation showed above.

b. Assuming in a pessimist scenario that RAN has more influence in the technical performance (which seems logical for some KPIS like Peak Data Rate, Spectrum Efficiency, Latency, Edge Throughput…), TR_{SB} must be such that it is it is true that $TR_{SA} < TR_{SB} < 3x\,\frac{TR_{SA}+TR_{SC}}{4}$. It seems reasonable to set a correlation such that the technical performance depends 65% on the RAN and 35% on the core, in other words:

$TR_{SB} = 0.65 \times TR_{RAN-SB} + 0.35 \times TR_{core-SB}\; ec(3)$ Where it is known:

$TR_{RAN-SB} = 0.75 \times TR_{RAN-SC}\; ec(4)$ and $TR_{core-SB} = 1 \times TR_{core-SC}\; ec(5)$

By weighting in the formula for each side, the new TR_{SB} is defined like:

$$TR_{SB-P} = \left(\frac{0.65 \times 0.75 + 0.35}{2}\right) \times (TR_{SA} + TR_{SC}) = 0.41875 \times (TR_{SA} + TR_{SC})\; ec(6)$$

Therefore, there will be to technical rates for scenario B: TR_{SB-O} & TR_{SB-P} where they represent the two assumptions (the optimistic and pessimistic situations respectively). This will set to possibilities for the intermediate scenario that will converge into one final scenario.

Next table will show the technical rates (TR) of each cross between case of use and scenario. Creating a simple script in Matlab, all the technical rates are calculated by setting the improvement percentages as constants and introducing in each case the weights assigned to the different KPIs in every case of use. This script will be very useful in next step in order to create the corresponding graphic. They will be displayed in (expected) chronological order:

Table 11. Technical rates correlations amongst scenarios and cases of use table

		Scenario A	Scenario B		Scenario C
			SB-O	SB-P	
Technical rate (TR)	EMBB	2.6863	38.6364	21.5704	48.8250
	MIoT	1.5204	18.0341	10.0691	22.5250
	MCS	2.2288	22.0028	12.2849	27.1083

There are some conclusions[41] that must be explained before going on:

- The lower technical rates in earlier scenarios may be misleading, as it is only a measure of the technical performance. Nonetheless, ON NO ACCOUNT do they mean that the cases of use requirements could be fully achieved. As an example, MCS will not be achievable in scenario A, as for the high latency and low data rate that eNodeBs grant.

- It must be recalled that there are around 20 identified cases of use for IMT. This means that some of them are midst the three big categories studied. Hence, this approach will not always reflect reality to all of them.

  This is the reason why EMBB has the highest technical rate with more than two times over the next case of use: as for the wide range of applications, there are a lot of KPIs that must be enhanced. Therefore, it will be more expensive (in terms of this kind of TR-related costs) to achieve it in any of the scenarios.

The corresponding graph that illustrates the results obtained by the operations has this appearance (all the overall conclusions will be explained in chapter 5 unless they are strictly necessary to understand the following steps of the analysis). There are two bars for each case of use, represented in blue on the left the optimistic scenario and represented in red on the right the pessimistic one:

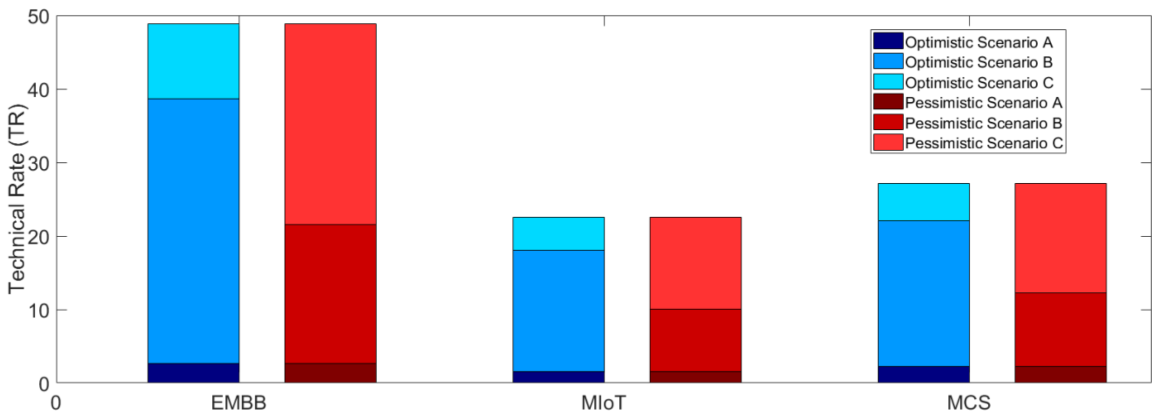


Figure 41. Technical rates correlations amongst scenarios and cases of use bars graph

It is now turn to discuss what cases of use should be attempted in each scenario. The objective on doing this is to find the order that would give the best commitment between cost and achievement of every case of use. At this point, the known data are: identification of scenarios, KPIs' weights in every case of use, technical rate of every case of use in every scenario and improvements of IMT-Advanced & IMT-2020 with regard to current LTE. To go on, there are other important data that may give a wider perspective to fully understand the analysis:

[41] Only necessary conclusions to go on with the analysis will be stated. The rest of them will be left to be explained in chapters 5-6 (result analysis and conclusions).

• Improvement rate between the TRs of a case of use among the different scenarios.

• Evolution of the improvement rate of KPIs through IMT-Advanced & IMT-2020

First, it will be useful to display the same graphic as in figure 41 but using lines instead of bars. The slope of each line representatively shows which case of use increments its technical rate most from scenario 3 to scenario 7. Each case of use is represented by a colour and a different marker (blue-circles for EMBB, green-stars for MIoT and red-crosses for MCS). Also, continuous lines are used for the optimistic scenario while dashed ones are used for the pessimistic case.

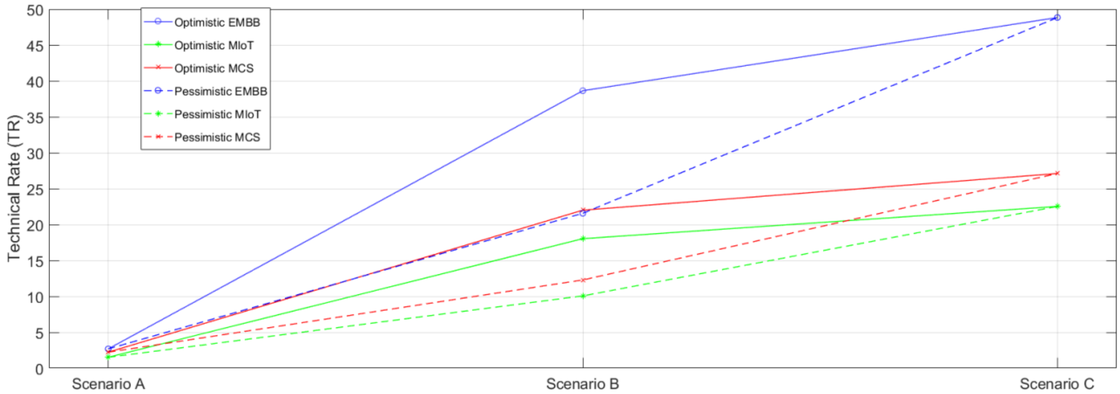


Figure 42. Technical rates correlations amongst scenarios and cases of use graph (lines)

As it can be seen qualitatively, MIoT is the one that increases its TR less (it has the lowest slope). Although it will be deeply discussed in next chapters, this starts to give evidence that, from a technical point of view, it is highly recommendable to start deploying, investing and researching in MIoT before the other two cases. This is explained by seeing that the slope, which represents the evolution of the performance (remember that it was weighted by the important of the correspondent KPIs), is less steep. Hence, there will be less difference in technical performance in MIoT.

No matter what assumption is considered (optimistic or pessimistic); it will be considered for every case of use. This means that only dashed lines can be compared with dashed lines and vice versa. Hence, in any case, qualitative comparisons and conclusions like the one on the previous paragraphs will be similar.

Secondly, to analyse the evolution of the TRs, the same kind of matrix will be displayed. However, it will be filled with the improvement ratio relation in terms of technical rates. It shows the improvement between scenarios: number of times that TR_{Si} is better than TR_{Sj} (previous scenario on the left, next on the right). This quotient will be named Technical Rate Correlation (TRC) and defined as:

$$TRC_{Si-Sj} = \frac{TR_{Si}}{TR_{Sj}} > 1 \; ec(7)$$

Once again, it must be noted that as there are two possible options for Scenario B, there will be two TRC_{SA-SB} and another two TRC_{SB-SC}. Furthermore, the greatest TRC for each scenario will be highlighted. Table 12 shows them:

Table 12. Technical rates correlations amongst scenarios and cases of use table

		TRC_{SA-SB}		TRC_{SB-SC}		TRC_{SA-SC} (total)
		Optim.	Pessim.	Optim.	Pessim.	
Technical rate correlation (TRC)	EMBB	14.3828	8.0298	1.2638	2.2635	18.1759
	MIoT	11.8613	6.6227	1.249	2.2388	14.815
	MCS	9.8723	5.5119	1.232	2.2066	12.163

As seen from the numbers above, EMBB is the one that improves most in both transitions between scenarios. Then, MIoT improves more than MCS in technical terms. Nonetheless, these numbers do not take into account possible constrains, for example: no matter if by these numbers it would be more recommendable to deploy MCS first if the KPI values are insufficient to give proper service in that actual scenario.

Therefore, next step is to discuss whether each of the cases of use might work properly within each of the scenarios. It may sound obvious that the three cases of use may be developed parallel. However, the aim is to find the optimum combination of resources and time to set an order of deployment. Until now, the TR has given a clue on how would a case of use perform in the different scenarios. This had a sense when comparing the different scenarios. However, as introduced in previous paragraph, the actual needs (absolute values) were not taken into account. As it is yet very difficult to give exact numbers on the proper performance of future cases of use, qualitative ideas will be discussed. Hence, some extra conditions and assumptions to be made:

• EMBB is the one that includes the most diverse set of cases of use. That is why so many KPIs hold such a big importance, many of which will increment up to 100 times compared to the current LTE networks. This does not mean that some of the cases of use within EMBB (such as enhanced voice/video calls) would be feasible in the early stages including scenario A and B.

• As for MIoT; the best scenario to implement it would be B. By this time, the new core would be already developed, while NR would be under development. This result is logical because the KPIs with less weight for MIoT have to do with peak data rate, throughput, etc., most of which depend on the RAN. Even with this conclusion, NB-IoT (which is already being used)[42] has proved to be a valid solution.

[42] 3GPP defined NB-IoT standard last June 2016, it is based on LTE one and it has been included in the Release 13 LTE-Advanced Pro

• Results carried out by these operations as for MCS may be surprising. While the improvement rate for MCS is less than for MIoT, on the other hand it can be found that MCS's slope is steeper, which means that it gradually improves more than MIoT but its TR was initially smaller. This gives evidence that it would be better to wait until the deployment of scenario B.

Taken these conclusions into an account, next table will show the same crosses between scenarios and cases of use. However, instead of showing the TR for each of the crosses, a mark will be granted from 1 to 3 for each of the cells. This mark will be a recommendation score (RS), where 1 means the less recommended option and 3 the most recommended one. Table 13 shows these recommendations:

Table 13. Technical Recommendation scores for every case of use-scenario situation

		Scenario A	Scenario B	Scenario C
Recommendation score (RS)	EMBB	2	2	**3**
	MIoT	**3**	3	3
	MCS	1	2	**3**

Some ideas that must be clear to understand this table properly and not being taken into mislead:

• This table represents a conclusion from the calculations and qualitative ideas discussed throughout the entire section. Therefore there is no direct source that shows these results. In the same way, it only represents a personal recommendation on whether it is recommendable or not to work in each of the cases of use at a particular scenario.

• These results are obtained by only assessing technical conditions. Needless to say, it is impossible to reach to conclusive ideas without considering the economic part, which is going to be analysed in next section.

• As stated at the beginning of this section, all cases of use could be best covered with the most complete scenario (C). Therefore, from a technical point of view, it seems logical that every case of use maximizes its RS when reaching scenario C. Hence, the interesting part of this table is to analyse the scenario in which the score becomes maximum. As an example, having a 3 for MIoT in scenario A means that there is not much sense in deploying scenario C for MIoT, as the case of use would be sufficiently covered in the first stages. In EMBB case, a RS of 2 for both scenarios A and B mean that many of the applications of EMBB will be already covered by the time the last scenario is deployed.

4.3 Economical analysis: cases of use

This section will complete section 4.2 by giving the economic perspective to the whole problem. The procedure will be identical, as many of the assumptions (that were used to develop previous section) and conclusions that were taken out of it will be used to develop this analysis.

First, the assumptions that will be used to discuss this section are:

- Same cases of use: EMBB, MIoT and MCS.
- Same relation between IMT figures and scenarios (3, 7 and 4).
- KPIs will still remain constant throughout the scenarios.

The starting point to proceed with this part of analysis will be assumption of cost statement from the European Commission's report "**Identification and quantification of key socio-economic data to support strategic planning for the introduction of 5G in Europe**" named as source [30]. This report estimates the deployment costs of 5G by extrapolating the currently known data from other generations. Figure 43, taken out directly from [30], shows the cost evolution throughout the years:

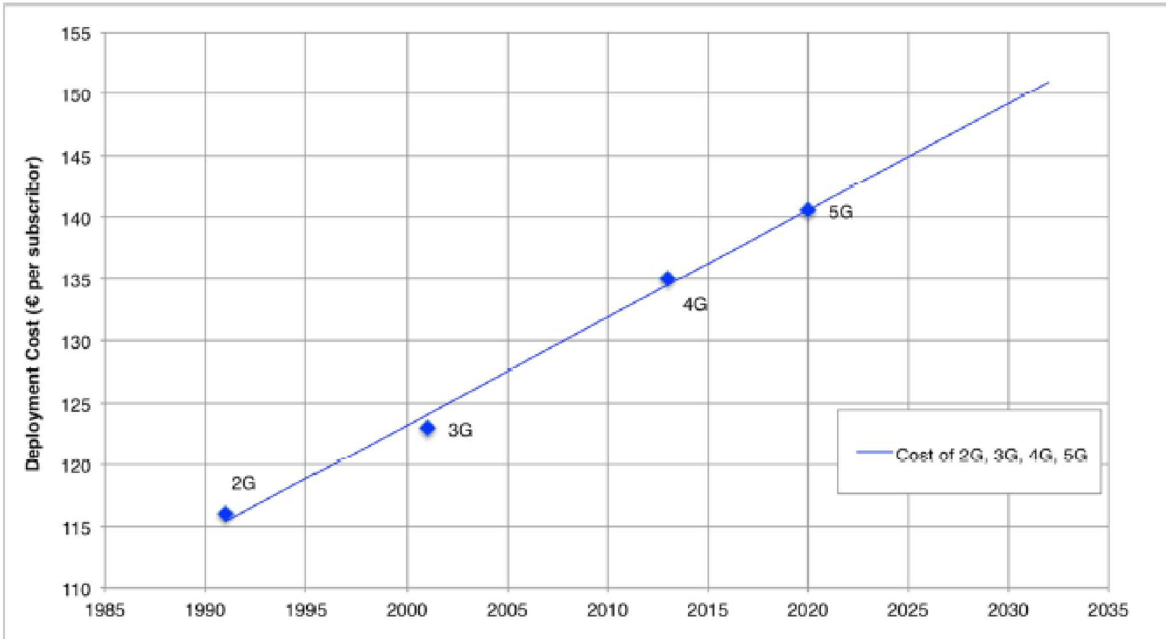


Figure 43. Estimating the cost of 5G deployment. Source [30]

Recalling the binding done in section 4.1.2 that unified scenarios and years, it is possible to associate a deployment cost to each scenario. Figure 43 represents "deployment cost" as it sets the generation milestones and then elaborates a linear regression. However, as it was mentioned in previous section, for some years (scenarios), these costs will not be deployment but enhancing costs.

Therefore, it is important to keep in mind the architecture of the different scenarios. As listed in figures 36 to 38, the disposal of the architectures is:

Table 14. Scenarios' architecture breakdown and cost analysis

	Scenario A-2018	Scenario B-2019	Scenario C-2020
Core	OLD(EPC)	NEW (NGC)	NEW (NGC)
RAN	OLD (LTE)	OLD (LTE) Developing NEW(NR)	NEW(NR) Assisted by OLD (LTE)
Costs (€/subscriber)[43]	138.875	139.75	141 (data from report)
Increase over 4G (*LTE:135€/s*)	1.0287	1.0351	1.0444

There are two options for each of the two parts of the network (enhanced old or new). In this sense, there are two types of costs to analyse:

• **Enhancing costs (EC):** this kind of costs is given when a side of the network is just enhanced. It must be recalled that even though the network's side's name does not change (LTE and EPC), they both will be enhanced, turning them into part of LTE-A Pro systems, in order to make them fulfil the IMT requirements.

• **Deployment costs (DC):** this kind of costs will be given in the transition between an old and a new side of the network (e.g. deploying the NGC to substitute the EPC). As for scenario B, NR is considered to be under development. It will be further justify whether it shall be considered as deployed or not; or if otherwise only a part of the deployment costs may be considered.

As introduced, the method will be similar. Instead of setting the spotlight on the absolute values, the aim will be to analyse the relative ratios between all the crosses (remember that a cross is defined by a case of use and a scenario). As long as each scenario is defined by a new or old RAN and core network, it would be relatively easy to analyse which is the most expensive and cheapest option amongst all of them. In this sense, there are some parameters (variables and constants) to be defined that will give mathematical sense to the problem:

• $\mathbf{EC_{RAN}}$: cost of enhancing the RAN (antennas, etc.) into a LTE-A (Pro) system.

• $\mathbf{DC_{RAN}}$: cost of deploying the new RAN with all its technologies.

• $\mathbf{EC_{CORE}}$ cost of enhancing the core into a LTE-A (Pro) system.

• $\mathbf{DC_{CORE}}$: cost of deploying the new RAN with all its technologies

[43] These costs have been calculated by finding $m = \frac{\Delta y}{\Delta x} = \frac{145-110}{2025-1985} = 0.875$ & $b = 110$ and substituting in a linear equation y=mx+b

(To see all the technologies named in each of the cases, please refer to chapter 3, where all the key issues are addressed and dissected into 4G and 5G). Now, there are four parameters that can be directly defined by crossing the four mentioned above:

• $\boldsymbol{\alpha} = \frac{DC_{RAN}}{EC_{RAN}}$ • $\boldsymbol{\gamma} = \frac{DC_{RAN}}{DC_{CORE}}$

• $\boldsymbol{\beta} = \frac{DC_{CORE}}{EC_{CORE}}$ • $\boldsymbol{\delta} = \frac{EC_{RAN}}{EC_{CORE}}$

Initially α and β should be the ones that would be more interesting to develop in this analysis, because they reflect the relative ratio in the costs that are more useful to compare: enhancement with deployment. However, γ and δ will perhaps be useful to discuss the benefits of enhancing and deploying one or another side of the network. Therefore, now it's time to give values to the costs. These values have been obtained by researching and interviewing experts in 5G across the mobile communications industries[44]. Before giving the figures, there are some premises to be considered:

• To set a relationship between the costs shown in table 9 and the parameters that have been just introduced the deployment assumptions will be taken into account. So any cost may be written in this way:

$$Cost_{Scenario\ i} = Cost_{RAN} + Cost_{CORE} \qquad ec(8)$$

Where any of the Costs must be DC or EC. There is an exception to this premise for scenario B. In this case, as considered the RAN under development, the cost for the RAN will have to contributions: EC_{RAN} & DC_{RAN} both of which will have to be weighted to keep a logical reasoning. In general we can assume that:

$$Cost_{RAN}^{S7} = \varepsilon \times EC_{RAN} + (1-\varepsilon) \times DC_{RAN} \qquad ec(9)$$

By applying this criteria to every scenario, and the particular equation to scenario B, a system of 3 equations appear, which has the following aspect:

$$Scenario\ A \rightarrow EC_{core} + EC_{RAN} = 138.875 \qquad ec(10)$$
$$Scenario\ B \rightarrow DC_{core} + (\varepsilon \times EC_{RAN} + (1-\varepsilon) \times DC_{RAN}) = 139.75 \qquad ec(11)$$
$$Scenario\ C \rightarrow DC_{core} + DC_{RAN} = 141 \qquad ec(12)$$

At this point, it is totally necessary to make an assumption. By consulting experts on mobile communications (Nokia, Ericsson) it was revealed that, due to the new techniques such as SDN and NFV (see chapter 3) on which the Cloud-Core will depend, the NGCN's cost will be very similar to the enhanced EPC's one. Hence, the costs for the new core will be mainly due to software engineering and programming, creation of cloud data centres and virtualization tasks.

[44] Experts asked worked for Nokia and Ericsson. They also revealed information that will be revealed throughout the analysis. Where quoted, note they are these experts mentioned here.

Therefore, the assumption can be easily explained by stating that $EC_{core} = \mu \times DC_{core}$. Here, μ represents a factor very slightly smaller than 1, which represents that deploying a new core will always be more expensive than enhancing the old one, even if their cost is similar. With these assumptions, the new system is:

$$Scenario\ A \quad \rightarrow \quad \mu \times DC_{core} + EC_{RAN} = 138.875 \qquad ec(13)$$

$$Scenario\ B \quad \rightarrow \quad DC_{core} + (\varepsilon \times EC_{RAN} + (1-\varepsilon) \times DC_{RAN}) = 139.75 \qquad ec(14)$$

$$Scenario\ C \quad \rightarrow \quad DC_{core} + DC_{RAN} = 141 \qquad ec(15)$$

- Parameter ε represents the costs related to the RAN in Scenario 7. The main difference with μ is that this relation gives the opportunity of deciding which the best option when spending money is. This means that, once known the constraint (more money spent in enhancing the RAN), it can be decided (**optimized**) how much money can be spent on enhancing and deploying.

4.3.1 The optimization problem

From now on, the economical analysis will be considered as an optimization problem. It will be solved by using the software Matlab. The reason why it is going to be faced this way is the fact that, when dealing with costs, the most common objective is cost saving. It must be noticed that this analysis does not take into account the effect on technical performance yet (i.e. how would reducing money spending affect the system's technical requirements and KPIs).

a. Input parameters

To adjust to an optimization problem's form, the conditions stated throughout this section must be transformed into a set of variables, an objective function, linear and nonlinear constrains (both equalities and inequalities), a starting point and lower and upper bounds for each variable. All these will constitute the input parameters for the optimization problem. In this case:

- **Variables**: there will be 6 variables in this problem that have been already introduced, which are EC_{Core}, EC_{RAN}, DC_{core}, DC_{RAN}, ε and μ. Note that EC_{Core} has appeared again because all of the six will be optimized.

- **Lower and upper bounds**: these are the limits that variables must stand within.

  - Costs must be greater than 0.

  - 0.5 < ε < 1 & 0< μ < 1 (already explained above)

- **Objective function:** the objective function to minimize in the problem is:

$$f(EC_{Core}, EC_{RAN}, DC_{core}, DC_{RAN}) = EC_{Core} + EC_{RAN} + DC_{core} + DC_{RAN} \quad ec(16)$$

Note that, from an economic point of view, the ultimate objective is to minimize all the costs. Therefore, the summation of all the costs must be minimal, no matter how early does the deployment of the new network starts.

• **Constrains:** they are conformed by all the equations from the scenarios plus extra logical assumptions. These are:

$$Scenario\ 3 \quad \rightarrow \quad EC_{core} + EC_{RAN} = 138.875 \qquad ec(17)$$

$$Scenario\ 7 \quad \rightarrow \quad DC_{core} + (\varepsilon \times EC_{RAN} + (1-\varepsilon) \times DC_{RAN}) = 139.75 \quad ec(18)$$

$$Scenario\ 4 \quad \rightarrow \quad DC_{core} + DC_{RAN} = 141 \qquad ec(19)$$

$$\mu\ definition \rightarrow EC_{core} = \mu \times DC_{core} \qquad ec(20)$$

$$Enhancement\ cheaper\ than\ deployment \rightarrow EC_{RAN} < DC_{RAN}\ \&\ EC_{core} < DC_{core} \quad ec(20)$$

• **Starting point (x_0):** only enhancing costs start at its levels for 4G (see figure 43) which will be assumed to be each one of them the half of the total cost (67.5 €/subscriber). The rest of the variables start at 0.

b. Optimizing

Next step is to actually solve the problem by optimizing the variables. To do so, there are two options that must be chosen to proceed: the function and algorithm to use.

• **Function**: choosing the function means what is the objective of the problem. In other words, it defines what it is going to be done with the objective function defined in section a. Two different functions will be tested in order to compare its results[45]:

- **fmincon**: this function finds the minimum of constrained nonlinear multivariable function. In this case it finds the minimum $f(EC_{Core}, EC_{RAN}, DC_{core}, DC_{RAN})$

- **fminimax:** : this function solves the minimax problem. This is, it tries to minimize the maximum cost. Note that it does not have to find the minimum cost like fmincon will do. Instead, it will keep the solution that that guarantees the smallest of the maximum $EC_{Core} + EC_{RAN} + DC_{core} + DC_{RAN}$ possibilities.

• **Algorithm:** the algorithm gives an answer to "how do I achieve my objective". In other words, it provides the mathematical procedure to find the answer. These are usually complex algorithms, so the algorithm chosen will be "interior point" (which is pre-set by default), and it will be kept as a black box.

[45] Functions are named in the document as they are named in Matlab syntax. In Appendix 1 the software script is fully explained. For more info, please visit https://es.mathworks.com/help/optim/constrained-optimization.html

c. Results

The function returns two measures. The first one is the actual value that returns f, the summation, while the second one is the value for each variable. An extra column was added (**fminimax**$_2$) changing the initial values for $DC_{core\&ran}$ to 40 as for a more conservative scenario. Next table and figure both show the results for each of the two functions:

Table 45. Variables comparison amongst functions

	fmincon	fminimax	fminimax$_2$
EC_{Core}	70.22	67.96	40.47
EC_{RAN}	68.66	70.91	98.40
DC_{core}	70.88	68.42	40.92
DC_{RAN}	70.12	72.58	100.08
ε	0.85	0.75	0.74
μ	0.99	0.99	0.99
f	279.875	279.875	279.875

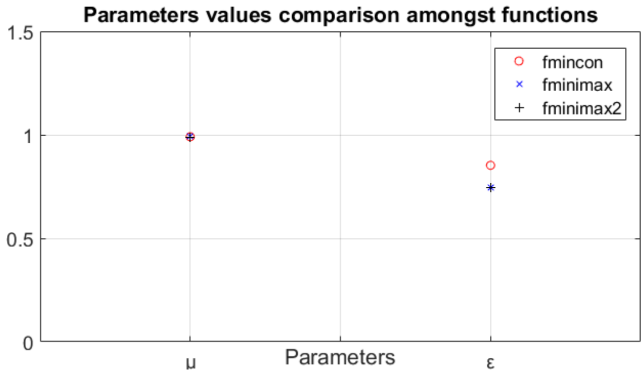


Figure 44. Parameters values comparison amongst functions

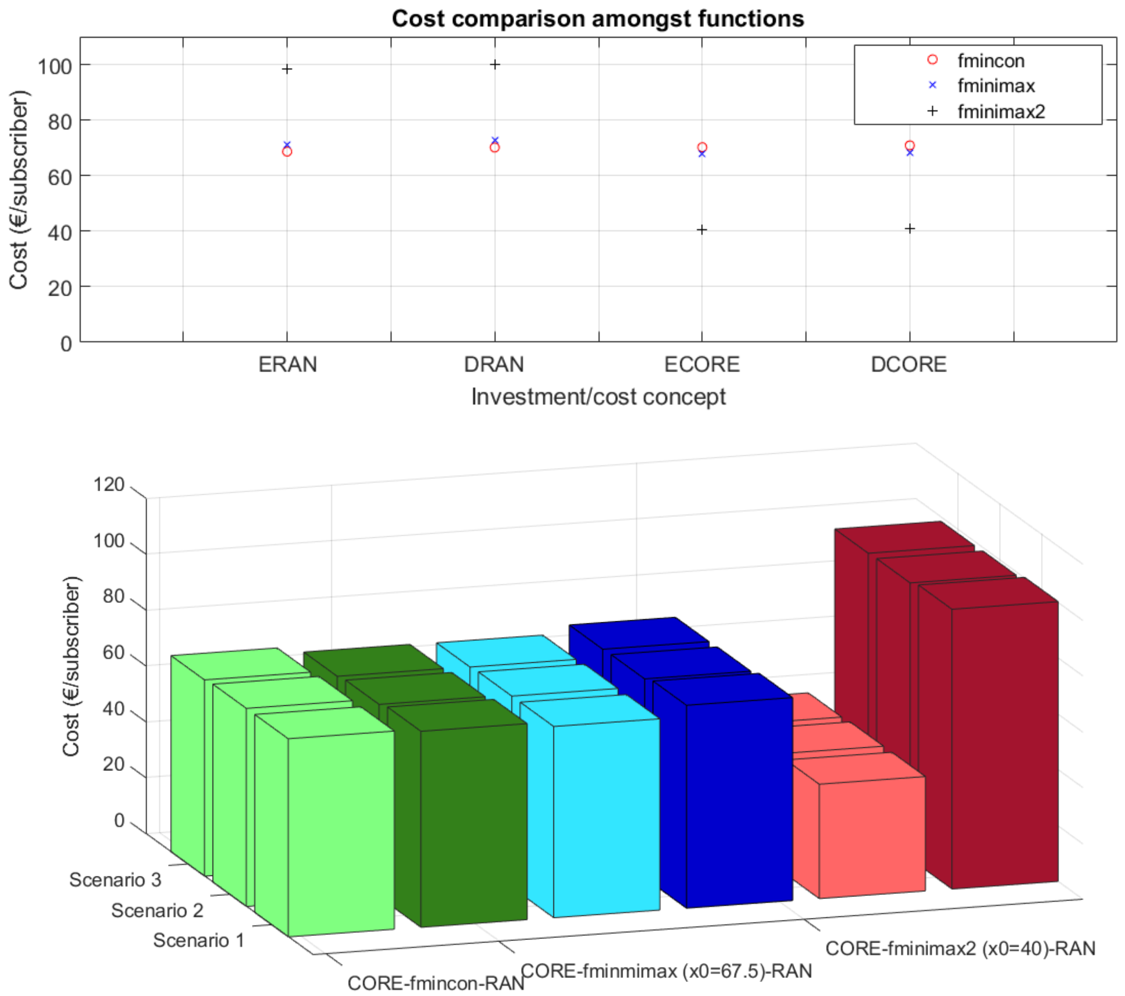


Figure 45. Cost comparison amongst functions (in 2D and 3D)

Figure 45 shows in two and three dimensions the comparison amongst the functions. As for the 3D graph, each function is represented by a colour (green, blue, red). Within a colour, the lighter bar represents the core while the darker bar represents the RAN. Furthermore, the Z axis represents the three scenarios for each function. It may be seen that cost increase (per subscriber) amongst scenarios is very small .And lastly, table 16 shows the four remaining parameters (for each of the three cases) that will give extra information: α, β, γ and δ:

Table 16. Parameters α, β, γ and δ for each of the functions and their mean

	fmincon	fminimax	$fminimax_2$	mean
$\alpha = \frac{DC_{RAN}}{EC_{RAN}}$	1.0212	1.0235	1.0243	1.0230
$\beta = \frac{DC_{CORE}}{EC_{CORE}}$	1.0093	1.0067	1.0111	1.0090
$\gamma = \frac{DC_{RAN}}{DC_{CORE}}$	0.9661	1.0608	2.4457	1.4909
$\delta = \frac{EC_{RAN}}{EC_{CORE}}$	0.9777	1.0434	2.4314	1.4842

d. Conclusions

There are several conclusions that can be taken out from these two figures and the corresponding table. The ones that are essential to continue with the analysis are going to be explained at this point (the rest, which include the ones that include techno-economic comparison are left to chapter 5):

• The optimal value that the two functions reach is the same. This means that the problem converges and has a finite solution, which obviously is the sum of the costs of Scenarios A & C (the ones that were respectively the sum of the enhancing and deployment costs). Nevertheless, this number does not give any information to this problem, as it was already known that the problem would converge.

• Parameter μ remains constant and almost equals 1. This confirms the assumption provided by experts on the similar price between enhancing the EPC and deploying the NGCN. This parameter could not equal 1 because the problem would not converge, but it could be approximated to 1.

• Parameter ε varies but is always above 0.75. In the highest case, it equals to 0.85, what would mean that the investment costs on the RAN would optimally incur in an 85% in enhancing it, and only a 15% in deploying the NR (for scenario 7). This also supports the previous assumption that stated that the NR could be considered non-deployed.

• As for fmincon function, the starting point x_0 does not affect the results because the function keeps searching for the lower cost. However, two trials have been developed with fminimax function. It turned out that, while ε and µ initial values did not affect the result, the initial values for the enhancing costs affected it. Particularly, the initial values (which were first considered 67.5 €/subscriber and then 40€/subscriber) determined the value of the core costs (both enhancing and deploying) which were very similar to the x_0 values.

This fits with the qualitative idea that enhancing the core will be easier due to the kinds of tasks to do it, while the RAN side will be more expensive because of new technologies, frequency bands, etc. Therefore, if the initial cost is higher (more money spent earlier in time), it would mean that it would be mainly used to improve the core rather than the RAN.

• For the two first cases, which seem to be the most feasible ones, it can be stated that each side of the network represents (economically) half the network indeed. Moreover, enhancing and deploying each of the segments would have a very similar cost (the absolute difference will depend on the number of subscribers). This also fits with the idea discussed throughout the entire document: "5G = LTE-A Pro+ NR/NGC"

• Relations α, β, γ and δ are very illustrative and reaffirm everything that has been concluded until now: deploying either of the segments of the network is slightly more expensive than enhancing it. On the other hand, the higher the initial investment is, the more expensive the RAN enhancing/deploying is, and the cheaper the Core results. This was explained due to the idea and assumption that the first investment will be headed to improve the core (not in vain did scenario B contemplate the new core deployed).

Furthermore, the higher the initial costs are, the higher parameter ε is. The higher this parameter is, the higher investment in the RAN is headed towards enhancements and the less is headed towards deploying in scenario B. As a consequence, the deployment cost for the RAN decrease in favour of the increase of the Core costs.

• Lastly, it must be at least mentioned that this economical analysis has only analysed the costs in which a service provider would incur. On no account does this analysis take into account the benefits that could be driven from one or other option. Moreover the costs are considered per subscriber. The increase on subscribers would increase the total cost and therefore results for forward analyses.

4.4 Techno-economic analysis: TR vs. costs

It is now time to contrast the technical and economical results. The aim in this section is to compare the technical performance in every scenario to the cost of each of them. As for this section, no scenarios will be prematurely discarded, so all the 9 crosses will be discussed. Conclusions will be left for next chapter.

Firstly, the most obvious graph will be presented: the total costs for the deployment of each scenario (without being disaggregated between RAN and Core) vs. every cases of use technical rate. In this case, X label will conformed by be the technical rates, while Y label will be labelled as the costs. This means that the aforementioned scenarios will be represented by asymptotes in the X label, since scenarios will have a fixed cost (i.e. fixed value in the X axis) regardless of the case of use.

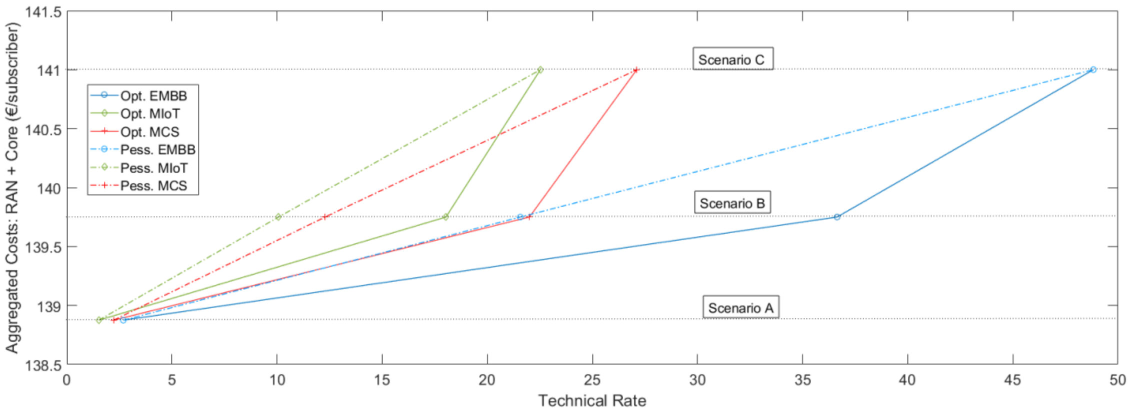


Figure 46. Total aggregated costs vs. TR for each case of use

The way to understand figure 46 is to notice that the objective is to have the less steep slope. This would mean that the technical rate would improve with less increase in the total costs. In this sense, all the cases highest improvement would happen between scenarios B and C. The eMBB case of use is the one that would improve more, both in overall and between the mentioned scenarios.

Next step is to disaggregate the costs so it can be analysed how the different core and RAN networks (and its price) affect the technical rate. At this point, there is other factor that must be considered: the different functions used in section 4.3.1. It must be recalled that one or another considered starting points affected a lot each sides' costs. However, although the prices for each of the functions aforementioned are different, the increase between prices (α and β from previous section) remain very similar (see table 16). Furthermore, parameters γ and δ (that change within the functions) vary because of x_0 so they cannot be properly analysed given lack of data.

Nonetheless, there is a shocking fact. The economic analysis gave evidence that there were different combinations to reach an aggregated cost that figure 46 relates to a technical rate. If x_0 had been increased more and more, a point where RAN-related costs reached to 0 would have been reached. This would have meant that all the costs and investments would have been incurred because of the core. From an economical point of view this doesn't mean anything as it was the function optimization at that point. However, from a technical point of view, it's a logical fact that improving and investing money in only one side of the network would be pointless, since performance relays on both segments of the network that must interwork. Henceforth, only the two first functions (fmincon and fminimax) will be taken into account, even being $fminimax_2$ a reasonable prediction.

Lastly it is also very important to discuss why the line remains with an almost constant growth for the pessimistic scenario. This is explained because the weight for the side that improves most (the core) becomes smaller, so the improvement effect is minimized for the first transition. As a consequence, in the second transition, where NR is deployed, the increase becomes greater (compared to optimistic scenario) so that the final TR is reached in both cases. If more weight was given to the Core, the TR at Scenario B would increase.

This completes the whole analysis. In next chapter, the conclusions will be developed and contrasted with the demand forecast so they can be focused properly.

5. <u>Result analysis</u>

Chapter 4 results were left as technical improvement percentages or cost increase ratio (also in terms of percentages). The aim in this chapter is to contrast these improvement ratios with demand forecasts. This will enable the opportunity of knowing not only how much better will networks be, but also to know how many people (subscribers) will be affected by these changes in mobile communications.

Chapter 5 is going to be divided into two well differentiated sections, each of which will deal with two of the aspects that must be issued in this document. The first section will analyse the social and economic environment. The aim of this section is to discuss the economic impact using the methodology aforementioned. On the other hand, the second section will address the social and economic impact of the project; this is the aftermath of the 5G networks.

5.1 Demand forecast: social and economic environment

The first approach to develop this section will be the discussion on the European situation. As for this environment, the amount of subscribers increase will be studied alongside the traffic increase. Secondly, the cases of use requirements will be discussed. These two development lines will give two perspectives, for scenarios and cases of use respectively. Thus, this social and economic data will complete the techno-economic analysis and the less feasible options will be discarded.

As for the Spanish situation, and extrapolation of the analysis and figures will be done. Once known Spain's role within the EU, it may be represented by a percentage. Specific conditions will be taken into account, and the socio-economic impact within the country will be assessed.

5.1.1 Total results: TR vs. Total Costs

As introduced, the first step is to get the real absolute data. The economic results in chapter 4 were given in euros per subscriber (€/subscriber). Therefore, it must be know the amount of subscribers for each year. Ericsson's Mobility Report was consulted (Source [31]) as it revealed data in terms of subscriptions/year. The difference between the number of subscriptions and the number of subscribers is due to inactive subscriptions, multiple device ownership and/or optimization of subscriptions for different types of calls [31]. Nonetheless, this difference will be neglected to facilitate the work. In next figure and its corresponding table, the numbers (not only subscriptions but also traffic amounts) are displayed. Furthermore, figure 47 displays the system to which subscribers are connected, so the number of LTE-5G connections through time could be calculated.

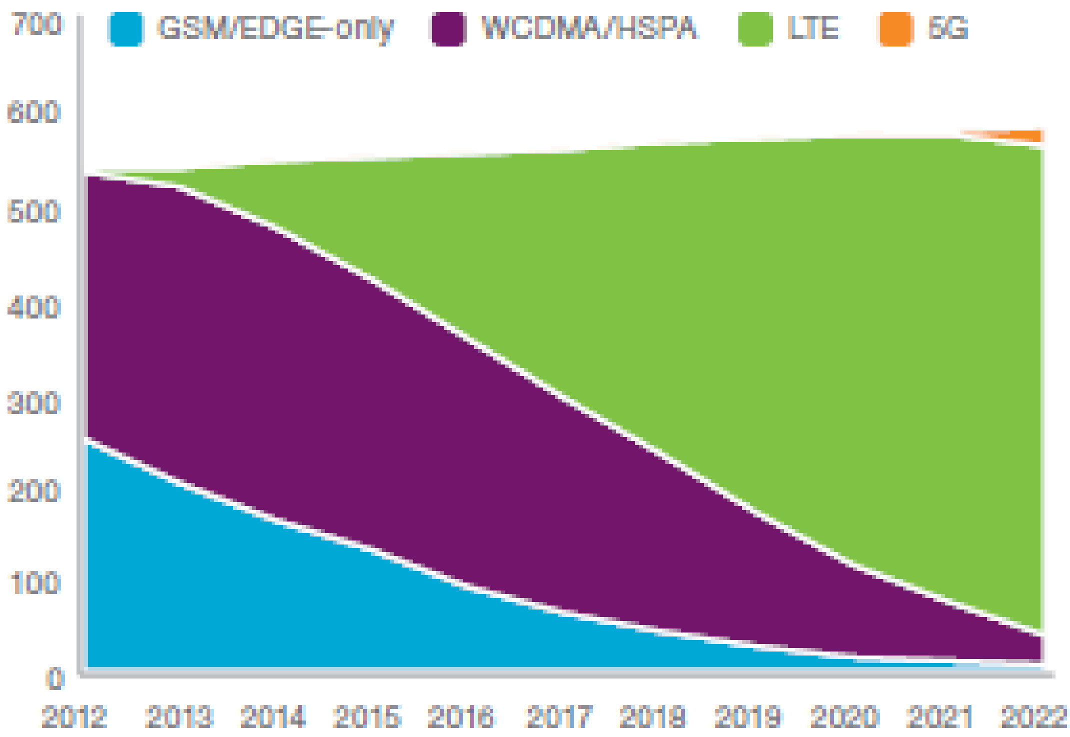


Figure 47. Western Europe mobile subscriptions. Source [31]

As it can be seen in figure 47, the trend will be the decrease of GSM and HSPA subscribers in favour of LTE subscribers. Moreover, it calls the attention the fact that 5G subscribers aren't forecasted until years 2021-2022. This prediction is perfectly feasible: IMT requirements are expected to be fulfilled by year 2020. On the other hand, Release 15 is not expected until 2018, and even later (by 2020) will Release 16 be terminated. Hence, until late 2020 no 5G system will be fully standardised, developed and deployed, and so won't 5G be commercialized until 2021 at least.

However, following the premise of 5G comprising LTE enhancement (and therefore its subscriptions), **LTE subscribers will be considered as well** for the analysis. To know the exact numbers, some operations will have to be done with the figures on table 17, which is displayed right below:

Table 17. Western Europe Subscription & Traffic figures. Source: [31]

	2016	2022	CAGR 2016–2022
Mobile subscriptions (million)	550	580	1%
Smartphone subscriptions (million)	380	480	5%
Data traffic per active smartphone (GB/month)	2.7	22	40%
Total mobile traffic (EB/month)	1.2	10	40%

First conclusion to be made upon table 12 is the high increase in mobile traffic, either in total or by user, compared to the subscriptions increase. Then, the first idea to get is that there will be slightly more devices generating 8.15 times nowadays' traffic.

Table 17 shows the CAGR (Compound Annual Growth Rate), which is defined as:

$$CAGR = \left(\frac{Ending\ Value}{Initial\ Value}\right)^{\frac{1}{years}} - 1 \qquad ec(21)$$

In this case, all the parameters are known for 2020, so the aim is to find the Ending Value for years 2018, 2019 and 2020. With simple calculations these values are the ones shown in table 18[46]. Moreover, the extra row calculates the approximate number of LTE-5G connections. Now, the different deployment costs for each of the scenarios can be contrasted by multiplying the number of LTE subscribers x cost per subscriber for each scenario. All this is shown in next table:

Table 18. Calculated subscriptions and deployment costs through the years. Reconstructed from table 17

	2016	2017	2018 (SA)	2019 (SB)	2020 (SC)
Mobile Suscriptions (million)	554,89 **555**	559,8235 **560**	564,8008 **565**	569,8224 **570**	574,8887 **575**
LTE-5G Subscriptions (million)	**185**	**260**	**325**	**420**	**455**
Costs (€/subscriber)[47]	**-**	-	138.875	139.75	141 (data from report)
Total cost (billions)	**-**	**-**	**45.134**	**58.695**	**64.155**
Percentage of LTE-5G	33 %	46.43 %	57.52 %	73.68 %	79.13 %

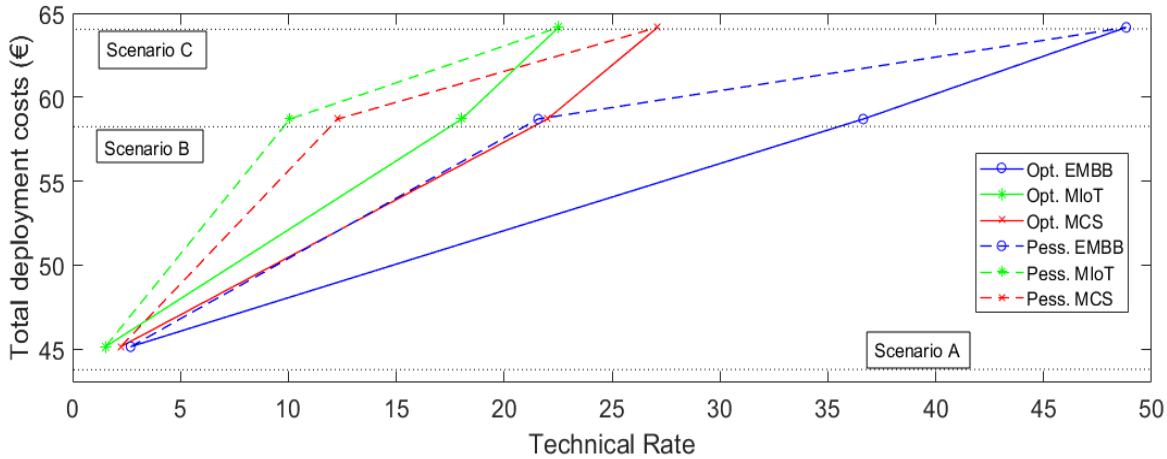


Figure 48. Total aggregated costs (billions €) vs. TR for each case of use

Figure 48 represents the total deployments costs versus the same TR from chapter 4.

[46] In the table, it is first displayed the exact number. Below, the approximation is shown, which is always rounded up to consider until the last of the subscribers.

[47] These costs have been calculated by finding $m = \frac{\Delta y}{\Delta x} = \frac{145-110}{2025-1985} = 0.875$ & $b = 110$ and substituting in a linear equation y=mx+b

Important conclusions to be made upon table 18 and figure 48:

1. *Meaning of figure 48:*

   Figure 48 is very similar to figure 46, as they both confront TR with costs. However, in figure 48 TR is confronted to total deployment costs. In other words, total subscribers are incorporated to the equation. This means that any difference in lines' aspect is due to the effect of the subscribers' difference.

   As in figure 46, the best options are identified by the less steep slopes, what would mean that the TR increases more than the costs do. Comparing both figures,

   • Lines are not only modified by changing the cost-intercept point but also by changing the shape. For the optimistic cases, the two-phase concave lines turn to the approximately straight lines. For the pessimistic cases, the opposite happens, but with a convex two-phase line instead for figure 48.

   • It can be therefore generalized that, regardless of the case, the growing amount of subscribers introduces a cost peak. This peak is noticed in scenario B, since Scenario A and C are the initial and ending points. Although subscribers grow for each year-scenario, its highest growth (according to source [31]) happens between scenarios A and B (13.461 billions) which is almost three times than the growth from scenario B to C (5.4 billions). This increments the total costs, what makes the slope become steeper.

2. *Pessimistic scenarios:*

   Pessimistic cases are represented in dashed lines for both graphs. Remember they were called so because it is a pessimistic hypothesis to deploy a transition scenario B such that more money is invested in the core but KPIs depend more on the RAN. Hence, money would not be optimally invested. For pessimistic scenarios, lines increased linearly through scenarios A, B and C. This was due to the core weight in the TR: it was so small (35%) that the TR would be slightly influenced by it. In this case, the TR increase is due to the RAN improvement.

   In this sense, lines become steeper for the first transition between scenarios A and B. Hence, all the TR improve most in the second transition, what makes more desirable to deploy the three cases of use as late in time as possible. Even so, eMBB remains as the more desirable one to deploy in two stages. Furthermore, in the second transition, there is a better commitment for MCS rather than for MIoT, what would make MIoT the best option to deploy first.

3. *Optimistic scenarios:*

Optimistic cases are represented in solid lines. In chapter 4 it was concluded, accordingly to figures 41 and 45, that eMBB improves most for both transitions, followed by MCS and lastly MIoT. Assuming that the explanation given in the previous page for the effect of the subscribers' increase is valid as well, the effect of the subscribers' increase has the following particular consequences for the different cases of use:

• eMBB: the line becomes almost straight, what means that costs and TR grow accordingly. This hypothesis represents perfectly the vision of a wide range of applications for eMBB. A improve in the technical KPIs would affect linearly in the costs, which is the most coherent approximation.

• MCS and MIoT: both lines follow the same trend. The convex lines turn into concave ones, fact that suggests that the cost peak in the transition between 4G and 5G will be reduced significantly. This also results obvious since all the transitions between generations have resulted in a cost peak for those service providers that have first implemented the new technologies. However the difference between cost and TR in the second transition is quite smaller than for the first one.

Between these two cases of use, is MCS the one to have a steeper slope in the second transition. Hence, when deciding between these two, it would be more recommendable to develop MIoT application at first, and leave MCS for the second development stage.

4. *Pessimistic versus Optimistic:*

It was already introduced the difference between the optimistic and pessimistic assumptions. The optimistic assumption would mean that, since the RAN has more weight in the TR, money would be optimally spent when choosing Scenario B as the transition. On the other hand, for the pessimistic assumption, the opposite would happen.

This is represented in figure 48, where the colour-corresponding dashed and solid lines have the same beginning and ending point, but different transitions. The more weight the RAN has, the less improvement for the first transitions and the more improvement for the second (where the NR is deployed). That is the reason why the lines for a same case of use draw a rhombus, except for eMBB, where the cost increase is such that it makes the solid line straight.

Besides, it must be mentioned that the first transition improvement for the optimistic assumption is different from the second transition pessimistic (and vice versa). This means that there is no direct relation between improvement phases within the two assumptions. Anyway, the ending point is the same. Hence, the total improvement remains constant, but the difference of improvement within transitions is more notorious for the pessimistic case. This supports the idea of wasting money in this assumption for the first transition.

This difference generates a discussion topic: what is the best transition scenario between 4G and 5G? Scenario B was chosen according to source [27]; however there could be other better scenarios, such scenario 8 (see section 4.1.1) if technical performance of the net depends more on the RAN. Chapter 6 will deeply discussed this fact.

5. *Differences between functions:*

Differences amongst economic optimization functions have been neglected in chapter 4, because these did not affect the development since the total cost would be the same. Furthermore, for the two first functions (fmincon and fminimax) costs were quite similar except for parameter ε. Furthermore, parameters α and β (see section 4.3 for definitions), give evidence that for the three functions, the difference between deploying and enhancing both segments of the net is almost null.

Hence the only differences between functions are parameters γ, δ and ε, which show the difference between segments (in terms of enhancement and deployment costs).

• For the two first functions, (slightly) more money in spent in the core ($EC_{CORE} > EC_{RAN}$ & $DC_{CORE} > DC_{RAN}$). For the optimistic assumption, this is the best option since both segments are considered to have the same weight for the TR. Nonetheless, for the pessimistic case, there would be more almost equal money spent in both segments, but being the RAN the segment that influences most. Therefore, it would be a very questionable money investment strategy.

• For function $fminimax_2$, the RAN represents the around the 70% of the costs. It is important to consider that this function differed from fminimax in the starting point, which is lower in this case. Therefore, shall the starting point be lower; the maximum cost would be minimized by investing in the RAN. This supports the evidence that the RAN has more weight in the network, since the RAN-related costs increase to minimize the maximum cost

5.1.2 Crosses selection & analysis

The aim of this section is to use all the conclusions obtained in previous section to reason which would be the best crosses[48] by analysing the commitment between TR and total costs. There will be four situations, since there are two functions[49] and two assumptions (optimistic and pessimistic). After analysing the four options, it will be discussed what are the ones that must be discarded.

To address this issue, it is going to be first given for each cross a recommendation score as in section 4.2, but taking into account the economic analysis asa well, summed up in figure 48. Therefore, for every case of use it is going to be determined the feasibility of each scenario what will correspond to a recommendation score.

a. eMBB:

This is the toughest case of use to discuss as for the wide range of applications. Figure 48 showed that the TR-cost commitment remained linear throughout time or it experimented a cost peak for the pessimistic assumption. For the optimistic case, eMBB would be equally desirable to spent money in. On the other hand, for the other assumption, it showed a better commitment in the second transition. It must also be analyzed how discriminative eMBB is with KPIs sub-optimal values.

First, due to new techniques such as LTE-WLAN Aggregation, traffic may be coursed directly through Wi-Fi 5 GHz acess point. This would mean that the net would have more bandwidth to offer. However, EPC would have to be sufficently enhanced to course the higher amounts of traffic. For other cases of use that imply outdoor environments were LWA can not be applied, both RAN and Core goals become more difficult to achieve. However, these outdoor applications for eMBB are the natural evolution of multimedia traffic, which would be gradually achievable and is not a very demanding case of use that might start working with worse performance. Therefore:

• For scenarios A and B, this case of use would be more or less covered. The less demanding applications, such as HD video (worse than ultra HD, 4K, etc.), or data (including voice) traffic would be covered. Also, more demanding applications like industry automation and working and playing in the cloud would be sub-optimally covered, but they could give a sufficiently good performance. Hence, a RS of 2 will be given for both scenarios.

[48] Remember crosses were defined as a scenario together with a case of use.
[49] Fminimax and fmincon will be considered as one function only due to its similarity in costs.

• Scenario C represents both the best technical performance and the best TR-deployment costs commitment. This shows that Scenario C would be optimal for both the most and less demanding applications. Undoubtedly, a RS of 3 must be given for this cross.

b. *MIoT:*

This case of use improves the least through the three scenarios. The most influential KPI was the connection density (remember this was measured in devices/km^2) which was intimately related to the cells' size. However, MIoT is the less demanding case of use, since it enerates very low traffic volume. In addition, future cells will be dessigned to meet eMBB and MCS needs, so they will by far meet MIoT needs.

It was discussed that NB-IoT, which is already implemented, has proved to be a valid solution for this case of use. Assuming that the amount of traffic coursed per device will be similar (edge throughput was weighted 1 in table 9), the only problem would be the increase in the amount of devices.

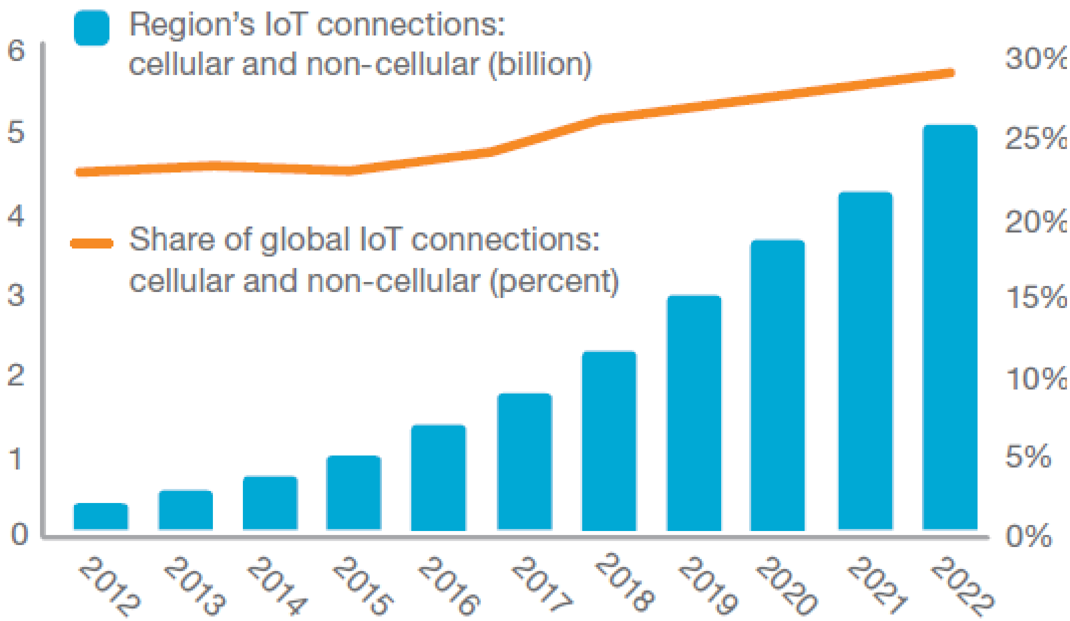


Figure 49. Total IoT connections. Source [31]

Figure 49 shows that, despite the IoT connections growth, IoT connections will still represent a 30% of the total connections. This data give evidence that the connections number won't be a problem either: they will still represent a small percentage of non demanding nor discriminative connections. Apparently, there will not be specific issues to adress (since connection density is not a excluive issue for IoT) in order to achieve the needs for this case of use.

Adding all this to what was concluded in section 5.2.1, MIoT is the most feasible case of use to be deployed within the two first scenarios. This would mean that by scenario C, this case of use would be almost optimally implemented and developed so (almost) no money should be spent in developing MIoT (less steady slope in figure 48).

In conclusion, RS will be 3, 2 and 1 respectively for scenarios A, B and C.

c. *MCS:*

This case of use follows a similar trend than MIoT in the TR-Cost commitment but increasing more its TR. In this sense, similar conclusiones could be driven than from MIoT case of use: for the pessimistic assumption, the second transition represents a better commitment. However, as concluded in previous section, MIoT slope is steeper in the second transition, what makes more desirable to develop (i.e. invest money in) MCS in the second transition rather than MIoT for the optimistic assumption. On the other hand, for the optimistic assumption, there is no doubt that the first transition represents a better commitment in this sense.

There is another important fact to consider. The same experts from Nokia and Ericsson that were consulted in chapter 4 highlighted the rigorousness on the MCS figures: MCS are intended to cover the most demanding applications. These applications have no room for fail or delay, since an error would have fatal consequences, for example, in autonomous vehicles.

The essential KPIs (weighted with a 3 in chapter 4) for MCS are latency and mobility. As seen in table 9, mobility increases 5 times from scenarios A-C and 1.4825 times between B-C. On the other hand, latency increases 100 times (it is 100 times lower) between A-C and 10 times between B-C. Obvoisuly it s not a trivial difference, specially in terms of mobility.

Taking into account both the assumption driven out from the analysis and the information given by the experts, it can be assumed that scenario A is totally discarded for MCS[50]. Scenarios B and C could be both feasible for this case of use, especially if mobility would be less important than latency. Nevertheless, applications like self-driven cars, drones and aircrafts give undeniable evidence that mobility is absolutely important. Hence, scenario B will only be granted a score of 1 while scenario C will have a 3 as it is the optimal option.

[50] Improvement between scenarios A and B is greater than the one between B and C. This starts to give evidence that, if scenario B was more or less feasible, scenario A would in any case be a valid option.

Last step is to display the same matrix as in table 13 but assuming all the conclusions explained so far. In this case, score 0 will be valid as well since there are some crosses that have been proved totally inefficient.

Table 19. Total Recommendation scores for every case of use-scenario situation

		Scenario A	Scenario B	Scenario C
Recommendation score (RS)	EMBB	2	2	**3**
	MIoT	**3**	1	0
	MCS	0	1	**3**

Table 19 summarizes perfectly the whole result analysis. The best combination to optimally address the 5G challenges is to start developing MIoT fully primarily, and start investing in eMBB applications. The second scenario is best oriented towards eMBB. This scenario represents a transitory period in which operators would incur in more costs without developing the full potential of any of the cases of use (and therefore not obtaining optimal revenues). However, it will enable the third scenario C, where eMBB will be fully deployed and MCS will be developed as well.

There is an advantage in this timeline: as eMBB has been being developed throughout the scenarios, less effort will be made upon eMBB so the spotlight can be turned to MCS. Furthermore, invests and technical efforts are divided so that there is a coherence between the developing time, technical performance and deployment costs for each case of use.

Another important conclusion must be made upon recommendation scores. As it was explained, it represents how desirable it would be to invest money in a particular case of use. Although it only represents the actual recommendation, it may be interesting to add the RS in every scenario and see the total (5 for scenario A, 4 for scenario B and 6 for scenario C).

In some sense, this would represent that scenario B is the one in which less developments and breakthroughs are triggered, what corresponds accurately with the idea of a transition scenario. On the other hand, the total RS for scenario A represents a high initial cost, while the 6 represents the highest cost corresponding to the development of the most demanding case of use[51].

[51] This paragraph just gives an estimated idea that in any case is quantitatively proven nor derived from the analysis from chapters 4 and 5.

5.1.3 Result analysis for the Spanish situation

Last step of this section is to extrapolate this analysis that has been made through this chapter (referred to Europe, according to GSMA and Ericsson sources) to the Spanish situation.

There are some parts of the analysis that remain constant:

- Technical Rates don't depend on the country analysed. They are just weights and improvements defined by ITU and 3GPPP, so they remain equal for both situations.

- Deployment costs from section 4.3 remain constant. These costs are referred to ones incurred in by deploying different systems. As long as Spain has the same systems than its European neighbours, the costs will be the same. Furthermore, the relative ratio between costs will be equal as well.

- TR-Cost commitment must remain equal since both sides remain equal as well.

Therefore, it seems obvious that changes arise when introducing the demand forecasts. There are two aspects to be considered in this sense:

- If Spain follows European trend line, results will be similar in shape: it will have lines similar to the ones in figure 48 but changing the cost-intercept point. Otherwise, lines will have different shape. In other words, as long as the trends are equal, relative results will be the same.

  By consulting CNMC (Comisión Nacional del Mercado de la Competencia) website, data are extrapolated [32] in which it can be seen that the trend until year 2015 (last data available) is very similar to the European one. Assuming these facts, **the relative results will remain equal**.

- Spanish subscribers represent just a part of the total European. Therefore, Spanish subscribers will represent a small percentage of the total European. Hence, once proven that the trend is the same, it must be found what the percentage that Spanish subscribers represent is in the whole European data.

  In this case, the source used will be the report from the European Commission named as source [30]. This report showed a table in which the economic investment of each country was shown. Although data slightly differ from Ericsson source, it is valid to extract the desired percentage Spain represents.

Table 20. Input-output effect. Taken directly from [30]

	5G Investment 2020 (€m)	Type I direct input-output effect (€m)	Type I direct input-output employment effect
Austria	970	2,170	25,200
Belgium	1,230	3,150	36,300
Bulgaria	840	2,320	128,900
Croatia	480	1,540	64,400
Cyprus	100	470	20,800
Czech Rep.	1,200	3,990	143,000
Denmark	620	1,480	14,800
Estonia	150	560	13,600
Finland	600	1,501	19,900
France	7,030	17,110	224,700
Germany	9,280	20,740	211,100
Greece	1,220	2,180	101,300
Hungary	1,130	3,450	134,600
Ireland	490	1,210	10,700
Italy	6,830	15,700	186,830
Latvia	230	570	16,800
Lithuania	330	700	28,200
Luxembourg	60	122	600
Malta	50	190	3,900
Netherlands	1,870	5,030	68,300
Poland	4,350	13,040	569,553
Portugal	1,170	3,730	127,300
Romania	2,270	4,660	252,300
Slovakia	620	1,980	71,500
Slovenia	240	610	14,700
Spain	5,190	14,600	329,400
Sweden	1,060	2,450	25,300
UK	7,040	16,520	172,100
EU28	56,640	141,840	2,394,800

Source: Calculations based on data from Stat. OECD, The World Bank's World Development Indicators (WDI),Eurostat [nama_10_gdp] and [demo_pjan].

Due to rounding columns may not total.

This table represents the input-output effect[52], in which the first column investment costs appear. By simply dividing Spain's investment by the total, it turns out that Spain represents a 9.1631%. Multiplying figure 48 results, the corresponding graph that explains the Spanish situations has this aspect[53]:

[52] Explained in the report, it is not important for this calculation
[53] Both situations will be shown to see the absolute difference

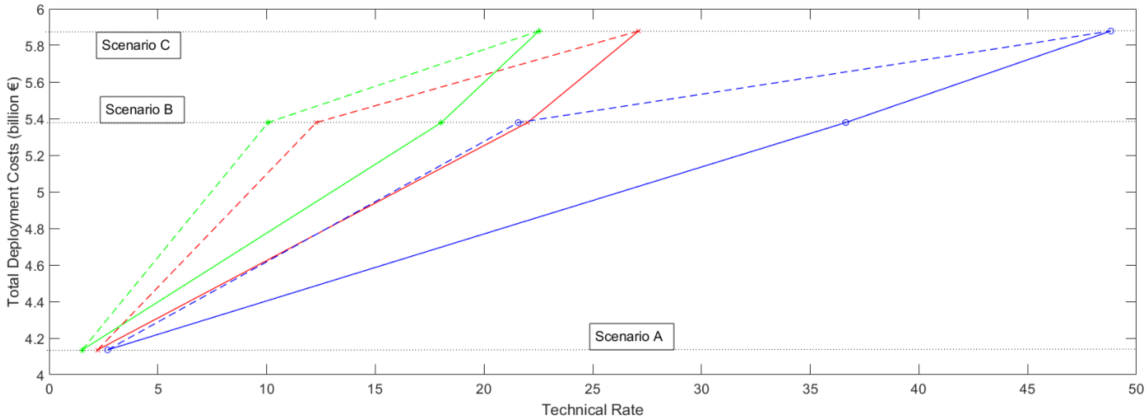


Figure 50. Total aggregated costs (billions €) vs. TR for each case of use (Spanish situation)

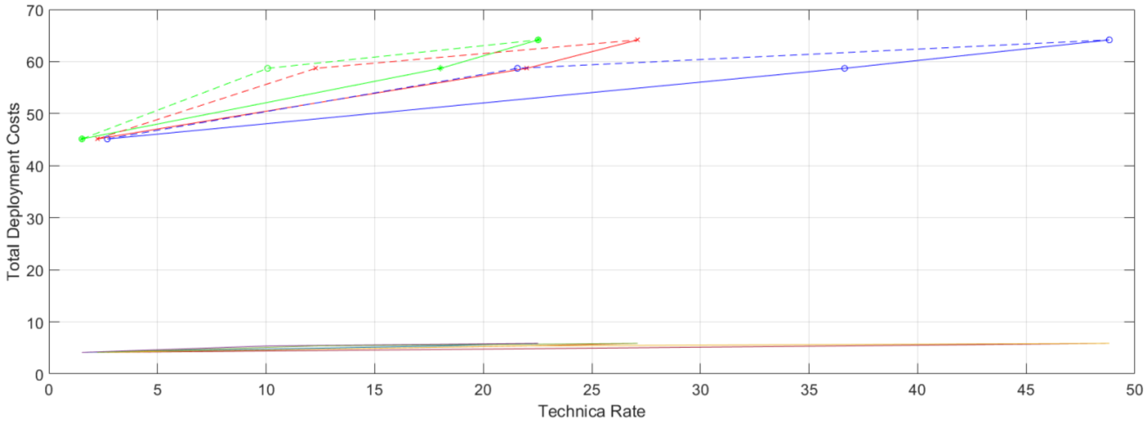


Figure 51. Total aggregated costs (billions €) vs. TR for each case of use (compared situations)

In both figures, the Spanish trend and the European trend (also shown in figure 48) are shown. In figure 50 it can be seen that the trends are exactly equal but changing y-axis values. On the other hand, in figure 51, both situations are compared. Needless to say that, due to the small percentage that Spain represents in the EU, the compared variation is almost unnoticeable.

5.2 Social and economic impact

The whole document has focused on costs and technical performance, both for the analysis and the results analysis. However, from a business and economic point of view, this information is incomplete: when deciding whether to invest or not in a particular project, revenues and costs must be compared. If revenues (through the years) turn out to be bigger than costs, the project will be viable.

Furthermore, there are some revenues that are not monetary. They take into account the impact that this project has in society: how it helps it to develop and evolve, what indirect benefits it triggers, etc. This section will discuss both the economic and social impact of this "project": the deployment of a 5G network in Spain.

5.2.1 Economic impact: expected output

As it was introduced, the estimated benefits/revenues that any project is predicted to trigger is one of the most important factors to consider, together with the costs. Hence, this section will be divided into the corresponding explanations of costs and revenues.

a. <u>*Costs*</u>

Costs have been partly analysed in this document. Particularly, deployment costs have been the studied ones. However, mobile communications systems have more associated costs. Like any other systems, they can be classified as Capital Expenditures (CAPEX) and Operational Expenditures (OPEX).

- CAPEX is a business expense incurred to create future benefit (i.e., acquisition of assets that will have a useful life beyond the tax year). They are incurred once (when acquiring an asset). In mobile communications market, CAPEX would be driven from the acquisition of base stations, servers, routers, etc.

- OPEX is referred to those expenditures required for the day-to-day functioning of the business, like wages, utilities, maintenance and repairs. In this studied case, OPEX would be driven from equipment maintenance, network infrastructure (e.g. cables, spectrum bands, etc.) rental and leasing.

At this point, the first thing to discuss is whether 5G costs and particularly deployment costs (analysed throughout the document) can be considered CAPEX, OPEX or both. First premise is that depending on the service provider will undoubtedly vary. However, the two types will always be incurred in.

Some ideas that have been discussed throughout the document[54] give evidence that the majority of the costs will be OPEX:

- SDN and NFV, which will be crucial in 5G networks, are based in software. Software has practically its entire costs associated to OPEX.

- eNodeBs will be enhanced and software-changed to support LTE-A Pro and 5G requirements respectively. Either these base stations are rented or have been earlier deployed (case in which CAPEX would be already taken into account), this costs would be regarded as OPEX.

[54] They are mainly explained in chapter 3 where the key factors of 5G networks are explained.

- Femtocells are a CAPEX special case. Although they can be considered as cell deployment costs, the final user will pay for it since it works as a router as well. Then, CAPEX in this case is reduced.

- The common Cloud-base converged core will simplify costs for every service provider and operators, since the costs will be divided by the providers that use the actual core.

- Network slicing will enable such flexibility that costs will be reduced, since only the needed slices will be implemented. In this sense, CAPEX costs will be reduced to the deployment of general centralized servers which will be software-configured to give only the needed functionalities.

- Efficiency (for spectrum, power and network) are studied KPIs defined for 5G networks which pretend to improve up to 100 times.

b. *Revenues*

Revenues are mainly based in the number of subscribers, who will pay for the services provided. In this sense, there is a good perspective since society's trend seems to be the increasing dependence on technology and communication systems. This trend can be seen in table 18, where subscriptions are expected to keep growing, but especially in traffic per user.

In terms of traffic, users have changed the trend by leaving voice traffic aside and focusing on data traffic (apps, web browsing, games, etc.). As a consequence, providers have focused their new fares in monthly data offering. This strategy has aroused more benefits since applications, video streaming and these kinds of cases of use carry on more and more data, due to the better quality and more extended service.

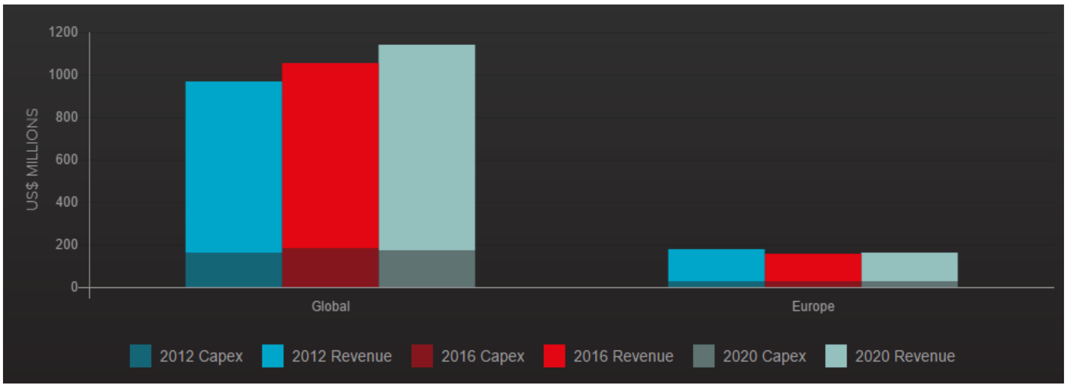


Figure 52. Global and European revenues and CAPEX comparison. Source: GSMA Intelligence [33]

Figure 52 shows the global and European relation between CAPEX and Revenues for years 2012, 2016 and 2020. For both scenarios, it is interesting to see how CAPEX increases from 15% to 17% between 2012 and 206 but decreases again to 15% in year 2020. This supports the conclusions explained in this section.

5.2.2 Social impact: GPT of a connected society

To complete this chapter, the social impact of 5G networks will be discussed. This term is referred to the effects and consequences (not only but also money-relate) that 5G networks will cause on society.

a. *<u>General Purpouse Technology</u>*

A General Purpouse Technology (GPT) is defined as catalyts for transformative changes that redefine work processes and rewrite the rules of competeitive economic advantage [28]. There is no doubt that mobile communications have become more and more important, playing an indispensable role in todays society and business environment.

However, 5G pretends to go beyond this importance and become a factor that will redefine business' models, innovations and services. This intention is based in the cases of use (discussed in chapters 4 and 5) that will not only cover mobile communication, but also other cases that will affect business development and performance: MIoT or autonomous vehicles. Furthermore, it will trigger a modernization in terms of regulation, which will have to contemplate the new needs of a connected society by smart cities and homes as well as new threads in the fields of public safety, cyber-security and privacy amongst any others. Figure 53 shows the importance of mobile industry to the GDP, reaching a total of 4.16 billion dollars (4.9% of the GDP) by 2020.

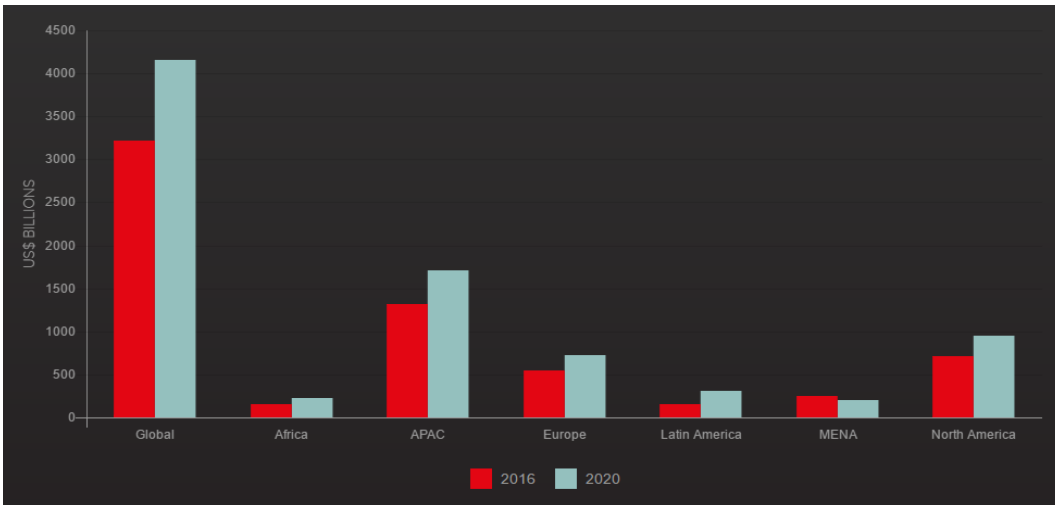


Figure 53. Mobile industry contribution to Gross Domestic Product. Source: GSMA Intelligence [33]

b. *Social goals: inclusive mobile communications*

Such an important role that 5G will play in the upcoming years will trigger an important responsibility. Both service providers by their own initiative and public institutions (European Commission in this case) will have to ensure the inclusion and achieving of social goals proper of GPTs.

On the one hand, GSMA Intelligence discusses in its report [33] ”Sustainable Development Goals”. It was a strategic plan which was adopted by the United Nations in 2015 in which mobile industry can help. Some of them are[55]:

- Financial inclusion by providing mobile money facilitates and access to financial services.
- Helping agriculture by MIoT, which will enable Smart Agriculture.
- Health improvement by enabling MCS cases of use, which included e-Health and improving equipment.
- Disaster Response by enabling MCS as well. It will enable (almost) errorless applications which will help in dramatic environment disasters.

On the other hand, public institutions must ensure social goals that also help to grant inclusiveness within all the countries and their respective regions. As for Europe, Digital Agenda plan, carried out by the European Commission has set some mandatory objectives that service providers must comply with by 2020. Some of these objectives are granting a basic coverage in rural areas as well as basic internet speed.

c. *Influence in vertical industries*

Becoming a GPT will mean that 5G will have a heavy weight in other industries' development. Sources [28], [30] and [33] discuss what are the exoected vertical industries that will be most affected by 5G. They all conclude that automotive/transport, healthcare and information & communications industries will be the ones that will enable a greater increase in the outcome.

Figure 54 shows the expected outcome of several indsutries by year 2035 by IHS [28]. Moreover, it includes a prediction that was founded by this company, which states that “**5G will enable $12 trillion of global economic acitivty in 2035”**.

[55] The whole 17 points are fully explained in report from source [33]. Only the most related ones to the cases of use analysed have been extracted.

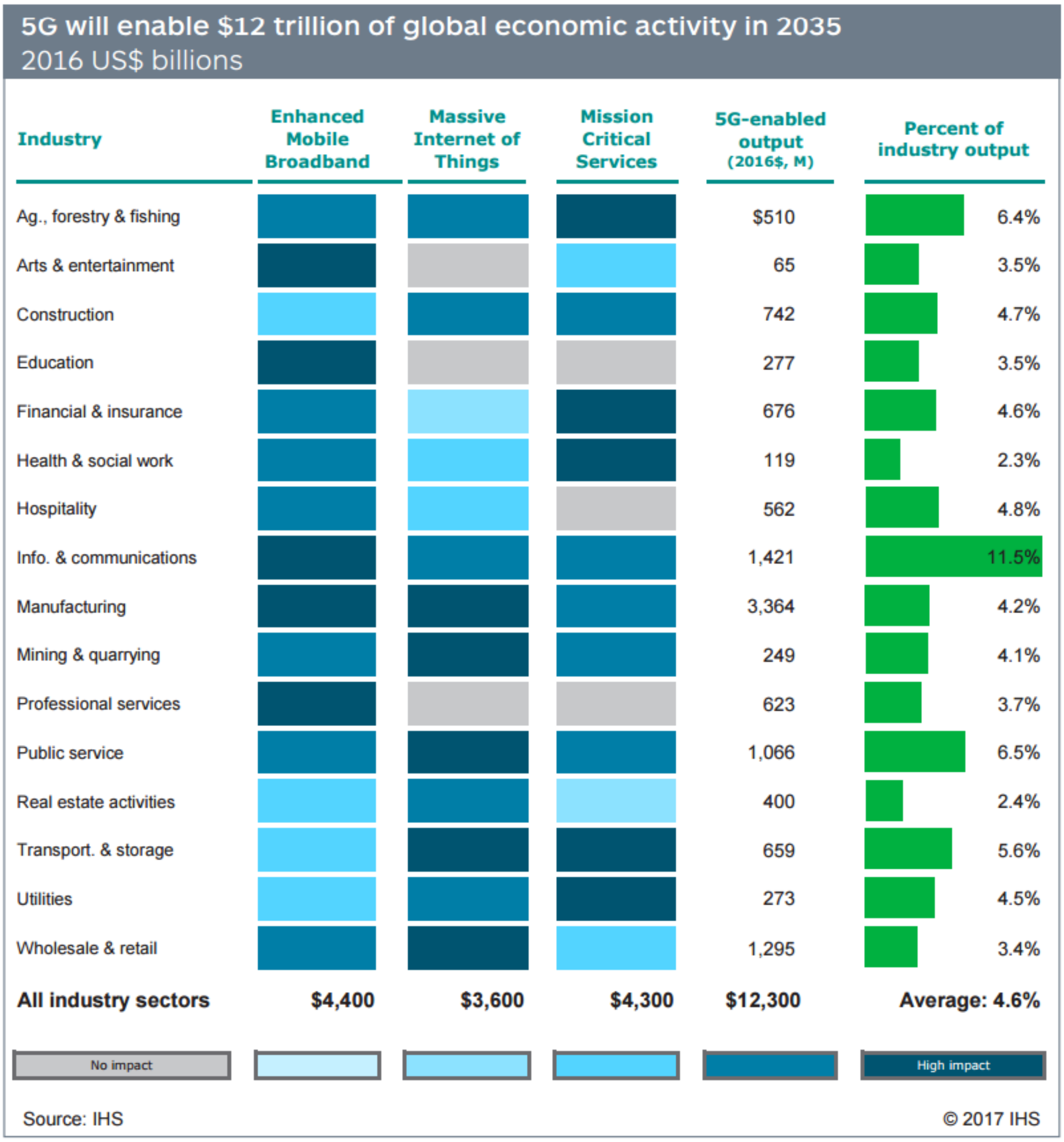


Figure 54. IHS's forecast on 5G impact on vertical industries by 2035. Source: IHS [28]

Therefore, as for the social impact of 5G, it can be concluded that:

• As for becoming a GPT, 5G will transform social behaviours and business' models. Its influence will go beyond economy and mobile communications area.

• Private and public sectors implementing 5G will have to acquire social responsibilities and address social challenges. 5G will provide unique opportunities to address these challenges by improving mobile communications.

• Regarding 5G's importance in rethinking business models, influence in vertical industries come a consequence of business remodelling. Each case of use is specified (by a colour) in terms of impact within each vertical industry. Besides economic impact (output), the more importance cases of use have, the more will they change industries.

6. Conclusions and future work

This last chapter will discuss two different issues. On the one hand, section 6.1 will explain the conclusions that have been reached on this analysis. On the other hand, section 6.2 will explain the constraints and limitations this study has faced as well as possible future investigation lines.

6.1 Conclusions

Many conclusions can be taken out from this study. First, quantitative conclusions regarding the techno-economic analysis and demand forecasts can be analysed. Nonetheless, qualitative ones on the advantages and drawbacks that 5G holds will also be explained.

6.1.1 On the 5G opportunity

Every new mobile generation has unveiled new chances and applications that have enhanced mobile communications to the point we know today. In this sense, every new generation was based on evolving and enhancing the previous one. However, 5G promises to be a true revolution, adding new cases of use and changing consumers' habits. Furthermore, 5G is set to trigger big changes in other vertical industries beyond mobile communications, pursuing to get considered as a GPT. It is a non-trivial objective to reach the figure of $12 trillion of global economic activity in 2035 (see section 5.2.1) by becoming a necessary enabler for other industries' output. Besides, 5G has an important role as an inclusive technology that will be indispensable to address social issues.

From a technical point of view, 5G represents a convergent technology. Benefits of a technology that will interwork with other key technologies such as Wi-Fi and will be sliced and customizable overwhelm drawbacks. Other important new technologies opportunities such software-designed networks and function virtualization enable a flexibility that will be key to upgrade all the services. And lastly, the different spectrum bands in which 5G will be able to work provide 5G of an even bigger flexibility and more opportunities to maximize technical performance.

From the economic point of view, 5G represents a good inversion as for the relatively small costs. Due to software-based networks implemented in part by SDN and NFV, operators will save in deployment costs (since ENodeBs will be reused). Furthermore, with the chance of operating in unlicensed spectrum (such as 5GHz), spectrum costs will be reduced. They will however incur in R&D (Research & Development) to achieve technologies capable of operating in those different frequencies and managing new inconveniences such as air interference in new bands or cyber-security in the Next Generation Core. In conclusion, that cost peak that operators have in transitional periods between generations will be smaller and will last less time.

Lastly, from a standardization and regulatory point of view, there are some challenges to consider. Public institutions must modernize their regulations and laws to adapt to this new connected era. They must be agile enough to keep up with communications industry breakthroughs in order to enable them with the appropriate regulatory framework. This obviously includes agile and coherent band harmonization, updating spectrum management models and international consensus to achieve a global technology. As for standardization, 3GPP and ITU are key-role players that must be supported, since 5G depends largely on them.

6.1.2 On the analysis methodology

This document has followed a top-approach methodology to address the techno-economic analysis, since only KPIs but not technical components (e.g. ENBs) were analysed. The entire work, which would have included every single technical and economic concept of the RAN and the Core, would have been impossible to achieve. First, because of the work dimension the study would hold. And secondly, because of the still diffuse state of art.

Assumptions: scenarios selection and binding

The study could not have been feasible if some assumptions had not been made. The most useful assumptions were scenarios selections and the binding between years-scenarios. These assumptions (made in chapter 4) linked the different players (European Commission, ITU and 3GPP) work and connected the theory discussed throughout chapter 1 to 3 with the analysis.

As for the binding, it was explained that this was not extremely precise, since deployment dates are still unknown and will vary upon countries and service providers. However, they were sufficiently good since they were chronologically ordered in the dates that ITU proposes for IMT plan. Anyway, there is a slightly error range assuming that service providers might deploy these scenarios in same order but earlier or later.

Scenarios selection was the other big assumption to start the analysis. Despite being sufficiently justified, some service providers could follow another deployment trend. Scenario A and C are logical scenarios to start and end the analysis. Hence, Scenario B, which represents the transitional period, is the only one whose validity is questionable.

6.1.3 On the results

Results that were reached seemed congruent since the followed methodology was logical[56], but there were two options: optimistic or pessimistic assumptions (that were based on the technical rate (TR) weight of each segments of the network).

---

[56] Methodology was logical since it was based on comparing different scenarios and figures

Consulted Nokia experts suggested that "a cost peak (smaller than in the transition to LTE) would be incurred by service providers that will first deploy the first 5G networks". This information gives evidence that the pessimistic assumption is more feasible.

As for the scenarios-cases of use recommendation scores (Table 15 in section 5.1.2), it is going to be validated in this section. Figure 55 below, from GSMA, shows the requirements that several applications (included within the three cases of use studied) will hold.

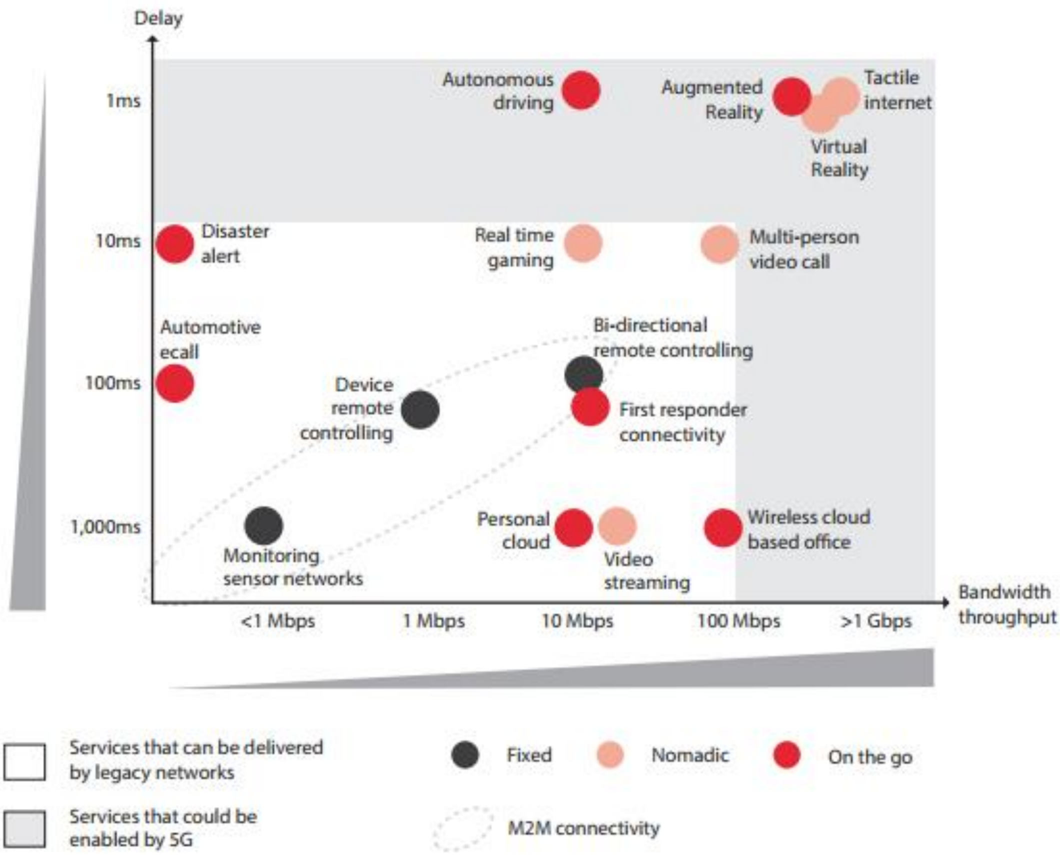


Figure 55. Different applications' requirements. Source [33]

Although figure 55 axes are delay and bandwidth throughput (both KPIs for ITU), there are two regions shadowed in white and grey. As the legend shows, the services inside the white region can be delivered by legacy (Scenario A, the first one) networks, while the ones inside the grey area could be enabled by 5G. Needless to say, these last services will be the most discriminative and demanding ones.

Furthermore, there is a sub-region inside the white one, limited by a dashed line. This region contains machine-to-machine (M2M) applications. In other words, IoT services. This supports the result from chapter 5 that concluded that Scenario A is optimum to develop MIoT.

As for MCS, the applications to focus are autonomous driving and disaster alert mainly. As for autonomous driving, there is no doubt that is one of the most demanding applications. Disaster alert stays in the limit between the two shadowed regions (region which might be considered to be the transitional scenario). Although it does not totally agree with the RS given to MCS, there would be sufficiently accurate results. Moreover, mobility is not considered in figure 54, yet it holds the biggest possible weight for MCS applications.

Tactile Internet, Augmented Reality and Virtual Reality belong to eMBB case of use. But video streaming, personal cloud, real time gaming or personal cloud belong to it as well. The different locations of these services in figure 54 clearly support the aforementioned diversity of eMBB case of use. The first mentioned group will be the most demanding apps and will be therefore only achieve by advanced 5G networks (Scenario C, last step). On the other hand, each service has a different location and could therefore be achieved by deploying either Scenario A (initial stage) or B (transitional scenario).

In conclusion, RS given to eMBB are supported by figure 54: the best recommendations score is given to the final scenario which will achieve to serve the most demanding applications, while a RS of 2 is given to the other two scenarios that will be sufficient to enable some of the less demanding applications.

6.2 Future work

Future work and constraints of this study are intimately related. The possible lines to work further on this issue would only be plausible by overcoming the constraints. There are three main directions to extend this work:

• **Ascertaining assumptions:** this line is based in ascertaining some aspects and premises that were treated as assumptions in this document. The main constraint why it has not been possible to follow this direction is the complexity of the issue. One aspect that could be objective of future work is the identification of each segment's weight in the technical performance of the entire network. Remember than Scenario B's technical rate were calculated by making an optimistic and pessimistic assumption. Pessimistic assumption turned out to be more feasible according to cost peak statement. However, there were many other weights combinations that could satisfy this statement.

On the other hand, there were 12 scenarios and 3 cases of use (combination of even more applications), which would result in 36 crosses. Although they were reduced to 9 crosses, there could be some scenarios that would work optimally for a specific application within a case of use. Needless to say that following this direction, the analysis could have resulted endless.

But, does it make any sense to study all the crosses? Due to network's flexibility, it is known that many scenarios could be deployed by slicing the network and programmes the necessary software in each node. Hence, almost any combination of case of use and deployment scenario becomes feasible.

• **Developing 5G in next yeras**: this investigation direction will be naturally followed in the future. It will come alongside 5G standardization progress, with the outcome of Releases 14, 15 and 16, where 5G technologies will be standardized. By the time when more technical aspects are already clarified, it will be easier to detail each segment's costs as well as the entire network architecture.

Furthermore, known standardization will provide the possibility to better select deployment scenarios. It seems logical that the other option for a transitional period would be that scenario with New Radio connected to Enhanced Packet Core (enhanced EPC). However, there is another alternative. It would consist on analysing the situation of an operator that would skip this transitional period and directly deploy Scenario C. Once again, knowing architectural changes between scenarios, it could be reasonable to analyse whether an operator should skip or not this transitional scenario.

• **Deepening into hardware and software**: this investigation line would consist on specifying the technical details of the elements of both segments of the network. By detailing and studying each element architecture and components, it could be addressed the specific operations to achieve the KPIs. For example, if all eNodeBs' components were studied (filters, modulation type, carrier spacing, channel codification), a very deep study would get to know what changes would be necessary to increment the peak data rate, spectrum efficiency, etc.

There could be two investigation lines regarding this down approach trend: RAN and Core. And what is more: within them many of the topics that could be addressed (filtering, security, channel codification) will have to be investigated collaboratively to make the two pieces of this jigsaw fit.

Bibliography

[1] Ericsson, “Ericsson Mobility Report, Western Europe”, November 2016

[2] Jian A. Zhang, Peng Cheng, Andrew R. Weily, Y. Jay Guo. 2014. *Towards 5th Generation Cellular Mobile Networks* . Australian Journal of Telecommunications and the Digital Economy, Vol 2, No 2, Article 34. http://doi.org/10.18080/ajtde.v2n2.34(link is external). Published by Telecommunications Association Inc. ABN 34 732 327 053. https://telsoc.org

[3] ITU, “MEMBERSHIP: CONTACT. COMMIT. CONNECT. BECOME A MEMBER OF ITU”, ITU, 2017 in http://www.itu.int/es/about/Pages/membership.aspx (Accessed January, 2017)

[4] ITU, “What does ITU do?”, ITU, 2017 in http://www.itu.int/en/about/Pages/whatwedo.aspx (Accessed January 2017)

[5] ITU GT IMT-2020, “Focus Groups on IMT-2020”, ITU, 2017 in http://www.itu.int/es/ITU-T/focusgroups/imt-2020/Pages/default.aspx (accessed January 2017)

[6] 3GPP, “About 3GPP Home”, 3GPP 2017 in http://www.3gpp.org/about-3gpp/about-3gpp (accessed January 2017)

[7] Inside5G, “3GPP publishes initial timeline for 5G”, Inside5G 2015 in http://inside5g.com/3gpp-publishes-initial-timeline-for-5g/ (accessed in January 2017)

[8] European Commission, “Digital Single Market-European 2020 strategy”, EC 2015 (last update May 2017) in https://ec.europa.eu/digital-single-market/en/europe-2020-strategy (accessed February 2017)

[9] ETSI,” ETSI Summit on 5G Network Infrastructure”, ETSI 2017 in http://www.etsi.org/news-events/events/1148-etsi-summit-on-5g-network-infrastructure (accessed February 2017)

[10] Ministry of Internal Affairs and Communications (Japan), “Principal Uses and Characteristics of Radio Wave”, Telecommunications Bureau of the Ministry of Internal Affairs and Communications2017 in http://www.tele.soumu.go.jp/e/adm/freq/search/myuse/summary/ (accessed February 2017)

[11] ITU, “World Radiocommunication Conferences (WRC)”, ITU-R 2017 in http://www.itu.int/en/ITU-R/conferences/wrc/Pages/default.aspx (accessed February 2017)

[12] ITU, “Regionally harmonized bands”, ITU-R 2017 in http://www.itu.int/net/ITU-R/index.asp?category=information&rlink=emergency-bands&lang=en (accessed February 2017)

[13] “Outcome of the World Radiocommunication Conference 2015”, M. Maniewizc at DSA Global Summit

[14] EC, “EC Press Release”, EC 2014 (last update February 2017) in http://europa.eu/rapid/press-release_IP-14-957_es.htm (accessed February 2017)

[15] Xataka-Movil, “Así está el panorama del 4G en España, su cobertura y últimas novedades”, Xataka-Movil 2015 in https://www.xatakamovil.com/mercado/asi-esta-el-panorama-del-4g-en-espana-su-cobertura-y-ultimas-novedades (accessed February 2017)

[16] Andy Sutton-Chief Architect’s Office, TSO, “5G Network Achitecture And Design”, SlideShare 2017 in https://www.slideshare.net/3G4GLtd/5g-network-architecture-and-design (accessed March 2017)

[17] LTE World, “Radio Resource Control (RRC) | LTE World”, LTE World 2017 http://rcrlte.wpengine.com/wp-content/uploads/2013/01/Network-diagram.jpg (accessed March 2017)

[18] NetManias Reports, “Analysis of LTE-WiFi Aggregation Solutions”, NetManias 2016 in http://www.netmanias.com/en/post/reports/8532/laa-lte-lte-u-lwa-mptcp-wi-fi/analysis-of-lte-wifi-aggregation-solutions (accessed March 2017)

[19] A.G. Armada and L. Hanzo, “A Non-Coherent Multi-User Large Scale SIMO System Relying on M-ary DPSK,” IEEE ICC, Jun. 2015 pp 2517-2522

[20] Open5GCore, “Fundamental Core Network Functionality”, Open5GCore 2017 in http://www.open5gcore.org/en/open5gcore/fundamentalcorenetworkfunctionality (accessed March 2017)

[21] 5G PPP Architecture Working Group: “View on 5G Architecture”, July,2016

[22] Inside5G, “Ericsson Looks To Slice The Cake for 5G”, Inside5G 2015 in http://inside5g.com/ericsson-looks-to-slice-the-cake-for-5g/ (accessed March 2017)

[23] Ericsson Technology Review: “A vision of the 5G Core: Flexibility for new business opportunities”

[24] Nick McKeown; et al. (April 2008). "OpenFlow: Enabling innovation in campus networks". ACM Communications Review. Retrieved 2009-11-02.

[25] SDX Central, "NFV and SDN: What's the Difference?", SDX Central 2014 in https://www.sdxcentral.com/articles/contributed/nfv-and-sdn-whats-the-difference/2013/03/ (accessed March 2017)

[26] Andy Sutton-Chief Architect's Office, TSO, "5G Network Achitecture And Design", SlideShare 2017 in https://www.slideshare.net/3G4GLtd/5g-network-architecture-and-design (accessed March 2017)

[27]TR-161266 via http://blog.3g4g.co.uk/2016/09/5g-new-radio-nr-architecture-options.html

[28] IHS Economics & Technology "The 5G Economy: How 5G technology will contribute to the global economy", January 2017

[29] TelecomCloud, "Scaling IoT with LTE/MTC", Telecom Cloud 2017 in http://www.telecom-cloud.net (accessed April 2017)

[30] European Commission by Tech4i2, Real Wireless CONNECT, Trinity College Dublin & InterDigital "Identification and quantification of key socio-economic data to support strategic planning for the introduction of 5G in Europe", 2016

[31] Ericsson "Western Europe Ericsson Mobility Report", November 2016

[32] CNMC DATA, "Informe Annual 2016", CNMC 2016 in http://data.cnmc.es/datagraph/jsp/inf_anual.jsp (accessed May 2017)

[33]GSMA & GSMA Intelligence "The Mobile Economy", 2017 in https://www.gsma.com/mobileeconomy/

Appendixes

Appendix 1: *Matlab Optimization script*

This appendix will explain more in detail the self-programmed Matlab script that was used to optimize the economic problem. First thing to explain is that Matlab's task is to find the optimum value for the desired function. This means that it does not compute whether a result is valid or not as long as it satisfies the conditions and constraints.

As stated in section 4.3 there were two functions that used one common algorithm; interior-point algorithm. Both the algorithm and pair of functions can be used to solve nonlinear multivariable functions. As a matter of fact, syntax is the same for both functions, which booth look like:

x = fminimax(fun,x0,A,b,Aeq,beq,lb,ub,nonlcon,options)

x = fmincon(fun,x0,A,b,Aeq,beq,lb,ub,nonlcon,options)

Parameters are the same:

- **fun** is the function which value has to be minimized (or its maximum cost minimized for fminimax). In this case the function represented the total deployment costs amongst the three scenarios:
$$f(EC_{Core}, EC_{RAN}, DC_{core}, DC_{RAN}) = EC_{Core} + EC_{RAN} + DC_{core} + DC_{RAN} \quad ec(16)$$
- **A & b** are matrixes that represent linear inequations:
$(EC_{RAN} < DC_{RAN} \ \& \ EC_{core} < DC_{core})$ ec(18)
- **Aeq & beq** are matrixes that represent linear equations (scenarios equations)
- **lb & ub** represent lower and upper bounds (defined by $0.5 < \varepsilon < 1$ & $0 < \mu < 1$)
- **nonlcon** are the nonlinear conditions, both equations and inequations. In this case, our nonlinear condition was scenario B definition
$$DC_{core} + (\varepsilon \times EC_{RAN} + (1-\varepsilon) \times DC_{RAN}) = 139.75 \quad ec(20)$$
- **options** was defined by function 'optimoptions', where the algorithm and other plot options were defined

The output parameter x was a vector of 6 positions containing the optimal value for all the 6 variables to optimize (EC's, DC's, ε and μ) and the optimal value for function f

Function fmincon is easier to perform: it took 22 iterations to reach the optimal value. On the other hand, it took 132 iterations for fminimax to reach the optimal value, since minimax has to find the smallest maximum cost, which is more difficult to evaluate.

Appendix 2: *Harmonization in the EU: Digital Agenda & ETSI*

This appendix will give complementary explanations to European Union's technological and standardization works. Needless to say, its situation is almost defined in sections 2.1.1 & 2.1.2, like the overwhelming majority of countries all around the world. Nevertheless, there are other institutions that must be considered to fully understand the current situation in this our region. It shall be mentioned reader could extrapolate to other regions' cases from this situation, changing the supranational legislative-regulatory body for its homologue in that actual region (e.g. Federal Communications Commission in USA). This legislative organism for Europe is the European Commission, and another group which is specialized in telecommunications issues: CEPT (European Conference of Postal and Telecommunications Administrations).

The Digital Agenda is part and one of the seven pillars of an even bigger project of the European Commission called Europe 2020. The Digital Agenda proposes to best exploit the potential of Information and Communication Technologies (ICTs) in order to foster innovation, economic growth and progress.

Figure 56. Pillars of Europe's Digital Agenda. Reconstruction of EC.Source [8]

In figure 56, the pillars of the Digital Agenda are displayed. The aims are not important as a standalone matter of study, but it is important to understand that European Commission is the one to set the society's challenges, many of which won't be possible without 5G: ultra-fast internet access for all, ICT (Information and Communication Technologies) benefits, interoperability…

Hence, it is clear the different roles of those who analyse society needs (EC), the ones who question technological needs and set objectives (ITU) and the ones to make standards (3GPP among other bodies). There is another link in this chain, a standard elaborator that works alongside 3GPP, and is different in each world's region. In Europe, this actual link is ETSI.

The European Telecommunications Standards Institute (ETSI) is an independent, not-for-profit, standardization organization in the telecommunications industry (equipment makers and network operators) in Europe with worldwide projection. ETSI produces globally-applicable standards for Information and Communications Technologies (ICT), including fixed, mobile, radio, converged, broadcast and internet technologies. ETSI was created by CEPT in 1988 and now works as an independent organisation (collaborating with other regulators). Its labour is important as long as it participates in 3GPP standardization process and releases, basing its contribution on the necessities Europe demands and EC propose. ETSI bases its work in writing what they call deliverable types, which are different types of documents that may contain from proposed standards to reports or standardization guides.