NTNU
Norwegian University of
Science and Technology

5G and its Economic Aspects

Literature Review and Selection of a
Connection Portfolio Under Risk

Jenny Antoinette E. Hansen

Industrial Economics and Technology Management
Submission date: June 2017
Supervisor: Alexei A. Gaivoronski, IØT

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

Problem Description

The objectives of this thesis are to study how the fifth generation of mobile networks (5G) is currently defined according to the existing literature, identify economic aspects of relevance, and show an example of how financial portfolio optimization can be applied to plan future 5G mobile networks, under risk.

Preface

This master's thesis is written in the spring of 2017 at the Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management. As I have taken my specialization in Financial Engineering, supported by courses in Digital Economics and ICT Economics, it seemed appropriate to choose a topic for my final thesis that combine these different fields.

The overall objective of this thesis has been to explore the next generation of mobile networks, 5G, and its economic aspects. The starting point was to study existing literature, to understand how 5G currently is defined. Based on this literature review, economic aspects that should be subject for further research were identified. As I have also taken introductory courses in Operations Research and Optimization, I have chosen to model a planning problem for a mobile network provider in the context of 5G. This is my quantitative contribution, as the rest of this thesis is based on the conducted literature review. As I unfortunately not have taken my specialization in Applied Optimization, I have not been able to implement the proposed model and perform a numerical analysis of it. However, the work with the model has still been very interesting, and I have learned a lot from it.

The reader of this master's thesis is presumed to have knowledge of the economic and business aspects of telecommunication at a university level, as well as some general knowledge about mobile networks and the field of optimization.

I would like to thank my supervisor Professor Alexei A. Gaivoronski for the help he has given me in all aspects of this thesis, especially with the model of the planning problem. I would also like to thank my family and friends for their support during this time.

Jenny A. Engeland Hansen

Oslo, Norway

June 10, 2017

Abstract

In this thesis, we study the economic aspects of the next generation of mobile networks, the fifth generation (5G). Mobile networking is an interesting topic, as telecommunication is an indispensable part of the infrastructure in our society, both today and in the future. 5G is needed as the existing 4G networks do not have enough capacity to handle the future amounts of mobile data traffic. It is important to understand the economic aspects of 5G, to ensure that the next generation of mobile networks will be economically sustainable. Although 5G still is a few years away from being commercially deployed, now is a good time to start addressing certain economic aspects. For instance, it is wise to begin the considerations of the cost aspect of the future networks early, while the enabling technologies are still under research and development, and before 5G is standardized. This is especially important for the mobile network operators, who provide the infrastructure. Similarly, the business potential of 5G is of interest for all actors that intend to generate revenues related to 5G. Some examples of these actors are service providers, content providers, companies selling telecommunication equipment (hardware), and again, the mobile network operators.

The three overall objectives of this thesis are to explain how 5G is currently defined, identify the most important economic aspects of 5G, and formulate an optimization planning problem for a mobile network operator operating in the 5G context, by using portfolio theory. Our main intention is to provide an overview that can be used as a starting point for further studies of 5G economics.

First, using existing literature about 5G as a starting point, we explore how 5G currently is defined, and how the recent literature address its economic and business related aspects. 5G is not standardized yet, but the literature shows that most actors and stakeholders agree on both the use cases and the actual requirements for 5G. We see that until now, most of the 5G literature is concerned with proposing technologies to be included in 5G, so that 5G can meet the requirements. Heterogeneous networks, device-to-device communication, software-defined networking, network function virtualization, cognitive radio, transmission over the millimeter frequency bands, advanced multiple-input multiple-output techniques and software-defined radios are examples of important technologies that are likely to characterize 5G.

An important finding is that there has not been conducted much dedicated research on the economic aspects of 5G yet. However, there are a few articles dedicated to this field, which are studying the profitability of proposed new technologies: device-to-device communication and the 5G heterogeneous networks. Certain economic aspects are also mentioned in technology-oriented publications, such as the increasing expenses the mobile network operators face, and the need for these costs to be at a reasonable level in the future. Cloud concepts with software-based implementations and virtualization technologies are described as promising solutions for the cost issues, along with lowered cell costs

in the 5G heterogeneous networks. If the costs of deploying and operating base stations are lowered, the decision of deploying new base stations in rural areas will be easier to make. This have direct impact on socio-economic development, for which connectivity is an important enabler. Further, there is a certain focus on green aspects in the literature, especially in terms of power consumption in the base stations. These considerations are also directly relevant for the economic context of 5G.

Regarding the 5G business perspective, we see that the improved network performance in 5G is likely to open up new business opportunities. In addition to an improved user-experience for existing services, new opportunities arise from applications that are first enabled when the network delay is minimized and the reliability is increased. This includes mission critical communications, a huge number of connected devices arising from the Internet of Things, and mobilization of different industries. The business context of 5G may also be characterized by new types of customers and partnerships.

Finally, as an example for further economic research of 5G, we provide a portfolio theory model for the planning of future 5G networks. In this model we show how portfolio optimization can be applied to formulate the following problem: How a mobile network operator can select an optimal portfolio of available connections to offer to its customers, to provide the necessary connections for the network traffic related to the customers' service demands, while also meeting their quality demands. The application of portfolio theory to allocate traffic in a mobile network is not novel. Our main contribution is to include the Quality of Experience aspect in this type of model, to make the model feasible for the planning of *5G* mobile networks.

Sammendrag

I denne masteroppgaven utforskes økonomiske aspekter ved neste generasjons mobilnettverk, den femte generasjonen, mest kjent som 5G. Mobilnett er et interessant tema for mange, siden telekommunikasjon er en uunnværlig del av infrastrukturen i samfunnet både nå og i fremtiden. 5G er nødvendig fordi dagens 4G-nettverk ikke har nok kapasitet til å håndtere den fremtidige mengden av mobil datatrafikk. Videre er det viktig å forstå økonomiske aspekter ved 5G for å sikre at neste generasjons mobilnettverk er økonomisk bærekraftige. Selv om 5G fortsatt er noen få år unna en kommersiell lansering, er nå et godt tidspunkt for å starte å vurdere enkelte økonomiske aspekter. For eksempel er det klokt å vurdere kostnadsaspektet for fremtidige nettverk tidlig, mens teknologien bak fortsatt er under forskning og utvikling, og før 5G standardiseres. Dette er spesielt viktig for mobiloperatørene som eier infrastruktur. Tilsvarende er forretningspotensialet ved 5G av interesse for alle aktører som har til hensikt å generere inntekter via 5G. Eksempler på slike aktører er tjenesteleverandører, innholdsleverandører, selskaper som selger maskinvare til telekommunikasjon, for eksempel produsenter av mobiltelefoner, og igjen: mobiloperatører.

De tre overordnede målene med denne studien er å forklare hvordan 5G er definert på nåværende tidspunkt, identifisere de viktigste økonomiske aspektene ved 5G, og formulere et planleggingsproblem for en 5G-operatør, ved bruk av porteføljeteori og optimering. Hovedhensikten med oppgaven er å skape en oversikt som kan benyttes som utgangspunkt for videre økonomiske studier av 5G-nettverk.

Ved å ta utgangspunkt i eksisterende litteratur om 5G utforsker vi hvordan 5G er definert på nåværende tidspunkt, samt hvordan litteraturen adresserer økonomiske og forretningsrelaterte aspekter. 5G er ikke standardisert ennå, men litteraturen viser at de fleste aktører og interessenter er enige om både brukssituasjoner og de faktiske kravene for 5G. Vi ser at mesteparten av 5G-litteraturen hittil har fokusert på å foreslå teknologier som bør inkluderes i 5G for å imøtekomme kravene. Heterogene nettverk, direkte kommunikasjon mellom enheter, programvaredefinerte nettverk, virtualisering av nettverksfunksjoner, kognitiv radio, bruk av millimeterfrekvensbåndene, avanserte MIMO-teknikker og programvaredefinerte radioer er eksempler på viktige teknologier som sannsynligvis vil karakterisere 5G.

Et viktig funn er at det ikke har blitt forsket så mye spesifikt på de økonomiske aspektene ved 5G hittil. Det finnes imidlertid noen få artikler som studerer forretningslønnsomheten ved følgende foreslåtte nye teknologier: direkte kommunikasjon mellom enheter, og 5G heterogene nettverk. Visse økonomiske aspekter er også nevnt i teknisk orienterte publikasjoner, som for eksempel økningen i mobiloperatørenes kostnader, og nødvendigheten av at disse kostnadene forblir på kontrollerte nivåer i fremtiden. Skyløsninger med programvarebaserte implementasjoner og virtualiseringsteknologier beskrives som lovende

løsninger for kostnadsreduksjon, sammen med lavere cellekostnader i 5G heterogene nettverk. Hvis kostnadene for å sette opp og drifte en basestasjon reduseres, er det enklere å ta valget om å sette opp flere nye basestasjoner der hvor det trengs, særlig i rurale områder. Dette har direkte innvirkning på samfunnsøkonomisk utvikling, da muligheten for nettverkstilkobling er en faktor for dette. Videre er det et visst fokus på miljørelaterte aspekter i litteraturen, særlig med tanke på strømforbruk i basestasjonene. Disse betraktningene er også direkte relevante for 5G sin økonomiske kontekst.

Det forventes at forbedret nettverksytelse i 5G vil åpne nye forretningsmuligheter. I tillegg til en forbedret brukeropplevelse for eksisterende tjenester, så vil nye muligheter oppstå i form av applikasjoner som først kan brukes på trådløse nettverk når nettverksforsinkelsene minimeres og nettverkspåliteligheten økes. Dette inkluderer misjonskritisk kommunikasjon, et stort antall tilkoblede enheter tilknyttet Tingenes Internett, samt prosessen ved å utvide bruken av trådløs kommunikasjon i ulike industrier. Nye typer kunder og forretningssamarbeid kan også karakterisere forretningskonteksten ved 5G.

Til sist, som et eksempel på videre økonomisk forskning på 5G, så utleder vi en porteføljeteorimodell for planlegging av fremtidige 5G-nettverk. I denne modellen viser vi hvordan porteføljeoptimering kan anvendes for å formulere følgende problem: Hvordan en mobiloperatør kan velge en optimal portefølje av tilgjengelige forbindelser å tilby til sine abonnenter for å overføre nettverkstrafikken assosiert med kundenes etterspørsel av tjenester, og samtidig møte krav om gitt kvalitet på forbindelsene, gitt at operatøren ønsker å maksimere sin totale profitt. Modellen inkluderer valg av pris for å tilby ulike tjenester til ulike markedssegmenter, samt beslutningen om hvorvidt operatøren bør investere i utvidet nettverkskapasitet. Anvendelsen av porteføljeteori for å allokere trafikk i mobilnettverk er ikke ny. Vårt hovedbidrag er å inkludere *Quality of Experience*-aspektet i en slik type modell, for å gjøre modellen egnet for å planlegge fremtidige *5G*-nettverk.

Table of Contents

List of Tables

List of Figures

Abbreviations

3GGP = The 3rd Generation Partnership Project
5G = 5th Generation Mobile Networks
AI = Artificial Intelligence
AWA-HetNet = Advanced Wireless Access Heterogeneous Network
CAPEX = Capital Expenditure
CMaaS = Connectivity Management-as-a-Service
CN = Core Network
CR = Cognitive Radio
D2D = Device-to-Device
DC-DC = Direct D2D Communication with Device Controlled Link Establishment
DC-OC = Direct D2D Communication with Operator Controlled Link Establishment
DR-DC = Device Relaying with Device Controlled Link Establishment
DR-OC = Device Relaying with Operator Controlled Link Establishment
EE = Energy Efficiency
HetNet = Heterogeneous Network
ICT = Information and Communications Technology
IP = Internet Protocol
ISP = Internet Service Provider
IT = Information Technology
LTE = Long-Term Evolution
M2M = Machine-to-Machine
METIS = The Mobile and Wireless Communications Enablers for the Twenty-Twenty Information Society
MIMO = Multiple-Input and Multiple-Output
MPT = Modern Portfolio Theory
MU-MIMO = Multi-User Multiple-Input and Multiple-Output
MVNO = Mobile Virtual Network Operators
NGMN = The Next Generation Mobile Networks Alliance
NFV = Network Functions Virtualization
OPEX = Operating Expenditure
RAN = Radio Access Network
RANaaS = Radio Access Network-as-a-Service
RAT = Radio Access Technology
RRM = Radio Resource Management
SDN = Software-Defined Networking
SDR = Software-Defined Radio
SE = Spectral Efficiency
SINR = Signal-to-Interference-and-Noise Ratio
UDRANETs = Ultra-Dense Radio Access Networks

QoE = Quality of Experience
QoS = Quality of Service

Chapter 1

Introduction

At this time, the 4G cellular networks have existed for several years, and it is time to look forward and see what the future will bring regarding the next generation of cellular networks: the fifth generation, most often referred to as 5G. Electronic communications and especially the Internet have an important position in our society, and hence 5G is a a topic of interest for many, not only the telecommunications industry itself. Wireless communication networks are perhaps the most critical element in the global ICT strategy (Wang et al., 2014), and the telecommunications industry has over the years grown to become an enormous industry on a global basis, and one of the fastest growing and most dynamic sectors in the world.

Today, “everyone” are already online, and with the emerging Internet of Things, every*thing* will also be online - everywhere, and always. There is a demand for ubiquitous access to mobile services. As a further introduction and a starting point to understand the road towards 5G, the evolution of earlier generations of cellular networks will be briefly explained.

Since 1G emerged in the 1980s, a new generation of mobile networks have been launched approximately every ten years. In the beginning, voice services or mobile telephony was in focus, and during the years, a clear shift from fixed to mobile telephony has occurred. By the end of 2010 there were four times more subscriptions for mobile telephony than fixed-line telephony, and there are no circumstantial evidences saying that this trend will be reversed (Sharma, 2013). Now, voice services is simply a part of the total communication need the users rely on mobile networks to cover. The user device is no longer mainly a mobile telephone, but rather a small computer. Users expect a wireless connection of the same quality as the maximum of what the device is capable of, similar to the quality of a wired connection on a desktop at home or in the office (Patel et al., 2012). Further, if users suffer from a poor WiFi or wired connection and are located in areas with good cellular coverage, they will easily switch to a cellular connection, relying on the fact that this will solve their connection trouble.

According to Sharma (2013), the cellular wireless generations (G) refer to a change in the fundamental nature of the service, non-backwards compatible transmission technology and new frequency bands. This is backed up by Chávez-Santiago et al. (2015), who describe the evolution from 1G to 4G as mainly characterized by a shift in the multiple access method, and improved modulation and coding schemes. Multiple access methods, or channel access methods, are methods that handle how several terminals can transmit over the same transmission medium. They are based on methods for multiplexing, i.e. techniques that combine multiple signals into one signal, which is then transmitted over a shared medium. Hossain (2013) describes that the cellular generations have traditionally differed from each other in terms of radio access, data rates, bandwidth and switching schemes. This implies that a new generation of cellular networks requires a much bigger technological change than a simple improvement of the existing performance of the last network generation. 5G is not expected to be an exception. To further understand the path towards 5G, the evolution of earlier network generations will be briefly explained.

1.1 The Evolution of Cellular Networks

The first generation of wireless telephone technology were analogue systems using frequency division multiple access (FDMA) for voice call modulation (Hossain, 2013). These systems were designed for voice only application (Nokia, 2014). Sapakal (2013) describes 1G with the following characteristics: (...) low capacity, unreliable handoff, poor voice links, and no security at all (...) in a retrospective manner. This indicates how simple the 1G systems were, compared to the standard of the current 4G systems. The world's first cellular system was launched by Nippon Telephone and Telegraph (NTT) in Japan in 1979, and a few years later in Europe and the United States (Sharma, 2013).

The second generation of cellular networks was introduced in the end of the 1980s (Sharma, 2013). The 2G system used digital modulation schemes: time division multiple access (TDMA) and code division multiple access (CDMA) (Sapakal, 2013). The digital signals used for voice transmission improved the voice service. It was the first generation to provide text messaging, more commonly known as Short Message Service (SMS), and it also enabled circuit switched data access (Nokia, 2014). Compared to 1G systems, 2G provided higher spectral efficiency and more advanced roaming (Sharma, 2013), which lead to an improvement of the voice services. 2G is generally associated with the global system for mobile communications (GSM) standard, which describes the 2G protocols. An important step between 2G and 3G was the implementation of a packet-switched domain, the general packet radio service (GPRS), along with GSM. The latter is referred to as 2.5G, as this technology expanded the existing 2G systems considerably.

Due to low data rate services, 2G did not fulfill the need for mobile Internet access of decent quality. This lead to new requirements and a demand for the 3G standards. The first 3G networks were launched in 1998. The characteristics of 3G are more capacity for voice services, and affordable mobile Internet with faster data services (Nokia, 2014). Sharma (2013) describes that 3G is not simply one standard, but rather a family of standards which can all work together. 3G networks enable network operators to offer the users a wider range of more advanced services, while an improvement of the spectral efficiency makes it possible to achieve greater network capacity. The Enhanced Data Rates for GSM Evolution (EDGE) technology is a part of the 3G definition, although it is considered a pre-3G technology. Several standards are associated with 3G, but the most important one is probably the Universal Mobile Telecommunications Service (UMTS) system, which was developed to complement the old GSM standard.

The fourth generation of cellular networks, 4G, is mainly associated with the Long-Term Evolution (LTE) system (Sapakal, 2013), which provides high capacity and higher rate data services for mobile multimedia (Nokia, 2014). Mobile networks are composed of two main components, the radio access network (RAN) and the core network (CN). In 4G networks, the RAN and CN components are referred to as LTE and System Architecture Evolution (SAE), respectively. These two components together are called the Evolved Packet System (EPS), which is an all-IP network, supporting packet-switched connectivity only (Yazici et al., 2014). This means that no part of this generation is based on circuit switching, which was an important intention with 4G.

The overall objective of 4G is to provide a coherent and secure IP-based solution that support voice, data and multimedia services to users everywhere, anytime. The fact that 4G is IP-based enables faster data transfer than the earlier generations. Along with LTE, Worldwide Interoperability for Microwave Access (WiMax) is also considered an important 4G technology. However, these technologies are actually pre-4G technologies. The improved standard LTE Advanced meets the International Telecommunication Union's (ITU) requirements for real 4G compliance. In some countries, this standard is marketed as 4G+. As intended, the improved data experience of 4G is the most noticeable difference from 3G from a user's point of view. This shows how the data use cases have become more important than voice and messaging services alone. While 3G was an operator-centric concept, 4G is a service-centric concept (Janevski, 2009).

1.2 Why 5G is needed

So, if 4G already provides decent mobile broadband, is there really a need for 5G? Yes, and this is due to several reasons. When 4G was developed, mobile broadband was the main driver. For 5G, mobile broadband is still an important driver, but with more respect to reliability, latency, throughput, data volume and mobility (Rost et al., 2014). The future 5G systems are expected to provide significant gains over 4G in terms of higher data rates, much better levels of connectivity and improved coverage (Chávez-Santiago et al., 2015). Overall, the need for increased network capacity is the main challenge 5G must solve. There has been an enormous growth in the use of wireless Internet over the last years, and there are no signs that this tremendous growth will slow down, as the demand growth for both user data rates and network capacity is still continuing. The growth in data traffic in addition to the increasing volume of connected devices necessitates the evolution of LTE towards 5G (Yazici et al., 2014). Rost et al. (2014) have summarized the following reasons for why the demand for mobile data increases:

- More connected devices (especially due to the Internet of Things).

- The connected devices are more powerful and more complex than before.
- The demanded services are more diverse, more complex and more bandwidth-hungry.
- Devices are integrated into more sides of the society, including the industry.
- User terminals are used as a gateway to access cloud services.

The use of wireless technology has exploded, with a 92 % growth in mobile broadband per year since 2006 (Wang et al., 2014). With the existing growth, the capacity limits of the existing mobile networks will soon be reached. As today's 4G networks will not be able to handle the predicted amount of traffic in the future, 5G is needed to overcome the limits of the current systems (Nokia, 2014). To obtain a significant increase is the capacity of a wireless system, there is a need for an increased number of wireless infrastructure nodes, increased use of the radio spectrum and an improvement in the link efficiency (Bhushan et al., 2014). Yazici et al. (2014) describe the networks' inflexible and expensive equipment and the fact that the current solutions for the networks' control plane are complex and non-agile, as two important reasons for the capacity problems. These factors make it hard to improve the capacity of the existing networks.

The strategy for 5G's capacity and performance growth must be economically sustainable. This mean that it should offer increasingly better coverage and a superior user experience at a lower cost than the existing wireless systems, LTE included (Bangerter et al., 2014). As the work on this strategy is ongoing and the transition period to 5G has started, Bangerter et al. (2014) declare that the 5G era effectively already have begun. This is a natural part of the network evolution as the 4G technology has matured over the years that have pasted since it was first launched.

We are moving towards a networked society, where there will be unconstrained access to information. Data sharing will be available for everyone, everywhere. This change implies that future mobile systems will need to handle vastly different challenges and expectations than the existing systems (Olsson et al., 2013). The main purpose of 5G is to design a fully mobile and connected society (NGMN, 2014), and the best wireless world, free from limitations and hindrance of the previous generations (Hossain, 2013), to obtain the World Wide Wireless Web (WWWW) (Sharma, 2013). 5G is also positioned to empower socioeconomic transformations in many different ways, including transformations for productivity, sustainability and well-being (NGMN, 2014). Compared to the service-centric 4G concept, 5G is assumed to be user-centric, i.e. the user will be on top of all (Janevski, 2009).

There seems to be a consensus in the existing research that 5G is expected to be released around 2020 (Sapakal, 2013), when the limits of 4G is predicted to be reached (Nokia, 2014). This also makes sense based on the fact that earlier mobile network generations have been launched approximately every tenth year. All forecasts indicate that South Korea will be the first country to deploy 5G. Their biggest operator, SK Telecom, cooperates with leading partners like Ericsson, Intel, Nokia, Samsung, Qualcomm and Deutche Telecom, to test 5G technology. In 2016, SK Telecom claimed that their aim was to deploy a full-scale pre-commercial 5G network by the end of 2017 (Vollen, 2016). A different South Korean operator, KT, claimed in December 2016 that they will launch 5G commercially by 2019 (Marek, 2016). The operators in South Korea are also planning to perform their biggest test of 5G so far at the Olympic Winter Games in Pyeongchang, February 2018 (Lucas, 2017). A successful test at this event will give several hundreds of thousands of spectators the possibility to experience 5G, and it will indeed give South Korea huge advantages in the global 5G race.

1.3 Outline

The rest of this thesis is organized as follows. In chapter 2, existing literature defining 5G will be presented, with focus on use cases, requirements, proposed technical solutions, and economic aspects, including the green perspective and the business aspects. In chapter 3, the presented literature will be discussed, with focus on the economic aspects. In chapter 4, we present a portfolio optimization model for network planning from a mobile network operator's point of view, in the context of 5G. Conclusions are drawn in chapter 5, where recommendations for further research are also presented.

Chapter 2

Literature review

In this chapter, relevant literature regarding 5G is reviewed. 5G is not standardized yet, and this affects the existing literature is several ways. Most 5G actors and stakeholders agree on the use cases for 5G and what the actual requirements for 5G are or will be (Carlton, 2017), but the technologies behind are not formally decided yet. In the second half of 2017, The 3rd Generation Partnership Project (3GPP) will shift the focus of their work to deliver the first set of 5G standards, in *Release 15* (3GPP, 2017). 3GPP is an international project that unites several organizations that develop telecommunication standards. Much of the existing 5G literature are proposing different technologies for 5G, especially solutions for how the required increase of capacity can be obtained. In comparison, there is conducted considerably less dedicated research on the economic and business related aspects of 5G.

To explore the 5G literature, the review is organized in the following manner: Use cases, requirements and some of the suggested technologies are presented, before we move on to the green perspective, and the economic and business aspects. The review continues in chapter 3, where the presented literature is discussed.

2.1 5G Use Cases

To give the reader an idea about which communication challenges 5G is expected to solve, the most important use cases proposed for 5G will be presented. Use cases are descriptions of interactions between an actor, typically the user, and a system, to achieve a goal. They also make it clearer to understand the background for the 5G requirements, and they are likely to be the driver for the 5G technology that will be developed (Nokia, 2014). Some of the 5G applications will be old and familiar, while an introduction of new and more diverse services is also expected. The existing dominating scenarios of human-centric communication will be complemented by an enormous increase in communication directly between machines or fully automated devices. Hence, the use cases for 5G may be split into two following main categories: mobile broadband and the Internet of Things (IoT). The use case category of mobile broadband consists of use cases where the user access the Internet directly by intention, for instance by browsing with a personal device like a computer, tablet or smartphone. It is expected that for 5G, the mobile broadband will be improved, and new services will be introduced.

2.1.1 Machine-to-Machine Communication and the Internet of Things

Compared to the mobile broadband use case, the IoT category covers the use cases where devices and sensors communicate directly with each other, without a user being present at all time. This is referred to as machine-to-machine (M2M) communication (Rost et al., 2014), and this type of communication will be an important part of the emerging IoT. Here, devices and items embedded with sensors, electronics, software and network connectivity collect and exchange data. Sensors or devices for M2M communications will be integrated in daily use objects such as cars, household appliances, textiles and health-critical appliances.

The result of the IoT and M2M communication is that everything that can benefit from a wireless connection, will be connected (Olsson et al., 2013). As a result, the IoT adds anything as an additional dimension to connectivity, in addition to anywhere and anytime (Agyapong et al., 2014). M2M communication will also create a huge increase in the number of subscribers (Chen and Zhao, 2014), as there are forecasts of a total of 50 billion connected devices by 2020 (Osseiran et al., 2014). The coexistence of human and machine communications will lead to a large diversity of communication characteristics. This is especially because the IoT requires more reliable communication links, but also lower transmission delays or latencies, as machines can process information much faster than people (Nokia, 2014).

Although the IoT is predicted to define 5G in several ways, it should be clarified that the IoT already have emerged over the last years. However, improved wireless connectivity is definitely needed to realize the full potential of the IoT and to satisfy all the possible users who may want to benefit from the IoT in their daily lives. What we have seen yet is probably just the beginning, and the IoT is expected to grow further in the next years, especially when the capacity of wireless networks is improved. Wearable devices with artificial intelligence (AI) capabilities are one example of new types of devices expected in the 5G era.

2.1.2 General Use Cases

The Mobile and Wireless Communications Enablers for the Twenty-Twenty Information Society (METIS), the EU flagship 5G project, mentions healthcare, security, logistics, automotive applications and mission-critical control as important 5G use cases. They have also identified five scenarios, or objectives, that is expected to drive the 5G research direction (Osseiran et al., 2014):

- Amazingly fast: 5G should provide very high data-rates for the users of mobile broadband. The users should perceive the response as instantaneous.

- Great service in a crowd: Users should get a reasonable broadband experience in crowded areas, like stadiums, shopping malls, outdoor festivals and other public events with many people attending.
- Best experience follows you: Users on the move should experience high levels of service.
- Super real-time and reliable connections: 5G should enable new applications and use cases with strict latency and reliability requirements.
- Ubiquitous things communicating: A large number of different devices with unique requirements should be handled efficiently.

These five scenarios are located in the intersection of use cases and the actual requirements, and should give the reader an idea about the enormous 5G expectations. Wang et al. (2014) specify that 5G should support communications for special scenarios the 4G networks do not support. For instance, 4G only supports traveling scenarios up to a velocity of 250 km/h, while existing high-speed rails easily can reach a pace of 350-500 km/h. This specifies the scenario “Best experience follows you” by Osseiran et al. (2014).

NGMN (2014) have identified twenty-four use cases for 5G, which they have grouped into eight use case families:

- Broadband access everywhere.
- Broadband access in dense areas.
- Higher user mobility.
- Massive Internet of Things.
- Extreme real-time communications.
- Lifeline communications.
- Ultra-reliable communications.
- Broadcast-like services.

These eight categories are not meant to be exhaustive, but rather intended as a tool to ensure that the required level of flexibility is well captured. Some use cases are covered by several use case families, and seen together, the use case families give an overview of the scope of the use cases.

Nokia (2016) mentions autonomous vehicles as a relevant use case in the context of mission critical communication. This is one of the 5G use cases that captures the requirements of high reliability and very low latency. Other use cases in the same context are augmented reality (AR) and virtual reality (VR). AR enhances a real-world view with graphics, and real-time information is displayed based on the location or vision of the user. VR creates a new user experience, where the user is in a fully immersive environment.

Both AR and VR devices need to track the user movements accurately, process the movements and receive images, and then display the response immediately. AR can be used to enhance existing service experiences. Shoppers can see how clothes would look on them without physically trying them on, firefighters can use AR to see ambient temperature and construction details for a building, and police officers can use AR with facial recognition to identify a suspect before arresting. VR can be used for virtual classrooms where the teacher is remote, or to design and prototype products before they are actually built. The latter shortens development time and cost considerably (Nokia, 2016).

Other mission critical use cases are remote surgeries, remote robotics and industry control and automation. In the future, robots may be used for tasks that have typically been performed by humans in the past. Two examples of remote robot tasks are bomb disposal or firefighting (Nokia, 2016), which could have fatal outcome if the connection is not reliable or have too high latency.

2.1.3 Device-to-Device Communication

In the 3GPP Release 12, the use case of device-to-device (D2D) communication is seen for the first time. Yazici et al. (2014) describe this as a significant step in the direction of evolving LTE toward 5G. The objective of D2D communications is to allow nearby devices to establish local links, so that the traffic can flow directly through them, instead of through the network infrastructure. D2D will be effective for traffic offloading (Li et al., 2014) and may help diminish traffic congestion in the cellular core network (Chávez-Santiago et al., 2015), as it will ensure that the base stations no longer are the bottleneck between source and destination. There are proposed D2D concepts for both direct communication between two proximate devices, and communication between sender and receiver through other devices, where intermediary devices in the latter case will work as relays in the network. D2D can potentially improve the user experience by reducing latency and power consumption, increasing the peak data rates, and creating new proximity-based services such as proximate multi-player gaming (Bhushan et al., 2014).

2.2 5G Requirements

As the 5G use cases are broader than ever before, 5G will require more diverse link characteristics, and hence, the requirements in general are more diverse than for earlier network generations. Evolution and innovation through the years, as well as new technology, have opened up opportunities that few would have imagined as possible when 1G first became a reality. The diverse 5G applications will have a wide range of characteristics and requirements, which add up to the general 5G requirements. Some applications will need lower latency and high reliability, while lower reliability is sufficient for other applications. The new and varying Quality of Service (QoS) requirements will need to be taken care of in the system design (Olsson et al., 2013).

While some 5G applications can be supported by the existing mobile broadband or the evolution of it, other applications will impose the new requirements. The use case classes

IoT and VR/AR are examples of drivers for the new and diverse requirements 5G will have to support. The future networks will need to handle new types of connected devices, while ultra high resolution video will require better mobile broadband (Chávez-Santiago et al., 2015). Nokia (2014) expects that in 5G, we will see a shift away from best-effort mobile broadband towards truly reliable communication. 5G must support a wide range of data rates, and scalability and flexibility of the network are required. If scalability and flexibility is properly handled, the networks will be able to support the many new devices with low complexity and requirements for very long battery lifetimes in an efficient manner (Osseiran et al., 2014). Rost et al. (2014) agree that flexibility ans scalability are fundamental requirements for 5G, in order to allow for network optimization for individual scenarios in different time and space.

According to forecasts, the growth in data traffic volumes will be exponential for the next years. In the next decade, the total volume of mobile traffic is expected to increase a thousand times the traffic volume today, mostly due to the increasing number of connected devices (Olsson et al., 2013). Hence, it is also expected a thousand-fold capacity increase for 5G (Li et al., 2014). There was a thousand-fold increase of the capacity of wireless systems from 2000 to 2010, and now the same order of increase is necessary again. The requirement of improved capacity is related to the need for access to more spectrum, over higher carrier frequencies. 5G must be designed to enable deployment in higher frequency bands (Nokia, 2014).

In addition to the thousand-fold capacity increase, Wang et al. (2014) list up the following widely agreed requirements for 5G, compared to 4G: 10 times the spectral efficiency (data volume per area unit), energy efficiency and data rate, and 25 times the average cell throughput. Another 5G requirement presented by Chen and Zhao (2014) is to handle higher assymetry in the data traffic. The ratio of downlink to uplink traffic have been 6:1, but this may rise to as much as 10:1 over the next years, due to the growing proportion of video traffic. Further, Chávez-Santiago et al. (2015) mention lower outage probability and better coverage, higher bit rates over larger coverage areas, higher versatility and higher reliability of the communication links in the network. Agyapong et al. (2014) add massive device connectivity as a specific requirement to the list.

The specific 5G requirements identified by the METIS project are (Osseiran et al., 2014):

- 1000 times higher mobile data volume per area.
- 10 to 100 times higher number of connected devices.
- 10 to 100 times higher user data rate.
- 10 times longer battery life for low power massive machine communication.
- 5 times reduced end-to-end latency.

This is compared to the performance of 4G.

Nokia (2016) points out that the requirements for reduced end-to-end latency and higher reliability are especially important in the context of mission critical control, or other use cases related to healthcare, security, logistics and automotive. This is because ultra-reliability is vital for safety, while low latency ensures that applications are usable. Latency also determines the perception of speed, and the increasing use of real-time functionality demands the lowest possible delay in the network. Further, high reliability makes the users confident in depending on wireless communications even in critical or actual life-threatening situations. Nokia (2014) explains that delays of less than 10 ms are absolutely required, because the human eye is sensitive to delays for multimedia applications. However, for the AR and VR use cases, an end-to-end latency of more than 5 ms increase the risk for that the users may experience cyber sickness, also known as virtual reality sickness. This is a state that causes symptoms similar to symptoms of motion sickness, and is an uncomfortable and nauseating experience for the users. Cyber sickness has a negative impact on the user experience, which means that delays in AR and VR applications should be minimized. These are example of how the human communication will be more demanding in the future.

There are not only proposed technical requirements for 5G, but also requirements regarding the costs. Osseiran et al. (2014) specify that the requirements they have listed shall be fulfilled at similar cost and energy dissipation as today. Yazici et al. (2014) describe that the 5G architecture should be an agile solution of high capacity, and to a low cost. If these requirements are met, both user satisfaction and the profitability of the mobile service providers will be ensured. Energy-efficient operation, resource efficiency and cost efficiency are also requirements mentioned by Demestichas et al. (2013). Chávez-Santiago et al. (2015) specify that especially the costs for infrastructure deployment should be lowered, while Agyapong et al. (2014) explain that it should be an objective for 5G to reduce both the capital expenditures (CAPEX) and the operating expenditures (OPEX) for network provision.

2.2.1 Intelligent Quality of Service Management

Hossain (2013) defines Quality of Service (QoS) as a network's ability to achieve maximum bandwidth and handle other network performance elements, such as latency, error rate and uptime. QoS also involves control and management of network resources. This is done by setting different priorities for specific types of data transmitted over the network. QoS also include dedication of bandwidth, control of jitter, ensuring low latency and improvement of loss characteristics (Hossain, 2013).

Today, users select wireless interface manually, and only one at a time. Patel et al. (2012) describe the possibility of being connected to multiple technologies at the same time, and also be able to switch between them for different services, as a main feature of 5G. An ultimate request for 5G is to provide intelligent QoS management over different networks. Until now, mobile networks have typically been dimensioned with respect to peak capacity, and the performance have been evaluated in "hard" metrics, such as peak data rates, coverage and spectral efficiency (Bangerter et al., 2014). In the future, they should rather satisfy the increasing traffic demand by a flexible availability of capacity in time and space (Olsson et al., 2013). As 5G has many other requirements than increased maximum throughput (Hossain, 2013), the devices should store different QoS measures in a database (Javaid, 2013). Such measures could be delay, jitter, losses, bandwidth and reliability (Gohil et al., 2013). Then, the 5G devices could test the service quality based on these measures.

Further, if these data were combined with intelligent algorithms, and given constraints for required QoS and personal costs, the best wireless connection could be identified by the 5G terminal. This will make the concept "always best connected" in heterogeneous networks a reality, as "best connected" is associated with the best quality (Sharma, 2013). To have a connection in compliance with QoS requirements is important for many different use cases, and especially for real-time services. As the future services become more complex, this will be even more important. Yazici et al. (2014) suggest some of the same ideas, referred to as connectivity management-as-a-service (CMaaS). As every wireless system has its distinctive characteristics and roles, the choice of the most appropriate technology will vary for different services at different places at different times. Gohil et al. (2013) mention that it also could be possible to choose among different mobile network operators at the same time. As 5G shall accomodate the QoS requirements for all possible types of applications, another ultimate goal for 5G would be to interconnect all wireless heterogeneous networks to provide a seamless and consistent telecommunication experience for all users (Javaid, 2013).

Wang et al. (2014) also discuss the challenge of evaluation of wireless communication networks and optimization of performance metrics, as they agree that several metrics is needed for a complete and fair assessment of the 5G systems. In addition to QoS, they suggest spectral efficiency, energy efficiency, delay, reliability, fairness of users and implementation complexity. They argue that the several metrics, the better result, and that there should be a trade-off among all the available metrics. This will require high-complexity joint optimization algorithms and long simulation times.

2.2.2 Quality of Experience

Bangerter et al. (2014) suggest that expanded performance metrics regarding the user's quality of experience (QoE) will be seen in 5G. QoE describes the user's subjective perception of how well a service or application is working. QoE factors are highly application and user specific, and hence they cannot be generalized, although low latency and high bandwidth generally will improve the QoE (Agyapong et al., 2014). In 5G, more intelligent use of network data will be seen. 5G should facilitate optimal use of network resources for QoE provisioning and planning. To meet the users' QoE needs, the future networks will need to be more aware of application requirements and QoE metrics (Bangerter et al., 2014). Connectivity is increasingly evaluated by the users, in terms of QoE. One example is the IP telephony service Skype, which gives their users the possibility to rate the experienced quality after communication sessions. The users tend to evaluate connectivity regardless of time or location, and they also tend to be unforgiving towards their network operators or providers when their expectations are not met. Thus, the mobile networks operators constantly face the risk of churn, i.e. losing their subscribers to competitors (Agyapong et al., 2014). The importance of keeping the subscribers satisfied to control the risk of churn may be even more critical in the future, when the users are likely to rely even more on their Internet connections than today.

According to Bangerter et al. (2014), QoE factors could also include ease of connectivity with nearby devices, and improved energy efficiency. 5G systems should be context-aware, to provide a high QoE for services. Context information should be utilized in a real-time manner, based on network, devices, applications, the users and their environment. The increased context awareness will contribute to more user-centric and personalized services, and may also allow for improvements in the efficiency of existing services (Bangerter et al., 2014). To deliver the required level of personalization and offer a more user-centric and context-aware experience, the 5G networks will need to cooperate in new ways. This requires closely coordinated radio access technology (RAT) selection and management at both the network and device level, in order to maintain an optimal user experience.

Both Demestichas et al. (2013) and Agyapong et al. (2014) agree that consistent QoE provisioning and QoE satisfaction are requirements for 5G. Rost et al. (2014) explain that the complexity of the user scenarios for 5G will make it challenging to manage and operate the mobile networks efficiently, while also providing the demanded quality of experience. Hence, proper QoE management is both a requirement and a challenge for the 5G networks.

2.3 5G Technology

As there are vast expectations to 5G, there are many challenges that need to be solved. As mentioned, most of the existing 5G literature so far are concerned with suggesting different technological solutions to solve the different challenges. This section will give an overview of some of the most featured and promising technologies. The overview is not meant to be exhaustive, but rather give the reader an idea of which solutions that seem most likely to be included in 5G, according to the literature. First, the general expectations will be summarized, before selected topics are described more thoroughly. These topics are the user terminal, the Internet protocol, heterogeneous networks, device-to-device communication, software-defined networking and network function virtualization, cognitive radio, the radio spectrum, network densification, advanced multiple-input and multiple-output (MIMO), indoor scenarios and ultra dense radio access networks (UDRANETs), beam division multiple access (BDMA), radio access network-as-a-service (RANaaS), architecture of the core network, and technology from the perspective of mission critical communications.

2.3.1 General Expectations

NGMN (2014) describes that 5G will operate in a highly heterogeneous environment, characterized by the existence of multiple types of access technologies, multi-layer networks, multiple types of devices and multiple types of user interactions. In such an environment there is a fundamental need for enablers to achieve seamless and consistent user experience across time and space. This have motivated the research on technology for 5G. Li et al.

(2014) explain that the main driver of the capacity growth for 5G is expected to come from advancements in the network architecture. Key techniques will be heterogeneous networks and convergence of information and communication technology. Chávez-Santiago et al. (2015) describe 5G as a convergence of wireless technologies. Bangerter et al. (2014) expect a shift towards network efficiency, with 5G systems based on architectures for dense heterogeneous networks. This is also a promising solution regarding the low cost perspective, while meeting the need of capacity growth and delivering an uniform connectivity experience to the users.

The METIS project has compiled one of the most thorough reviews of what 5G is expected to be like. Their paper seem like a suitable further introduction to the technology aspect of 5G. The METIS project's overall approach towards 5G is to evolve existing technologies, and complement them by new radio concepts, as a single new radio access technology neither will be able to satisfy all requirements, nor replace the existing networks. They have carried out research on technology components targeted to each of their identified 5G scenarios. The suggested concepts are massive MIMO, Ultra Dense Networks, Moving Networks, Device-to-Device, Ultra Reliable and Massive Machine Communications. These concepts will allow 5G to support the expected increase in the mobile data volume, while at the same time broadening the range of application domains mobile communications can support beyond 2020 (Osseiran et al., 2014).

To obtain higher data rates and lower latencies beyond what the 4G networks are capable of, 5G has to adopt revolutionary ways of using the radio spectrum. The current consensus is that to address the radio access network challenges, there is a need for a combination of more spectrum resource, higher spectral efficiency, network densification and offloading (Agyapong et al., 2014). It is a major challenge for 5G to increase the system capacity and quality, given the limited available frequency spectrum (Hossain, 2013), but 5G must still solve the challenge. Wang et al. (2014) agree that the physical scarcity of radio spectrum allocated for cellular communication is a crucial challenge. Chávez-Santiago et al. (2015) describe that the already allocated frequency bands can be exploited in better ways by enhanced modulation schemes and traditional network scaling. Further, introduction of cognitive radio (CR) will provide an opportunistic access to the under-utilized spectrum. This will also increase the capacity of the systems. CR is further described in 2.3.7. Chávez-Santiago et al. (2015) also describe that different offloading techniques will contribute to an almost unlimited access to large amounts of multimedia data. Reconfigurable platforms for the implementation of CR in underlay cellular networks and traffic offloading solutions will be enabled by software-defined radio (SDR) technology. Software-defined networking (SDN) will facilitate the programmable operation of the core network. SDN is described further in 2.3.6.

Bhushan et al. (2014) also present the path to 5G as an enhancement and scaling of current 4G technologies, with a few new components. They expect that this combination can raise the overall user experience to a whole new level. According to Chávez-Santiago et al. (2015), the features of extensive use of 5G technologies like CR, SDR and SDN, combined with evolution technologies, such as new modulation schemes and traffic offloading, will differentiate 5G from previous network generations. Bangerter et al. (2014)

mention small cell base stations as one of the technologies expected to be a key technology for 5G that is already available. Other key technologies will be developed and deployed over the coming years. 5G will require technology that enables advanced source coding, advanced radio access networks (RANs), such as heterogeneous networks (HetNets), and advanced radio access technologies (RATs). Bangerter et al. (2014) further describe that transport technologies at cell sites, both fronthaul and backhaul, need to be significantly improved in terms of both speed and deployment flexibility. Yazici et al. (2014) envision fully decoupled, independently scalable and programmable control and user planes for the 5G network. Both Chen and Zhao (2014) and Agyapong et al. (2014) describe solutions for a flexible separation and coordination of the control and user planes in the network, to address the capacity and data rate challenges in 5G. Li et al. (2014) agree that an architecture where the control and user planes are separated can be used as a general framework for the 5G HetNets.

Technology like SDN, NFV, Big Data and all-IP change the way networks are being constructed and managed. NGMN (2014) describe that these changes will enable the development of a highly flexible infrastructure that enables cost-efficient development of networks and associated services, as well as an increased pace of innovation.

2.3.2 The User Terminal

As the mobile networks will be improved, the user devices should be upgraded correspondingly to ensure that the improvements are fully utilized (Li et al., 2014). Hossain (2013) describes that the terminals should be dynamically upgraded and adapted to new situations, and designed to self-explanatory operate in different heterogeneous access networks. Javaid (2013) expects that in 5G, the user terminals will be more in focus than in the earlier generations. According to Janevski (2009), the 5G devices will have software defined radios and modulation schemes, as well as new error-control schemes that can be downloaded from the Internet. The terminals will have access to different wireless technologies at the same time, and they will be able to combine different flows from different technologies, as described earlier. The networks will handle user mobility, while the device will make the final choice for which access technology to use for a given service. The latter will be handled by open intelligent middleware in the devices.

When connection selection based on intelligent QoS management becomes a reality, the terminals will also be able to change the QoS related constraints during a single end-to-end connection, and thus also change the connection if necessary. As described by Sapakal (2013), it may seems like the initial Internet philosophy of keeping the network as simple as possible, and giving more functionalities to the end nodes, will become reality in 5G. Janevski (2009) further suggests that while the networks handle the user mobility, the terminals may also choose between different network providers for different services, in addition to different kinds of connections. The role of the terminals described in this section should be interpreted as a part of the proposed variants of CMaaS as mentioned in 2.2.1, to optimize the QoS (and potentially the QoE). NGMN (2014) describe the same idea, that the user application should be seamlessly and consistently connected to the RATs and or access points providing the best user experience without user intervention.

2.3.3 The Internet Protocol

It is a premise that 5G will have an All-IP architecture in the same way as the existing 4G systems (Javaid, 2013), and 5G must support IPv6 (Patel et al., 2012). According to Sharma (2013), the existing concept of different IP addresses is likely to be continued: Each terminal will have a permanent home IP address, and an addition care-of address (CoA), where the latter represents the actual location of the device. Incoming communication to the device will be handled in the following manner: First, a packet is sent to the home address. Then, the server on the home address will forward the packet to the actual location. The server will also send a packet back to the transmitter to inform about the current care-of address, such that future packets will be sent directly to the device. Patel et al. (2012) describe flat IP architecture as a key enabler to make 5G acceptable for all kinds of technologies. Flat IP architecture makes it possible to identify devices by symbolic names, unlike the hierarchical architecture of the conventional IP format. The main benefits of this architecture are lower costs, reduced system latency and decoupled radio access and core network evolution. Patel et al. (2012) also point out that security is the negative side of IP, as attacks are a challenge both today and for the future mobile networks.

Further, Patel et al. (2012) expect that we will see the technique multihoming in 5G. This is used to increase the reliability and performance or to reduce costs of the Internet connection for an IP network. In multihoming, the host is connected to more than one network, i.e. to several operators. The network configuration assigns multiple IP addresses to the different available wireless technologies. Then, if one link fails, the other IP addresses will still work and provide Internet access. The topic of vertical handovers is related to the discussion of how devices in the 5G era will support different access technologies simultaneously and handle the possibility to switch between them. While Patel et al. (2012) think that 5G will support vertical handovers, Janevski (2009) explains that they should be avoided, as they are infeasible in cases with many different technologies, operators and service providers. Yazici et al. (2014) introduce a unified approach to mobility, handoff and routing management for 5G.

2.3.4 Heterogeneous Networks

Although LTE achieved large gains in lowering the expenditures for mobile network operators and increase the end user experience with higher data rates and lower latencies, Yazici et al. (2014) argue that fresh thinking is needed toward 5G, mainly due to the enormous demand in wireless data. This drives the paradigm of heterogeneous networks (HetNets). HetNets are network deployments with different types of network nodes, equipped with different transmission powers and data processing capabilities, that support different radio access technologies (RATs) and are supported by different types of backhaul links (Li et al., 2014). The 5G HetNets will consist of macrocells, traditional microcells or picocells, a large number of new local IMT small cells and relay nodes or other low-power nodes. Small cells have been proposed as the key to achieve the 5G requirements for traffic volume, frequency efficiency and reduction of energy and costs (Chen and Zhao, 2014), and they support aggressive spectrum spatial reuse (Bangerter et al., 2014). The smalls cells are believed to play a very important role in 5G. These cells are likely to be designed differently than the cells in 4G (Chen and Zhao, 2014).

In addition to small cells, cloud-based Radio Access Networks (RAN-as-a-Service, RANaaS), device-to-device communication and multiple radio access technologies are considered as important components for the evolution of the proposed architecture of HetNets (Bangerter et al., 2014). RANaaS is further described by Rost et al. (2014), who have dedicated an entire paper to this field of research, as shown in 2.3.13. Chen and Zhao (2014) specify the need for heterogeneous layer coordination, to ensure that the different layers in the 5G HetNets are coordinated properly. If the heterogeneous layer coordination is handled well, it may give benefits in terms of system performance, energy savings and service provision. HetNets are among the most promising low-cost approaches to meet the industry's capacity growth needs and deliver a uniform connectivity experience (Bangerter et al., 2014).

Another proposed new cell concept to be included in the 5G HetNets is the mobile femtocell (MFemtocell) described by Wang et al. (2014), which is a combination of the mobile relay and femtocell concepts. This concept is meant to accommodate high-mobility users, i.e. traveling users. MFemtocells are located inside vehicles to communicate with other users inside the vehicle, while large antenna arrays are located outside the vehicle to communicate with outdoor base stations. An MFemtocell and its associated users are seen as a single unit to the base station, and from the user point of view, an MFemtocell is seen as a regular base station. MFemtocell can potentially improve the spectral efficiency of the entire network (Wang et al., 2014).

The intelligence aspects of 5G, and especially for HetNets, is investigated by Demestichas et al. (2013). They provide an overview for delivering intelligence toward 5G, by taking intro account the complex context of operation and essential requirements such as QoE, energy efficiency, cost efficiency and resource efficiency. Intelligence shall provide energy- and cost-efficient solutions, at which a certain quality provision is achieved. They describe that cognitive, intelligent management mechanisms are necessary to efficiently control and manage all these networked devices in the future.

2.3.5 Device-to-Device Communication

According to Nokia (2016) and Chen and Zhao (2014), direct device-to-device (D2D) communication will be an important method of communication in 5G. In addition to the mentioned potential for reduced latency, reduced power consumption and increased peak data rates, D2D may also contribute to the requirement of high reliability. Bangerter et al. (2014) add increased area spectral efficiency and improved cellular coverage to the list of key benefits. D2D leads to dense spectrum reuse (Bhushan et al., 2014). Chen and Zhao (2014) envision that cellular D2D will be integrated into the existing cellular networks, as a supplemental part of the systems. They also point out that relaying technologies is a relevant topic of research for D2D communication in 5G. Nokia (2016) describes that low latency is achieved through direct communication between devices over short distances with minimal propagation delay, and no involvement of network elements for processing the traffic data (i.e. no processing in the user-plane by the network elements), or any transport network that introduces delay. Further, to increase reliability or to extend network coverage, an additional D2D link can be used as a diversity path (Nokia, 2016). A D2D ad-hoc network where the devices serve as relays can be used as a backup solution, if the network infrastructure becomes unavailable or there are other failures in the network.

In their paper *Device-to-Device Communication in 5G Cellular Networks: Challenges, Solutions and Future Directions*, Tehrani et al. (2014) envision a two-tier cellular network. This involves a macrocell tier for conventional communication through base stations, and a dedicated device tier to handle device-to-device (D2D) communication. Hence, D2D communication will serve as another cell tier in the 5G HetNet (Bangerter et al., 2014). Tehrani et al. (2014) explain that the growing demand for higher data rates and capacity requires unconventional thinking for the 5G systems, and the proposed device tier could be a possible answer. The device tier in a two-tier cellular network will be an example of cooperative communications, which is a new class of wireless communication techniques. To realize the full potential of cooperative communication, device relaying must be implemented. With this technology, the user devices could serve as transmission relays for each other instead of having all communication in the licensed cellular bandwidth going through the base stations. In this way, the base stations would no longer be a bottleneck in the communication systems.

D2D communication has not been considered for the first four network generations, mainly because its potential seemed to be a cost reduction of local services provision, which never was a large cost in the past. This attitude has changed, and D2D functionality now seems much more beneficial, as it is widely believed that D2D communication can facilitate effective sharing of resources, which is a key issue for 5G.

Both Chávez-Santiago et al. (2015) and Tehrani et al. (2014) describe D2D as an underlay to cellular networks in licensed frequency bands. The most important difference between WiFi or Bluetooth and the proposed D2D functionality will then be that the first two work in unlicensed band, where interference cannot be controlled. Unlike D2D, WiFi and Bluetooth cannot provide any security and QoS guarantee. Mobile ad-hoc networks also operate similarly as D2D, but in the unlicensed spectrum (Chávez-Santiago et al., 2015). However, Bangerter et al. (2014) suggest that D2D technology can either reuse the same spectrum as the macrocell tier, or use unlicensed spectrum.

Tehrani et al. (2014) propose four main types of device-tier communications:

- DR-OC: Device relaying with operator controlled link establishment.
- DC-OC: Direct D2D communication with operator controlled link establishment.
- DR-DC: Device relaying with device controlled link establishment.
- DC-DC: Direct D2D communication with device controlled link establishment.

For DR-OC and DC-OC, the operator assists the link establishment, i.e. the base station is involved in the initial communication, while the devices handles this themselves in the latter two cases. DR-OC and DR-DC are about device relaying, where several devices could used as transmission relays between the sending and the receiving devices. DC-OC and DC-DC on the other hand, is simply about direct communication between the sending and receiving devices, i.e. no device relays between them. The different device-tier communication possibilities are shown in figure 2.1, where the gray lines to the base stations show communication in the macrocell tier and the black lines show direct communication between the devices in the device tier. In the same manner as Tehrani et al. (2014), Nokia (2016) separates between two different approaches for the radio resource management in D2D: centralized and distributed resource allocation. With the distributed solution, the devices can transmit immediately, and there are no association procedure or dedicated control channel. On the other hand, the centralized solution will effectively avoid packet collisions, but at the expense of increased complexity and average delay.

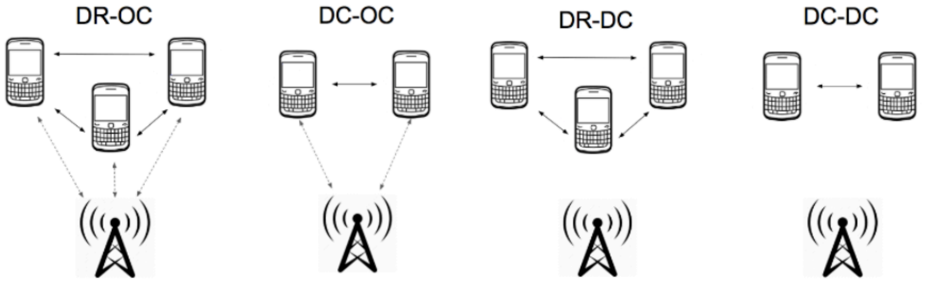

Figure 2.1: An overview of the four different device-tier communications proposed by Tehrani et al. (2014).

An important challenge with the proposed D2D communication, is security. Since user data is routed through the devices of other users, privacy must be maintained. This could be solved with closed access, where each device keeps a list of trusted devices which will be the only devices it can communicate with in the device layer. Another challenge is interference management, which should be handled in a smart way to have the least possible impact on the performance of the macrocell tier (Tehrani et al., 2014).

2.3.6 Software-Defined Networking and Network Function Virtualization

Software-Defined Networking (SDN) and Network Function Virtualization (NFV) are two emerging network architecture concepts based on cloud technology, and they are often mentioned in the context of 5G. The main connection between these two concepts is that NFV is a complementary technology to SDN, as NFV can provide the network infrastructure on which SDN can run (Demestichas et al., 2013). Also, SDN will enable better NFV opportunities and network programmability (Yazici et al., 2014). As both concepts are complex, the technology behind will not be explained in detail here. The key takeaway point is that both concepts provide flexible network deployment and operation, and are therefore relevant for 5G (Agyapong et al., 2014). Rost et al. (2014) agree that NFV and SDN should be leveraged for the RAN in 5G.

Demestichas et al. (2013) describe SDN as a network architecture with a centralized network controller in the control plane, which is responsible for allocating traffic to network elements in the separated data plane of the network. Denazis et al. (2015) further describe that the SDN approach allows network administrators to programmatically initialize, control, change, and manage network behavior dynamically via open interfaces and abstraction of lower-level functionality.

The brief definition of NFV is that the concept uses IT virtualization technologies to virtualize classes of network node functions into building blocks that may connect, or chain together, to create communication services. Demestichas et al. (2013) describe more thoroughly that NFV technology is a new way to build an end-to-end network infrastructure with evolving standard IT virtualization technology, to enable the consolidation of many heterogeneous network devices onto industry standard high-volume servers, switches and storage. The network function of a network device is implemented in a software package running in virtual machines. This makes it easier to introduce or test a new network function by simply installing or upgrading a software package, which is run by the servers.

According to Yazici et al. (2014), an all-SDN network architecture with hierarchical network control capabilities will give the opportunity for different grades of performance and complexity when core network services and service differentiation is provided. Yazici et al. (2014) further describe that if the cloudification opportunities in the form of NFV are pursued, the 5G challenges regarding expenditures, agility and flexibility will indeed be met. With NFV, mobile network functions can be moved from dedicated hardware platforms to virtual machines running on generic hardware. One example of NFV is virtualization and pooling of baseband processing in the base stations, more commonly referred to as the cloud radio access network (C-RAN). If the benefits of cloudification are utilized to take full advantage of the distributed transport capacity, there will be big differences between 4G and 5G in how the overall system control is designed and managed (Yazici et al., 2014). Demestichas et al. (2013) describe that the main benefit of NFV technologies is that carriers can build and operate a network with reduced power consumption and equipment costs. This is due to consolidation of equipment and exploitation of the economies of scale in IT. The carriers can also benefit from new services that are easy to create, test and deploy with a short time to market, as well as easy management of the network infrastructure and services. Demestichas et al. (2013) also mention reduced space costs as a possible benefit.

2.3.7 Cognitive Radio

Although many small cells will be deployed in 5G, increasing the radio spectrum is not enough to ensure flexible spectrum usage. A promising technology to deal with this issue is cognitive radio (CR) (Chen and Zhao, 2014). Hong et al. (2014) describe CR as a promising technology to cope with the challenges in 5G networks, especially the challenges regarding capacity issues and QoS requirements. CR is a technique that allows a cellular network to dynamically lease the under-utilized frequency bands, without causing harmful interference to the incumbents. Hence, cognitive cellular networks are cellular networks that employ CR to lease additional spectrum outside the licensed cellular bands. Hong et al. (2014) illustrate that the concept of cognitive cellular networks is promising for 5G systems. The cost of leasing spectrum is expected to be much lower than the cost of purchasing a licensed band, as the leased spectrum is fundamentally opportunistic and unreliable. CR allows an expansion of a cellular network, on demand, and at a relatively low cost. This is a promising solution to handle unpredictable amounts of random and diverse mobile data traffic.

Spectrum leasing bills and electricity bills are both envisioned to be significant parts of the OPEX of cognitive cellular networks. Spectrum leasing bills are related to bandwidth, or spectral efficiency (SE). Electricity bills are related to energy efficiency (EE). Spectrum and energy are natural resources that underpin all wireless systems, and SE and EE are related to the network constraints of reliability, delay and overhead. Hong et al. (2014) argue that understanding the EE-SE trade-off can provide direct guidelines to the OPEX management of cognitive cellular networks. Analysis of EE-SE trade-off can also give interesting insights into system design and optimization. Hong et al. (2014) show that insights from EE-SE analysis are useful in the provision of guidelines for cognitive radio resource management.

CR is also mentioned by Wang et al. (2014), who describes it as an innovative software-defined radio technique that may improve the spectrum utilization of the existing radio spectrum, as a large portion of the radio spectrum is under-utilized most of the time. Patel et al. (2012) agree that use of CR or software-defined radio (SDR) will result in more efficient radio communication systems in the future. Chen and Zhao (2014) also consider CR as a promising technology to obtain a network design that allows for flexible spectrum usage. Another term for CR is smart radio (Gohil et al., 2013).

2.3.8 The Radio Spectrum and Millimeter-Wave Frequency Bands

The traffic growth in mobile communications has drawn attention to the millimeter-wave frequency bands, as the available spectrum for cellular usage in the lower frequency bands has become scarce. Several authors have addressed this. There is a large amount of underutilized spectrum in the millimeter-wave bands, and they are a potential solution for achieving an enormous capacity increase in the cellular networks. According to Chen and Zhao (2014), new available radio spectrum for the future mobile network generation is mainly located above 3 GHz, or even higher. Yazici et al. (2014) describe that the carrier frequencies for 5G will be from the cellular bands used today, below 5 GHz, to millimeter-waves at 60 GHz and beyond. Bangerter et al. (2014) mention the exploration of higher carrier frequency, such as millimeter-wave bands from 30 to 300 GHz, as an important area for future studies regarding 5G. Li et al. (2014) assume that a 3-10 times increase in the size of allocated spectrum can be seen over the next decade if the millimeter-wave bands are utilized. They also mention that channel propagation characteristics are different in higher frequency bands, compared to bands below 3 GHz. This requires new design of the air-interface and the network architecture.

Roh et al. (2014) agree that the millimeter-wave bands are a potential candidate for 5G, as tens to hundreds times more capacity could be achieved with these bands, compared to the capacity of 4G. They conclude that wide bandwidth is the most effective and straightforward method to provide the foreseen data demands for 5G cellular services. Roh et al. (2014) show how the millimeter-wave frequencies exhibit themselves as strong candidates for 5G with recent channel measurement, simulation and prototype results. Most of their article is dedicated to presenting the results of their millimeter-wave prototype, which feature a large system bandwidth in excess of 500 MHz at 28 GHz, and supports tens of antennas placed in plannar arrays at both ends of the communications.

2.3.9 Network Densification

To realize the 5G vision, there is a need for very dense deployments and centralized processing. Rost et al. (2014) discuss how these requirements shall be met, by utilizing cloud technologies and flexible functionality assignment in radio access networks to enable densification of the networks and centralized operation of the radio access network over heterogeneous backhaul networks. As most 3.5G and 4G mobile networks are based on the 3GPP standards, Rost et al. (2014) have based their network architecture and radio access suggestions on an evolution of the current LTE technology.

Bhushan et al. (2014) explore network densification over space and frequency for wireless evolution into 5G. Network densification means adding more cells, and is a combination of the two components spatial densification and spectral aggregation. Spatial densification is realized by increasing the number of antennas per node and increasing the density of base stations deployed. Nodes are user devices and base stations. Spectral aggregation involves using larger amounts of the electromagnetic spectrum. This spans all the way from 500 MHz and into the millimeter wavebands (30-300 GHz). Network densification gives higher capacity and improved coverage. Spatial densification and spectral aggregation is of little consequence unless they are complemented by densification of the backhauls, that connects the base stations to the core network. Densification of the backhauls is expensive, and because of this, the cost of network densification might not be sustainable. However, Bhushan et al. (2014) conclude that network densification the way they describe it will meet the challenge of a thousand-fold capacity increase, along with significant reduction in cost-per-bit delivered. This ensures the business viability of the proposed approach.

2.3.10 Advanced MIMO

MIMO stands for multiple-input and multiple-output, and is a method where the capacity of a radio link is multiplied, by the use of multiple antennas at both the transmitter and receiver, to exploit multipath propagation. MIMO techniques have been an important element in Wi-Fi, 3G and 4G. However, the research consensus is that its full potential is not yet realized. Bangerter et al. (2014) describe that advanced MIMO techniques are at the heart of achieving higher capacity for the future cellular systems, while Chen and Zhao (2014) add significantly improved system performance in coverage and user data rates as other benefits. According to Wang et al. (2014), MIMO methods can further improve the reliability, SE and EE in the network.

MIMO is often mentioned with different prefixes in the literature. The most used ones are advanced, MU and massive. Advanced MIMO is an umbrella term that covers different evolutionary MIMO techniques. One such technique is Multi-User MIMO (MU-MIMO), a MIMO technology where a set of users communicate with each other, and each node has one or more antennas. The opposite case, singe-user MIMO, is when a single transmitter communicates with a single receiver, both with several antennas. MU-MIMO offers increased multiplexing gains, and is suggested for 5G (Bangerter et al., 2014). Chen and Zhao (2014) describe MU-MIMO as the perfect solution for the challenge in asymmetric traffic, i.e. the ratio between uplink and downlink traffic. Massive MIMO is one of the techniques related to MU-MIMO. Wang et al. (2014) describe massive MIMO as a MIMO technique where the transmitter and/or receiver are equipped with a *large* number of antenna elements, typically tens or hundreds, hence the alternative expression: large-scale antenna systems. Massive MIMO is considered a promising candidate for 5G, as it minimizes interference while it maximizes the beamforming gain.

Wang et al. (2014) also suggest spatial modulation, which is a novel MIMO technique. This is a technique that can increase the data rates. Spatial modulation encodes part of the data to be transmitted onto the spatial position of each transmit antenna in the antenna array, instead of simultaneously transmitting multiple data streams from the available antennas. According to Patel et al. (2012), the higher throughput and reliability associated with MIMO systems is more beneficial for the base stations than the users, because of size and power consumption. They suggest group cooperative relay techniques as an alternative to MIMO systems.

2.3.11 Indoor Scenarios and UDRANETs

An important goal of 3G and 4G was to achieve constant coverage for the same services in both outdoor and indoor scenarios. According to Chen and Zhao (2014), 5G will be a heterogeneous framework, and backward compatibility will not be mandatory both indoors and outdoors. The improvement of user equipment is expected to provide the ability to support simultaneous connections, both indoors and outdoors. The previous barriers will be broken and there will occur a fundamental change in 5G.

Wireless users are located indoors 80 % of the time, and 60 % voice traffic and 70% data traffic happens indoors (Chen and Zhao, 2014). This may increase in the future, and there is a need for better indoor coverage. One reason for the indoor coverage issues is the fact that many modern office buildings are built in materials such as steel, different metals and even modern glass that contain metal. Unfortunately, these materials make the buildings work like so-called Faraday cages, where radio signals are effectively blocked (Seehusen, 2013). To solve the issue of indoor coverage, Wang et al. (2014) propose a potential cellular architecture for 5G that separates indoor and outdoor scenarios. This architecture should be assisted by distributed antenna system and massive MIMO technology (2.3.10).

Chávez-Santiago et al. (2015) also address the challenge of indoor coverage, and describe the proposed new concept of Ultra-Dense Radio Access Networks (UDRANETs). This technology is envisaged as low-power access nodes located a few meters apart for indoor areas. The main goal of UDRANETs is to provide an extremely high traffic capacity over highly-reliable short-range links (Chávez-Santiago et al., 2015). UDRANETs are likely to operate in the frequency range from 10 to 100 GHz.

2.3.12 Beam Division Multiple Access

Beam Division Multiple Access (BDMA) is a novel multiple access technique for wireless communication that was first suggested in 2008, by the South Korean research and development program *5G Mobile Communication Systems based on Beam-Division Multiple Access and Relays with Group Cooperation* (Patel et al., 2012). Limitations in available frequency and time resources causes communication challenges, and there is a demand for a technique using other resources to increase the system capacity. This have resulted in the BDMA concept. When a base station is communicating with mobile stations, an orthogonal beam is allocated to each mobile station. BDMA divides an antenna beam according to locations of the mobile stations to allow them to give multiple accesses, thereby significantly increasing the capacity of the system (Patil et al., 2012).

2.3.13 Radio Access Network-as-a-Service

Rost et al. (2014) describe the Radio Access Network-as-a-Service (RANaaS) or cloud-based RAN concept, which partially centralizes functionalities of the RAN depending on the actual needs, as well as network characteristics. The proposed concept is a trade-off between full centralization and the existing decentralization of today's networks, as the centralization of processing and management in 5G networks need to be flexible and adapted to the actual service requirements. The proposed RANaaS concept will affect the architecture of mobile networks, but for economic reasons, the 5G architecture is expected to be developed as an evolution of LTE. Rost et al. (2014) conclude that RANaaS is a key enabler for the future 5G networks.

2.3.14 The Core Network Architecture

Both Sharma (2013) and Patel et al. (2012) assume that the 5G core will be a reconfigurable, multi-technology core. They call it the SuperCore, while Hossain (2013) uses the term MasterCore and Patil et al. (2012) describe it as the NanoCore. The 5G core is likely to be a convergence of different technologies. According to Hossain (2013), the hardware and software should be upgradeable, and hence adapted to new situations. The upgradeability will be based on CR technology as described in 2.3.7. The proposed MasterCore should be able to change its communication functions based on factors such as network status or user demands. Hossain (2013) further suggests the use of a 5G Interfacing Unit between new deployments and the core network, for the purpose of easy network management. Advantages will be lower costs to establish networks, improved network efficiency and reduced complexity. The key takeaway point from Hossain (2013) is that the 5G core network is likely to be of a highly interdisciplinary character.

The NanoCore suggested by Patil et al. (2012) is a convergence of nanotechnology, cloud computing and an All IP platform. The main segments in cloud computing will be applications, platforms and infrastructure. These segments should be utilized to satisfy all the customer demands for 5G. Cloud computing will contribute to lower the CAPEX of 5G network deployment, which may create decreased billing to the end users for all utilized services (Patil et al., 2012). The flat IP concept described in 2.3.3 will make it easier for different RANs to upgrade into a single core network. Nanotechnology can be used as a defensive tool against the security issues related to flat IP.

Sharma (2013) explains that all network operators will be connected through one single core with one single infrastructure, regardless of the access technologies of the operators. The same idea is described by Patel et al. (2012), who envision one "super" core with massive capacity. Such a core would integrate the unique standards of different engineering practices.

2.3.15 Technology for Mission Critical Communications

In the white paper *5G for Mission Critical Communication: Achieve ultra-reliability and virtual zero latency*, Nokia (2016) has looked at technological solutions for the defined requirements. To achieve ultra-reliability, they propose massive MIMO and multi-connectivity, as well as advanced interference management and user/service-optimized retransmission mechanisms. The radio latency may be brought down to milliseconds if flexible frame structure is introduced. To ensure a flexible and reliable network, 5G needs a programmable multi-service architecture. The suggested key components for this architecture are network slicing, programmable networks, network resiliency and mobile-edge computing.

Nokia (2016) explain that research on improved radio access is important, as it is close to the user and has a significant impact on reliability and latency. Diversity and interference management should be included for high reliability. This is because the quality of the radio link between base station and the user terminal directly affects the overall reliability of the system. Signal-to-Interference-and-Noise Ratio (SINR) is used to measure the link quality. Higher SINR gives lower packet error probability, and hence higher reliability. A technique to improve SINR by interference management is to reduce the received interference from neighboring base stations or terminals. To cancel the one or two strongest interferers is usually sufficient to achieve most of the potential gain in SINR.

Massive MIMO with hundreds or thousands of antennas is recommended for 5G, to exploit the principle that more antennas give better coverage or space availability. An increase in the number of antennas gives better signal and higher SINR. Nokia (2016) further explains that the target for 5G must be virtually zero interruption time. This necessitates the make-before-break principle at every handover, to ensure that the connectivity is always available.

2.4 The Green Perspective

Olsson et al. (2013) have looked at the green aspects of 5G, as it is a big challenge to provide the necessary capacity increase to billions of devices in an affordable and sustainable way. The energy bill of the mobile operator is already an increasing part of their operating expenses (OPEX), and if nothing is done, there is a risk that it may increase even more. The power reduction is not only important from an economic perspective, but also from a sustainability perspective. Olsson et al. (2013) describe that low energy consumption is the key to achieve this. However, the minimization of network power consumption should not affect efficient QoS management in the network. In their article, they have tried to pinpoint some important focus areas and potential solutions for the design of an energy efficient 5G mobile network architecture. For the system architecture, it is suggested a separation of the data and control planes. For network deployment, heterogeneous ultra dense layouts are expected to have a positive effect. Regarding radio transmission, introduction of massive antenna configurations could be an important enabler. Further, there is a need for more energy efficient backhauling solutions than today (Olsson et al., 2013).

Chen and Zhao (2014) agree upon the challenge of the need for a future reduction of the increased energy consumption in the base stations. The energy consumption for future networks may need to be reduced on the order of several magnitudes. One way to achieve this, could be to offload majority data to local small cells, and letting inactive cells sleep. Further, as the deployment of advanced wireless technologies comes at the cost of high energy consumption, increased CO_2 emission is caused. This is a major threat for the environment. To minimize the energy consumption further, Wang et al. (2014) present separate indoor communication systems, visible light communication and millimeter-wave technologies as possible solutions to include in the design of 5G wireless systems. Finally, Agyapong et al. (2014) have also identified the need for intelligent energy management techniques, especially in the RAN. They describe that NFV and SDN can enable further cost reductions.

2.5 Economics

In this section, the economic aspects of 5G described in the literature are presented. First, general comments on costs and revenues will be presented, before two papers dedicated to certain economic aspects of 5G is described more thoroughly.

2.5.1 Costs

Several authors have addressed the cost issues for future mobile networks. Osseiran et al. (2014) explain that while 5G must satisfy the specified requirements, the growing cost pressure must also be addressed. According to Yazici et al. (2014), revenue source loss to service providers and increasing CAPEX and OPEX are some of the biggest challenges the mobile operators currently face, in addition to the general capacity challenge. Olsson et al. (2013) describe that the CAPEX and OPEX should to be at a level where services can be provided at a reasonable end user price, and with attractive business cases for the mobile operators. In addition to the factors that already drives the network costs up, Agyapong et al. (2014) describe that some new improvements also may be expensive. One such improvement is the optimization of network usage for fair and better QoE, which includes an integration of more functionality, such as deep packet inspection, caching and transcoding.

To lower the costs, Demestichas et al. (2013) have faith in the proposed cloud concepts, as it is a rationale that there can be significant savings in both CAPEX and OPEX if wireless networks are based on cloud principles. Bangerter et al. (2014) describe that a shift towards software-based implementations and virtualization technologies is required, as the cost and flexibility of deployment will be important factors in 5G networks. In other words, the economic motivation seems to be an important factor behind the focus on cloud concepts like SDN and NFV in 5G. Another cost reduction in 5G is that the cost of a small cell, including the cost of maintenance of cells, will be significantly reduced for 5G (Chen and Zhao, 2014). Reduction in cell costs is also described by Bhushan et al. (2014), who point out that picocells will enable much lower CAPEX and OPEX than macrocells. Bangerter et al. (2014) explain that small cells and D2D communication together will form a new underlay tier of low-cost infrastructure that complements the coverage and costs of conventional cellular networks.

Chávez-Santiago et al. (2015) have looked at the societal perspective, regarding mobile broadband access in rural areas. 5G does indeed have the potential to improve this. The complete coverage of rural areas have been delayed due to high CAPEX for deploying a large number of base stations, and low average revenue per user at the same time. Chávez-Santiago et al. (2015) describe that fewer base stations will be needed if TV white space technology and traffic offloading solutions are utilized. This gives better propagation conditions in the VHF/UHF spectrum, and hence 5G networks could be deployed in rural areas at a lower cost. Patel et al. (2012) also suggest that future base stations should be designed to cover bigger regions. In this way, fewer base stations will be needed, and some may be shut down.

Agyapong et al. (2014) also consider socio-economic development, for which connectivity is an important enabler. They point out that to make connectivity a universally available, affordable and sustainable utility, it is important to reduce the infrastructure costs as well as the costs associated with their deployment, maintenance, management and operation. 5G faces the challenge that huge improvements are needed to address the new requirements, while the customers unfortunately are unwilling to pay proportionally. To solve the capacity and data rate challenges with network densification (2.3.9) could be expensive in terms of equipment, maintenance and operations. The equipment costs could be reduced by minimizing the number of functionalities at the base stations. One suggestion is to only implement some functionalities in the regular base stations, and move functionalities of higher layers. This results in simpler base stations, and these could be deployed by users, and remotely or autonomously managed. In this way, deployment and operations costs could be reduced (Agyapong et al., 2014).

2.5.2 Revenues

Minimizing latency and increasing reliability opens up potentially lucrative new business opportunities for the industry, arising from new applications that simply will not work properly if network delays are too high (Nokia, 2016). As 5G will give the user a new and improved experience of wireless connectivity, simple network economics require that the mobile industry comes up with a strategy that enables new services, new applications and new opportunities to monetize the user experience (Bangerter et al., 2014). As 5G will provide a new level of connectivity, the new opportunities will not only affect the telecommunication industry and Internet actors, but also other industries such as automotive, healthcare, manufacturing and logistics, and even the government/public sector, as 5G is expected to cover use cases for these industries (Nokia, 2016).

According to Yazici et al. (2014), the proposed CMaaS concept may open up new paths for revenue generation, as it allows the control of mobility and routing of different flows or users differently in the network. Agyapong et al. (2014) suggest that pricing schemes can be used to manage and potentially reduce the increase in data consumption, which has already been demonstrated by the operators in the market. However, customers are more willing to pay for the provisioned service than the data volume, and hence pricing models may not be effective to suppress traffic in future (Agyapong et al., 2014).

2.5.3 D2D Pricing

Tehrani et al. (2014) identify pricing for D2D services as an important dilemma the operators will have to face, i.e. how to control and charge for D2D services. For the cases of DR-DC and DC-DC communication as described in 2.3.5, the operator will not have any control as the initial setup does not go through the base stations. Hence, they should not expect to make any profit from it. For the other types of D2D communication, DR-OC and DC-OC, where initial setup goes through the base stations, Tehrani et al. (2014) propose pricing schemes based on auction theory and game theory.

For DR-OC, the main challenge is to provide sufficient incentives for the users to be willing to share their own resources for the relay purpose. Tehrani et al. (2014) mention discounts on monthly bills based on the amount of data relayed through the user's device or free services as possible incentives. This is reasonable from the operator's point of view, as device relaying is beneficial for them, as it contributes to offloading the amount of traffic in the macrocell tier, and hence offload the base stations. By defining an utility function and plotting operator revenue as a function of average signal-to-noise ratio (SNR), Tehrani et al. (2014) have been able to show that for a two-tier network with DR-OC, both the operator revenue and the device revenue will increase compared to a conventional single-tier system.

For DC-OC, it is an important challenge to have tempting price schemes so that the users will use this instead of Bluetooth or WiFi. Tehrani et al. (2014) describe that the design of a pricing model can be viewed as a spectrum trading problem, where the seller wants to maximize profit, while the buyer wants to maximize his or hers utility of spectrum usage. To examine the DC-OC case numerically, they have plotted operator revenue as a function of the number of devices. They have found that the operator revenue in the two-tier network is 10 % higher than in a single-tier system. The gain in device revenues in a two-tier network is 48 % for 100 devices, and 6 % when the number of devices is 800.

Tehrani et al. (2014) conclude that DR-OC and DC-DC should bring significant gains for both the operator and the users. One shortcoming in this research, is that they assume a monopoly market (only one operator present), hence they have not covered the aspect of competition between several operators.

2.5.4 Cost-Effectiveness Assessment

Nikolikj and Janevski (2015) explain that the lack of microwave spectrum have forced recent publications to focus on millimeter-wave systems, to utilize the available spectrum in the bands from 30-300 GHz. The performance in terms of capacity and user data rates has motivated the evaluation of the financial ability of the proposed millimeter-wave systems integrated in the future Advanced Wireless Access Heterogeneous Network (AWA-HetNet). The objective is to see if this is a potential solution to bridge the revenue gap experienced by the mobile network operators.

Nikolikj and Janevski (2015) use techno-economic analysis to provide a model to assess the cost-effectiveness of the AWA-HetNets. They present comparative cost-capacity modeling, calculation of production costs, and business profitability. Nikolikj and Janevski (2015) consider an extremely high populated area from the future, with 250,000 citizens per km^2. This is more than 6 times the current highest population density in the world which is found at Manila, the capital of the Philippines, with 42,000 citizens per km^2.

Nikolikj and Janevski (2015) model the total cost of AWA-HetNets (C_{TOT}) as:

$$C_{TOT} = N_i \sum_{i \epsilon \phi} \sum_{k=0}^{(K-1)} \frac{\alpha_{k,i}}{(1+\beta)^k}, \tag{2.1}$$

where $\alpha_{k,i}$ is the sum of expenditures, in terms of CAPEX and OPEX occurred in year k for a base station or access point of type i, β is the yearly discount rate, N_i is the total number of base stations or access point of type i that is required, and ϕ is the set of all the available base station or access point configurations.

Various cost-efficient capacity enlargement strategies are also proposed. Their cost structure modeling is based on limiting the CAPEX and OPEX of the radio access network. The business profitability of AWA-HetNet is assessed by relating the total investment costs to the production costs and revenues. They have found that the 5G millimeter wave systems can ensure the network operators a sustainable profit, with more than 50 % profit margins.

Regarding the known issue of indoor coverage in modern office buildings in described in 2.3.11, Nikolikj and Janevski (2015) have looked at the cost-capacity and profitability of radio access technologies for office areas. Their result indicates that the proposed 5G picocell base stations could be a preferable deployment solution for heavily loaded office environments. In particular, they have assessed the profit margin of deployment needed to satisfy the demand known as the virtual reality office, where the requirement is 36.0 TB per user, per month.

Further, Nikolikj and Janevski (2015) also analyze cooperative resource sharing, by the application of Radio Resource Management (RRM), where radio resources are shared between mobile operators. This is for the purpose of maximizing the aggregate performance, in terms of capacity or profit. In particular, they study the cases of capacity over-provisioning, and they also determine principles to provide guaranteed data rates (in terms of QoS) to a number of individual users. Capacity over-provisioning is expected to occur often in the 5G AWA-HetNets, due to the huge amount of available bandwidth. Although the RRM approach could provide a perfect sharing of the available capacity, the operators will pay a higher cost for the capacity allocation, than the value added. This is due to the reduced capacity. Because of this, Nikolikj and Janevski (2015) have also looked at the use of RRM with priority queuing dynamically adjusted to the load of the mobile network operators. In this case, Nikolikj and Janevski (2015) have been able to show capacity improvements of 31 % when two operators handle comparable loads, compared to RRM with fixed resource sharing. This is a significant improvement.

2.6 Business Aspects

NGMN (2014) takes a closer look at the business aspects of 5G, and expects the business context beyond 2020 to be notably different from today. This is mainly due to the emergence of new use cases and business models, which are driven by the needs of the customers and operators, and enabled by the key technologies for 5G. Business orientation and economic incentives with foundational shift in cost, energy and operational efficiency should make 5G feasible and sustainable. Their 5G vision is that “5G is an end-to-end ecosystem to enable a fully mobile and connected society. It empowers value creation towards customers and partners, through existing and emerging use cases, delivered with consistent experience, and enabled by sustainable business models”.

Regarding the business models for 5G, NGMN (2014) expects that these will expand to new ones, and support different types of both customers and partnerships. They define customers as not only consumers and enterprises as of today, but also verticals and other partnerships. Regarding partnerships, they describe that mobile network operators will continue to develop own services, but also expand their business reach through partnerships for both the infrastructure and the application development aspects. Network operators have already started to leverage partnerships with service providers to deliver packaged services to end users. In the future, service providers are expected to continue delivering applications that require higher quality, lower latency and other service anhancing capabilities such as proximity, location, QoS and authentication, on demand, and in a highly flexible and programmable way. Topics regarding cooperation between different actors in the ecosystem of telecommunication, such as paid peering and vertical integration, will indeed be important in the 5G era. These topics address the challenge of how the operators can ensure revenues sufficient to pay for the necessary infrastructure in the future, when the largest profit is located a different place in the value chain, i.e. at the service providers, who often receive high revenues and have much less expenses. NGMN (2014) also expects that operators will contribute to the mobilization of industries and processes.

NGMN (2014) describe 5Gs ability to support the business models at a very low cost and provide the flexibility to enable the business models on demand, as a key requirement. Regarding value creation, they propose that the following service characteristics should be offered across all services and customer segments. 5G services should be (NGMN, 2014):

- Available anywhere and anytime.
- Delivered with consistent experience across time and space.
- Accessible on multiple devices and access technologies.
- Able to support multiple types of interaction.
- Delivered seamlessly and transparently across access technologies.
- Delivered in a contextual and personalized fashion.
- Enabled by secure and trusted communications.
- Supported by a highly reliable and resilient network.
- Delivered in a responsive and real-time fashion.

A global business model evolution of mobile operators' services will include the evolution of current services, as well as the emergence of new ones. According to Hossain (2013), the proposed upgradeability for the 5G networks could make it easier for network operators to introduce value-added services at a higher pace in the future.

2.6.1 Market adoption of 5G

Demestichas et al. (2013) have looked at how the market is likely to adopt 5G. They present two different approaches: gradual and disruptive approach. The gradual approach will be characterized by introducing the intelligence related aspects first. The disruptive approach is more likely to be considered in the long term. When sufficient profits and return on investment is achieved by the mature 4G systems, operators will be more willing to invest in radically updated equipment. Chávez-Santiago et al. (2015) also describe two simultaneous phases of the evolution towards 5G. These are the enhancement of the current networks, and the integration of evolved networks with complementary wireless communication systems. Nokia (2016) describes that in the first phase of deployment, 5G radio access with diversity, multi-connectivity, interference management and flexible frame structure will be introduced and integrated with the LTE core network. At the same time, network slicing, programmable networks, network resiliency and mobile-edge computing will be introduced by a programmable 5G multi-service architecture. In the next phase, 5G radio access will be integrated with the 5G core network, without the need for an LTE anchor. In this phase, the 5G requirements regarding low latency and high mobility and reliability will be met.

Chapter 3

Discussion

In this chapter, the literature defining 5G as presented in chapter 2 will be discussed. The main part of the discussion will be focused on the identification of different economic aspects of 5G. However, first we summarize the presented technologies.

3.1 The Technology Aspect

As described in the beginning of chapter 2, the use cases and requirements for 5G is clear, but 5G is not standardized yet. The 3GPP expects to release the first set of 5G standards within the next year. Therefore, most of the published 5G literature is concerned with the technological aspects, i.e. proposing enabling technologies and solutions to be included in the next generation of mobile networks, so that 5G can meet the requirements described in section 2.2. Although the requirements are high, they seem considered to be fully reachable. The technology research are approaching a consensus for some of the most important solutions for 5G, while others are not formally decided yet. The experts agree that 5G will not be an entirely new communication system built from scratch, but rather an evolution of the 4G systems, with some new components.

The role of the user terminal, or smartphone, as the main personal device is clear. It is also clear that in 5G, the number of personal devices will increase. Wearable devices with AI capabilities is one example of new kinds of personal devices. 5G must enable the IoT with a much higher number of connected devices than today, and therefore it must support machine-type devices in an efficient manner.

The technologies that are most often mentioned in the context of 5G, are HetNets with a dedicated tier for D2D communication, SDR, SDN, NFV, CR, massive MIMO systems and use of the millimeter-wave frequencies. These are regarded as some of the most important key enablers for 5G. Network densification, UDRANETs and how to meet the demand for indoor communication, the architecture of the core network, the BDMA technique and the cloud-based RAN concept, RANaaS, are other parts of the proposed 5G technology that is often mentioned in the 5G context. Therefore, we have chosen to present all these technology topics in chapter 2. However, as the focus of this thesis is the economic side of 5G, the mentioned technologies will not be further evaluated here, as this is beyond the author's field of expertise. Based on the literature review, we conclude that the research consensus seems to be that all these technologies are likely to be included in 5G.

3.2 The Economic Aspects

Until now, we have tried to answer the first part of the problem description of this thesis, i.e. explore how 5G currently is defined. In this section, we look at the next objective: identify economic aspects of relevance. Initially, the outline for this thesis was to study certain economic aspects deeper. However, during the literature review, it became clear that it has not been published much literature dedicated entirely to the economic sides of 5G. The main reasons for this are already described: 5G is not standardized yet, and the current focus area is the technology to be included. In other words, it is still early to look at economic aspects isolated. However, it is fully possible to *identify* important economic aspects of 5G, and that is the starting point for the following discussion here. Suggestions for further research will be summarized in section 5.2.

It should be pinpointed that although there are few papers dedicated to economic research on 5G, it does not mean that economic aspects have not been much considered for 5G yet. This perception arise from the fact that several papers mentions economic aspects briefly, without further elaboration. Economic considerations rarely occurs in abstracts or conclusions, which means that the publications often must be read thoroughly to find these considerations. Costs are one such example. As shown in section 2.5.1, many technically oriented authors agree that the issue of increasing expenses must be faced. Osseiran et al. (2014); Yazici et al. (2014); Demestichas et al. (2013); Chávez-Santiago et al. (2015); Agyapong et al. (2014) have formulated requirements for this, as described in section 2.2. It is highlighted by Bangerter et al. (2014) that the strategy for the capacity and performance growth of 5G must be economically sustainable. Yazici et al. (2014) describe that LTE did achieve significant gains in lowering the mobile network operators' expenses, but this gain is not sufficient for the future mobile networks, and future gains must be achieved in new and different ways. Although it is of interest to minimize the CAPEX and OPEX of cellular networks, the main constraint is that the actual 5G requirements in terms of network performance must be met. There will always be a trade-off between cost and performance.

Olsson et al. (2013) reflect upon the growing cost pressure from both the business perspective, i.e. the necessity of offering affordable end-user prices, and the green perspective, in terms of power consumption, sustainability and CO_2 emissions. The green perspective is related to economics in several ways. The easiest connection is that less power consumption is a green objective, that also leads to lower utility expenses, which are a part of the OPEX. A more indirect connection that is more clear in the long term, is the economic consequences of emissions from power generation, such as CO_2. Large emissions may be charged with fees, which are added to the total costs. And even if there are no fees present, large emissions may still have socioeconomic consequences, that reduces the social welfare or the social surplus. These are some of the reasons for why the green perspective should be included when economic aspects of 5G are studied. In addition, it should be mentioned that the focus on decreased power consumption is not only focused on the networks, but also on the user terminals.

Technologies that are described to have a positive impact on the mobile network operators' expenses in 5G, are cloud concepts, such as SDN and NFV (Demestichas et al., 2013; NGMN, 2014), and reduction of cell costs (Chen and Zhao, 2014; Bhushan et al., 2014). Yazici et al. (2014) describe that the expenditure challenge will be met with utilization of NFV. Liang and Yu (2015) agree that since wireless network virtualization will enable sharing of infrastructure and radio spectrum resources, the CAPEX and OPEX of both RANs and CNs can be reduced significantly. Dense, heterogeneous networks with small cells and D2D will enable low-cost infrastructure (Bangerter et al., 2014). Although the reviewed literature describing these concepts does not emphasize the economic benefits as the main reasons for the choice of these technologies as relevant for 5G, it still seems likely that there has been a certain economic motivation behind the development of the concepts in the first place. For instance, it would not be feasible to only develop solutions that solved all the capacity issues of mobile networking, but were too expensive to actually implement in full scale. Finally, as mentioned in section 2.3.7, Hong et al. (2014) describe that understanding the trade-off between energy efficiency and spectral efficiency can provide direct guidelines to the OPEX management of cognitive cellular networks. If this understanding is utilized, it should possible to obtain reduced OPEX in 5G.

An example of "hidden" economic considerations, is the cost aspects of network densification, as described in section 2.3.9. Network densification will give a significant reduction in cost-per-bit delivered Bhushan et al. (2014). However, to obtain the full effect of network densification, it should be complemented by densification of the backhauls. The latter is expensive enough to challenge the initial gains of network densification. This is an example of a technology where there are important economic considerations behind the proposed solution. It also shows how important it is to consider the total economic gains before an investment option is taken. To study the cost-effectiveness of proposed solutions, techno-economic analysis is a possibility. This is shown by Nikolikj and Janevski (2015), who have done this for the AWA-HetNets, where they also have looked at the business profitability. They have used the method of discounted cashflows, among others. Similar analyses could be performed for other possible 5G components, to assess whether the investment opportunities are estimated to be profitable. An advantage with techno-economic analysis is that the technological aspect is also ensured, and the technology can be evaluated against the requirements in the same model.

Regarding the revenues, Bangerter et al. (2014) describe that the mobile industry needs a strategy that enables new services, new applications, and new opportunities to monetize the improved user experience. How is network performance related to revenue generation? Does better network performance automatically guarantee higher revenues in the future? Many would think so, but this is not necessarily true. Voice and message services have already become a commodity in the market, where price is the only factor the suppliers compete on. This drives the prices down. There is an ongoing debate on whether data traffic already is a commodity too, or when it will become a commodity. However, it is still possible to differentiate data transmission in other terms than just the price, e.g. by quality.

The trade-off between cost and performance is also relevant for the end-users. As described by Nokia (2014), a shift away from best-effort mobile broadband towards network efficiency with truly reliable communication of high performance is expected. This means that for 5G, connection quality may be the most important differentiation factor the mobile network operators compete on. As shown in section 2.2.1 and 2.2.2, many authors have addressed the topics of quality, or QoS and QoE. Intelligent QoS management or connectivity management-as-a-service is also proposed, where the terminal chooses the best wireless connection given different quality measures for the available connection, constraints for the required QoS and constraints for personal costs. This is an important economic aspect of 5G. An understanding of how the users value the relationship between quality and cost gives valuable insight into how the 5G users' willingness to pay will be for the improved network performance associated with 5G. Further, as the maximum connection quality is expected to be very high, the users willingness to pay may finally become the strongest constraint, instead of the maximum capacity of the networks.

Tehrani et al. (2014) show that the mobile network operators should be able to earn a profit from D2D communication, if initial set-up goes through the base stations, in the macrocell tier. In that case, the operators can control and charge for traffic transmitted in the device tier. To realize this business potential, the D2D communication services must be more attractive for the consumers, than comparable alternatives. This means that the operators must answer the “pay for what” question (Tehrani et al., 2014). If the mobile network operators can ensure better performance, better QoS and sufficient network security, it should definitely be possible for the operators to gain profit by providing D2D communication. However, the user incentives for D2D communication must be good enough to attract the users, and still affordable for the mobile network operators. This will be an entirely new service instance, and the mobile network operators may consider to change their pricing schemes to monetize D2D communication. Such changes may also concern the M2M communication, especially associated with the IoT. Many IoT devices will need to be online at all time, and not everyone will have access to WiFi or Bluetooth connections. For some IoT devices, wireless connection through the cellular network is the only possibility, and new types of subscriptions for these devices could be a new business potential for 5G. This is an example of how the mobile network operators can monetize the IoT.

Nokia (2016) describes that *new* business opportunities in 5G will arise from new applications that are enabled by the minimized latency and increased reliability. These are applications that will not work properly if the network delay is too high, and have not been possible to use over earlier generations of wireless networks, 4G included. This is reasonable from a logical point of view. Many of the mission critical communication scenarios presented by Nokia (2016) belong to this category. This is also analogous to the assumption that 5G will contribute to the continuation of the mobilization of industries and processes, as described by NGMN (2014).

As described in 2.6, NGMN (2014) expects the business context for 5G to be different from the business context of 4G. They agree with Nokia (2016) in that the new use cases and the improved network performance will lead to the main difference. As the networks are improved, the users will expect more and develop new needs that 5G must solve. The improved network performance will enable the value creation in 5G. NGMN (2014) also describes that the business models will expand to new ones, and 5G must support them at a very low cost and provide the flexibility to enable business models on demand. The business models will include the evolution of current services, and the emergence of new ones. Further, NGMN (2014) also expects that new types of customers and partnerships will be seen in 5G. The mobile network operators are expected to continue the development of own services, but also expand the business reach by entering relevant partnerships. The latter, along with some other important outlooks for the business aspects of 5G, are defined out of the scope of this thesis, but will be briefly explained in section 3.3.

Some last examples of how 5G changes the business context for other actors than mobile network operators follows. Nokia (2016) mentions virtual reality (VR) as a use case for 5G. VR technology can be applied to design and prototype products before they are built. Use of VR in product development can enable shorter development time, as it makes it easier to demonstrate and show the product idea before a physical prototype is built. For the same reasons, it may also lower the development costs. Therefore, use of VR technology may simplify product development in the future, and 5G may be a key-enabler for this.

5G will also change the business context for actors producing wireless hardware. These actors are labeled “architects and devices builders” in the overview of Internet actors in figure 3.1, which is presented further in the next section. The producers of users terminals will have to design new and improved user terminals as described in section 2.3.2, to enable the complete 5G experience and let the users experience all the network improvements. Producers of IoT devices are also likely to benefit from 5G, as 5G is expected to enable a realization of the full potential of the IoT, which is described in section 2.1.1.

3.3 Limitations

The literature scope of this thesis has been literature that is focused on 5G explicitly. This choice was taken due to the objective of defining 5G, where the differences from 4G have been emphasized much more than the similarities. This is because the reader is assumed to have some general knowledge of 4G systems, and the question “what is new in 5G?” is probably more interesting in this context, than the opposite (“what is old?”). Therefore, an understanding of 5G’s unique characteristics is assumed to be more important as a starting point for future research - especially given the assumption that the researchers are already familiar with 4G. Because of this, other existing research directions that are highly relevant in the extended context of 5G have been left out of the scope. Two examples will be presented here.

The first example is economic research on 4G. As 5G is assumed to be an evolution of 4G in many ways, one possible angle for economic research on 5G is to study research on the economic aspects of 4G. Research problems where the background and initial assumptions are assumed to be similar for 5G as for 4G could be continued directly into 5G research. Research problems where the context will change from 4G to 5G, for instance in terms of different technology, could be modified to be applicable for 5G research.

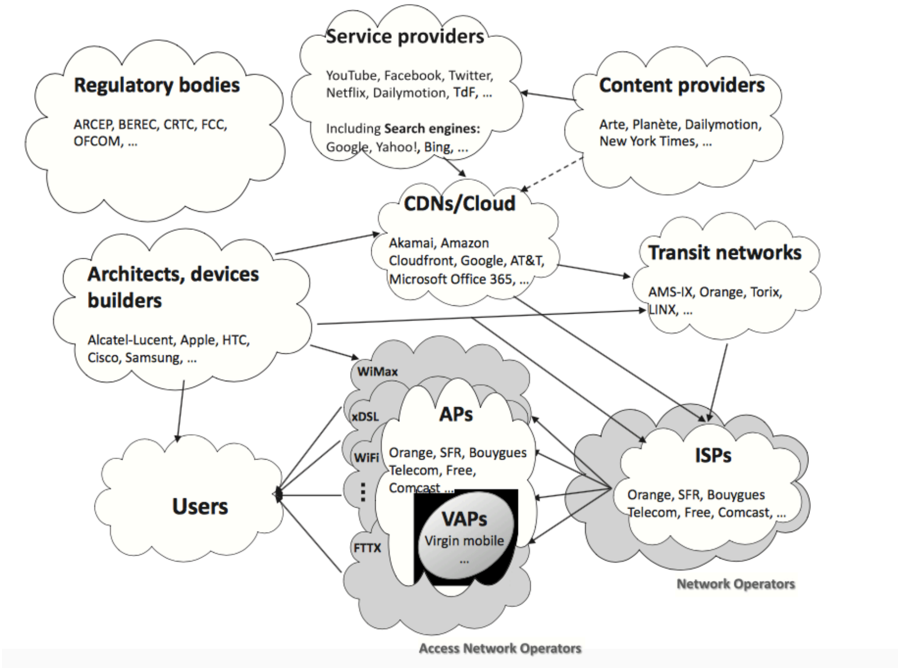

Figure 3.1: Overview of the Internet ecosystem (Maillé and Tuffin, 2014, p.16).

The second example is the "general" issue of the Internet supply chain, i.e. the uneven distribution of revenues and costs. The access providers who own the infrastructure necessary for providing Internet connectivity have much higher costs, both OPEX and CAPEX, than other Internet actors such as service and content providers. However, the latter actors receive a much higher revenues-to-expenses ratio, while they depend on the infrastructure provided by other actors in order to provide their services or distribute their content. Mobile network operators are both access providers and Internet service providers, offering wireless broadband. An overview of the different actors in the Internet ecosystem is shown in figure 3.1. The question "who shall pay for the Internet?" has been on the research agenda for several years, as huge investments in the infrastructure is necessary to operate, expand and improve the Internet into the future. One example of a proposed solution to fill what is referred to as the revenue gap, is vertical integration in the supply chain. For instance, Dai and Tang (2009) have studied vertical integration between telecommunication operators and service providers, while Maillé et al. (2016) have studied vertical integration between content distribution networks (abbreviated CDNs in figure 3.1) and network operators. Liang and Yu (2015) mention the relationship between mobile network operators and mobile virtual network operators (MVNOs), where the latter operators do not own infrastructure, but instead lease from the mobile network operators. They describe that the MVNOs may help the mobile network operators attract more users by providing specific services. The mobile network operators can increase their revenues by leasing isolated virtualized networks to the MVNOs.

Another example of a proposed solution, is paid peering. A simplified and brief explanation of paid peering is when two Internet companies connect their networks, and one of the companies pays the other company for prioritizing its traffic. This raises the debate on network neutrality, i.e. the basic Internet principle that all Internet data should be treated equally, with no discrimination or differential charging by user, content, website, platform, application, type of attached equipment, or mode of communication. This debate is already important, and is likely to still be on the agenda when 5G is deployed. The network neutrality issue is further described by Maillé and Tuffin (2014).

Chapter 4

Planning Problem and Mathematical Model

In this chapter, we show an example of how portfolio theory can be used for network planning in the 5G era. This is the quantitative contribution of this thesis. First, a network planning problem for a mobile network operator is described, before the developed mathematical model of the problem is presented. The problem is then compared to Modern Portfolio Theory and Portfolio Optimization to enlighten the similarities between these models. Finally, a brief outline for further analysis of the model is suggested.

4.1 Background

With the insight obtained from the literature review, it is clear that the end-users have enormous expectation to the next generation of mobile networks. In fact, many subscribers already expect much from the existing mobile networks. When 5G is launched, the end-users are likely to rely even more on their network operator, as they in the future are expected to become even more dependent on constant connectivity of high quality in their daily lives. The latter is a result of several factors, such as the emerging IoT and the digitization of the society in general.

The optimization problem chosen to describe in this chapter is a planning problem for a mobile network operator of the 5G era. As the end-users will expect more from their network connections in the future, the pressure on the mobile network operator to provide perfect connectivity to accommodate the needs of their users is likely to grow. To have access to a connection is not longer enough, because the connection should also guarantee low latency and high reliability, and hence provide the users a high quality of experience (QoE). QoE has become more important over the last few years, and is expected to play a significant role in the evaluation of the 5G experience. Hence, the overall objective of this part of the thesis is to expand a simple model for network flow and resource allocation to also take into account the perceived QoE of the provided services. In the problem presented here, the flow elements are service units, and the resources are network connections with a limited capacity.

The portfolio model for end-to-end differentiated services presented by Gaivoronski et al. (2016) has been used as a starting point for the model in this thesis. They study the selection of a service portfolio that maximizes the profit for a Internet service provider (ISP). The service portfolio, i.e. the set of all decisions, consists of the units of services offered to each market segment, the pricing of each service to each market segment, the provided quality of service (QoS) and whether additional network capacity is installed. In the model described in this thesis, revenues, network costs, the opportunity cost of not meeting demand and the option of expanding the network capacity is handled in the same manner. However, while Gaivoronski et al. (2016) focus on the two dimensions services and market segments, network connections are added as a third dimension in our problem. The main decision is the selection of connections to include in the network, instead of the service offer itself. Finally, to make the 5G context more clear, the quality parameter in the model is changed from QoS to QoE. We have also introduced an additional opportunity cost that occurs if the perceived QoE is not sufficient.

Although the context of our problem is 5G related, especially given the focus on QoE, the model might as well be used by a 4G operator today. In other words, there are no characteristics in the planning problem described here that do not apply for 4G, but the model is intended to be just as applicable for the future 5G network logistics.

4.2 Problem Description

The network planning problem we would like to resolve is that a mobile network operator chooses a portfolio of connections to offer to its users, to maximize its own profit. The users belong to different market segments, and they request different services. The mobile network operator invests in different connections, defined by different capacities, that is used to serve the service demand from the users. Revenues are generated when services are provided over the connections.

There are also a number of associated costs. We assume that there are both fixed and variable costs for providing a connection. The network is assumed to have a total capacity dedicated to service provision, but the operator may choose to invest in additional capacity if necessary. There are both fixed and variable costs for network expansion. There is also a fixed operating cost for the mobile network, and a variable cost for maintaining one unit of capacity. Finally, there are two different opportunity costs. The first is the opportunity cost of not meeting demand, i.e. not be able to deliver a requested service to a user. The second one is the opportunity cost of providing a connection of too poor quality to satisfy the QoE required by the user. These terms represent the risk in the problem, as not met demand or provision of too poor quality will lead to dissatisfaction of the customers. If capacity or quality issues seem to be persistent, churn will arise, i.e. loss of customers to competitors. This is a risk the mobile network operator must address, as churn decreases its revenues, and hence the profit. Therefore it is important to include the mentioned opportunity costs in the model, as an economic penalty. In reality, another significant risk is the demand itself. For simplicity in the model, the demand is expressed as a linear function of the price.

Further logic in the problem is that each connection has a given capacity, and the sum of the capacities for all the connections the network operator chooses to provide may not exceed the total capacity in the network. We assume that the QoE varies with the available capacity in a given connection, and that each market segment of users have fixed requirements for the QoE of each service.

4.3 Mathematical Formulation of the Model

In this section, the developed mathematical model for the described planning problem is presented.

Indices and sets

$i \epsilon I$	is the index for services.
$j \epsilon J$	is the index for available connections.
$k \epsilon K$	is the index for market segments.

Parameters

W_j	is the maximum capacity connection j can provide.
c_j^f	is the fixed cost associated with the provision of connection j.
c_j^v	is the variable cost for providing one unit of connection j.
c_{ik}^d	is the opportunity cost of not meeting a unit of demand from market segment k for service i.
c_{ik}^q	is the opportunity cost of providing QoE one unit below $\overline{q_{ijk}}$ when service i is delivered to market segment k.
c^f	is the fixed cost required for capacity expansion.
c^e	is the variable cost for expansion of capacity by one unit.
c^m	is the variable cost for the maintenance of one unit of capacity.
c^b	is the basic operating cost of the mobile network operator.
W_0	is the existing network capacity dedicated to serve the demand from all market segments.
W	is the additional network capacity the mobile network provider may install. Thus, the total network capacity is $W_0 + W$.
$\overline{q_{ik}}$	is the QoE market segment k requires from service i.
$\underline{q_{ik}}$	is the lowest possible QoE where market segment k will still demand service i.

Variables

p_{ik}	is the price charged per unit of service i provided to market segment k.
x_{ijk}	denotes the number of units of service i provided to market segment k over connection j.
z_j	is a binary variable that equals 1 if connection j is included in the connection portfolio of the mobile network operator, and 0 otherwise.
y	is a binary variable which equals 1 if the decision to install capacity is taken, and 0 otherwise.

Functions

The QoE, q_{ijk}, varies with the available capacity on the given link:

$$q_{ijk} = \begin{cases} \overline{q_{ik}}, & \text{if } \sum\limits_{i\epsilon I}\sum\limits_{k\epsilon K} x_{ijk} \leq (1-\alpha_{ik})W_j \\ \overline{q_{ik}} - \frac{(\overline{q_{ik}}-\underline{q_{ik}})}{\alpha_{ik}W_j}(\sum\limits_{i\epsilon I}\sum\limits_{k\epsilon K} x_{ijk} - (1-\alpha_{ik})W_j), & \text{if } (1-\alpha_{ik})W_j < \sum\limits_{i\epsilon I}\sum\limits_{k\epsilon K} x_{ijk} < W_j \\ \underline{q_{ik}}, & \text{if } \sum\limits_{i\epsilon I}\sum\limits_{k\epsilon K} x_{ijk} \geq W_j \end{cases}$$

for $i\epsilon I$, $j\epsilon J$ and $k\epsilon K$. $\alpha_{ik}\epsilon\{0,1\}$ is a constant that vary for both every service and every market segment. The larger (closer to 1) α_{ik} is, the higher the QoE requirement for service i from market segment k is.

It is assumed that several connections can be used to deliver a service to a market segment, if necessary. The average quality of each service delivered to each market segment of customers must be computed across the different connection. The fairest way to approximate this is to calculate the average quality by weighting the quality of each connection with the share of the service that is sent over each connection:

$$q_{ik} = \frac{\sum\limits_{j\epsilon J} q_{ijk}x_{ijk}}{\sum\limits_{j\epsilon J} x_{ijk}}.$$

The demand d_{ik} is assumed to be a linear function of the price p_{ik}:
$d_{ik} = d^0_{ik} + \beta_{ik}(p^0_{ik} - p_{ik})$,
where the parameters d^0_{ik} and p^0_{ik} are reference values for the demand and price, respectively, and β_{ik} is a sensitivity parameter.

Mathematical Formulation

Maximizing the profit of the service provider:

$$\begin{aligned} P = &\sum_{i\epsilon I}\sum_{j\epsilon J}\sum_{k\epsilon K}(p_{ik} - c^v_j)x_{ijk} - \sum_{j\epsilon J} c^f_j z_j - \sum_{i\epsilon I}\sum_{k\epsilon K} c^q_{ik} max\{0, \overline{q_{ik}} - \sum_{j\epsilon J} q_{ik}\} \\ &- \sum_{i\epsilon I}\sum_{k\epsilon K} c^d_{ik} max\{0, d_{ik} - \sum_{j\epsilon J} x_{ijk}\} - y(c^f + c^e W) - c^m(W_0 + yW) - c^b \end{aligned} \tag{4.1}$$

Subject to:

$$\sum_{j \epsilon J} W_j z_j - W y \leq W_0 \tag{4.2}$$

$$\sum_{i \epsilon I} \sum_{k \epsilon K} x_{ijk} - M z_j \leq 0, j \epsilon J \tag{4.3}$$

$p_{ik}, x_{ijk} \geq 0, i \epsilon I, j \epsilon J, k \epsilon K.$
$z_j, y \epsilon \{0, 1\}, j \epsilon J.$

The profit function (4.1) contains the following parts: The first term is the revenue minus the variable cost for each service unit provided, which net result is the margin. Then, the second term is the fixed costs for the connections the operator provides, followed by the opportunity cost of providing a connection of too poor quality and the opportunity cost of not meeting the demand. The fifth term is the costs of network expansion if the decision to expand is taken, the sixth term is the maintenance costs and seventh term is the fixed basic operating costs.

Restriction (4.2) shows that the sum of the capacity for all connections included in the connection portfolio must not exceed the total possible capacity of the network dedicated to service provision. Restriction (4.3) is a technical constraint that connects continuous and binary variables. M denotes a big number, that preferably should be chosen according to the area of feasible solutions, to make the restriction tight. Finally, there are non-negativity constraints for the decision variables p_{ik} and x_{ijk}, and binary constraints for the other decision variables, z_j and y.

4.4 Explanation of the Quality Function

In the model, it is assumed that the QoE perceived by the user, q_{ijk}, depends on the available capacity on connection j. To explain it further, we look at a simplified case with only one connection, one service and one user. The available connection has a total capacity W, which equals the maximum number of service units that can be sent over the connection simultaneously. The number of service units sent over the connection is denoted x. The user requires that the service is delivered with a quality $\overline{q}$. This is the case as long as the total number of service units provided is less than or equal to $(1 - \alpha)W$. The α parameter takes a value between 0 and 1, and is specific for the service and the user. The larger (closer to 1) α is, the sooner the perceived quality will decrease as more of the capacity of the connection is used. When the number of service units (x) provided over the connection exceeds $(1 - \alpha)W$, the quality will decrease linearly until the maximum capacity is reached. Then, the quality equals $\underline{q}$, which is the lowest QoE value where there is still demand. Figure 4.1 shows how q is a function of x. In this particular case, $\underline{q}$ is 0.

In the model, several connections, services and segments of users are considered. Hence, the QoE depends on all of these. To determine the available capacity on a connection j, we must summarize over all services i delivered to all market segments k over the connection j. We further assume that the parameters $\overline{q}$, $\underline{q}$ and α are specific for both each service i and each market segment k, and hence their notations in the model are $\overline{q_{ik}}$, $\underline{q_{ik}}$ and α_{ik}.

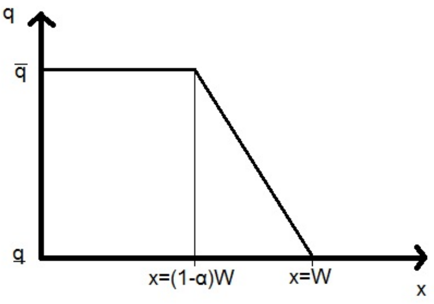

Figure 4.1: How the delivered quality is assumed to vary with the available capacity on a connection.

4.5 Similarity with Portfolio Optimization

The perceptive reader will notice the similarity between portfolio optimization problems and the model described in this chapter. Modern portfolio theory (MPT) or mean-variance analysis was originally described by the American economist Harry Markowitz in 1952. The objective of MPT is to maximize the expected return of a portfolio of assets, for a given level of risk:

Maximize $E(R_p) = \sum\limits_{i \epsilon I} x_i E(R_i)$

s.t.

$\sigma^2(x) \leq \omega$

$\sum\limits_{i \epsilon I} x_i = 1.$

Here, $E(R_p)$ is the expected return of the portfolio, x_i is the fraction of the budget invested in asset i, and $E(R_i)$ is the expected return of asset i. $\sigma^2(x)$ is the variance of the portfolio, and ω is the upper bound for the risk that the variance of the portfolio must not exceed. The portfolio return variance is calculated in the following way:

$$\sigma^2(x) = \sum_{i \epsilon I} x_i^2 \sigma_i^2 + \sum_{i \epsilon I} \sum_{j \neq i} x_i x_j \sigma_i \sigma_j \rho_{ij},$$

where σ_i^2 is the variance of the return of asset i, and ρ_{ij} is the correlation coefficient between the returns on asset i and j. Further, as shown by the second constraint, the sum of all the asset weights must equal 1. If it is not allowed to hold short positions in the assets, a non-negativity constraint is added for all the x_i variables. Due to the non-linear variance constraint, it is more computationally feasible to solve the inverse of the problem, i.e. finding the portfolio that minimizes the risk for a given lower bound of return. MPT is further described by Zenios (2008).

There are several common features of MPT and the model for the planning problem described in this chapter. The first similarity is that the MPT model chooses the fraction of the budget to invest in the different assets that is included in the asset portfolio, and in the planning problem the network operator chooses which available connections that should be included in the connectivity portfolio. In the MPT model, the investor seeks to maximize the expected return of his assets, while the network operator maximizes its expected profit. Further, both problems address the issue of risk. While the MPT model has a constraint for the variance of the portfolio, the uncertainty is embedded in the objective function in the planning problem. The demand is assumed to be a function of the price, and if the demand is not met, an opportunity cost accrues. The other risk in the planning problem is that the operator delivers too poor quality on their connection, and the QoE demands of the subscribers are not met. If this occurs, the profit function is penalized by another opportunity cost. Finally, there are limited resources in both problems. The weights of the assets cannot exceed 1 in MPT, and in the planning problem, the capacity of all the connections that are chosen to be included cannot exceed the total network capacity. Further, the number of service units sent over each connection cannot exceed the capacity of the given connection.

4.6 Outline for Further Analysis

To continue the work with the model, the next step would be to perform a numerical analysis by implementing the model in commercial software. For instance, the model could be implemented in MATLAB and solved with the nonlinear programming solver *fmincon* in Optimization Toolbox$^{\text{TM}}$. As this solver minimizes the objective function, the defined objective function (4.1) should be multiplied by -1. Further, the functions defined outside of the model, i.e. the functions for quality and demand, would have to be rewritten to be included in the model before implementation.

The model would obviously be most useful if empirical or historical data could be obtained and used as parameters. However, as these data could be hard to obtain, one could still gain a lot of insight in both characteristics of both the problem and the model by analyzing the model with example data. Table 4.1 shows example of different market segments and characteristics, which could be used as a starting point for setting different parameters and adjust them according to the different market segments and services.

Further, a similar overview could be made for different services, and the associated QoE demand for the particular service, for each market segment. If the parameters are made in a systematic and thorough manner, the results from the model will become more valid than if the example data were chosen in a more random manner.

Some examples of issues that could be investigated:

- The choice of different connections: For which capacities and costs would it be optimal with several connections with less capacity, compared to fewer connections with greater capacity?

Market segment	**Ability and willingness to pay**	**Services**	**General QoE demand**
Students/young adults	Low ability, low willingness	Social media, multimedia streaming, general entertainment	Medium to high
Older adults	High ability, medium willingness	Simple broadband services	Low to medium
Business segment	High ability and high willingness	Video conferencing	High

Table 4.1: Examples of three different market segments and their characteristics.

- Investigation of the investment decision isolated: how big should the demand be before it is feasible to expand the network capacity? When are the expansion costs so high that it is not feasible to expand, even if the demand is sufficient?
- How much capacity is needed to avoid the opportunity costs? For which values of the opportunity costs is it worth or not worth the risk that they might occur?
- Given fixation of the costs in different ranges, what is the optimal trade-off between increasing the prices or increasing the demand by lowering the prices?
- With a more stochastic behavior of the demand, than a linear relationship with the price variable: How much does the optimal solution (the profit) change if the demand changes within or outside a given interval?
- Other general sensitivity and “what if” analyzes.

These are simply a few suggestions of different difficulties, as there are many other interesting aspects that could be subject for further analysis as well. The list above shows the broad range of possibilities of how the behaviour of the model could be explored.

4.7 Shortcomings

The first shortcoming is that the model has not been implemented and analyzed numerically. Because of this, it is hard to evaluate how “good” the model is, or the computational performance of the model. Another shortcoming in the proposed model is that the QoE is simplified to depend on the available capacity of the connections. This is a clear contradiction to how QoE is defined in the literature as described in section 2.2.2, as QoE is highly subjective and may depend on many different factors, including latency and bandwidth. However, for the purpose of this model, it had to be simplified. For any model that seeks to maximize performance in terms of quality, several metrics will give a better result, but at some point there will be too many metrics to obtain a manageable runtime when the problem is solved. This trade-off should be carefully chosen. If QoE consideration is too complex for a problem, QoS is a good alternative, as it is more specific and less subjective.

As mentioned, network planning with QoS as quality measure is shown by Gaivoronski et al. (2016).

Chapter 5

Conclusion and Outlook

In this chapter, the conclusions that can be drawn from this thesis are presented. Suggestions for future research on economic aspects of 5G are also presented.

5.1 Conclusion

Even though 5G is not standardized yet, the work on the first set of 5G standards is expected to begin this year. Therefore, the main focus of the 5G research so far has been to understand which communication needs the next generation of mobile networks must cover, what the requirements are, and which technologies that may ensure that the requirements are met. The use cases and requirements are clear, which means that the current focus is the technologies, as these are not formally decided yet. Some of the new use cases that make 5G different from 4G are machine-to-machine communication including the Internet of Things, extended use of augmented reality and virtual reality technology, as well as the need for a significant improvement of the mobile broadband technology, to provide truly reliable mobile communication with extremely low latency.

In addition to the general requirements of higher data rates, reduced end-to-end latency, higher mobile data volume per area and being able to support a higher number of connected devices, it is also expected a much bigger focus on the overall quality of the connections. The ultimate request for 5G is to provide intelligent Quality of Service management. Examples of different measures to be included in a QoS assessment are delay, jitter, bandwidth and reliability. It is proposed that in 5G, different services may run on different connections simultaneously. The best available connection for a given service may be found by evaluating available QoS data against QoS requirements and restrictions for personal costs. Quality of Experience may be just as important as QoS when the quality of a connection is evaluated, as this is based on the user's subjective perception.

The literature review further shows that the following technologies are expected to be especially important to enable 5G: heterogeneous networks, device-to-device communication, software-defined networking, network function virtualization, cognitive radio, transmission over the millimeter frequency bands, advanced multiple-input multiple-output techniques and software-defined radios. Cloud concepts such as SDN and NFV are expected to lower the mobile network operators' expenses, as they enable sharing of infrastructure and radio spectrum resources. D2D communication will also contribute as low-cost infrastructure. Further, it is predicted that the cell costs will be reduced in the 5G heterogeneous networks. The possibilities for cost reductions are important, as the mobile network operators' expenses have increased over the last years. Even though it is of interest to minimize both capital and operational expenditures, it is even more important to obtain the required capacity improvement in the networks. The growing cost pressure is also addressed in literature focusing on the green perspective of 5G. As both power consumption and the related CO_2 emissions are also relevant for the economic aspects of 5G, the further research on green aspects should be payed attention to.

An important finding is that there has, unfortunately, not been conducted much 5G research fully devoted to the economic aspects yet. This makes sense, as it is natural that the technology is the main focus right before the standardization begins. However, the review of 5G economic research shows that some of the proposed technologies already seem to be profitable. Tehrani et al. (2014) have found that the mobile network operators should be able to earn a profit from device-to-device communication, given that the initial set-up goes through the base station, so the associated traffic can be controlled and charged for. Nikolikj and Janevski (2015) provide a cost-effectiveness assessment for the 5G heterogeneous networks, where they have found that the 5G millimeter wave systems can ensure the network operators a solid profit, with more than 50 % profit margins. This is promising regarding the need for 5G to be economically sustainable.

The business context of 5G will include both new and old services. New business opportunities arise from applications that are first enabled when the network delay is minimized and the reliability is increased. This includes mission critical communications, a huge number of connected devices arising from the Internet of Things, and mobilization of different industries. The telecommunication industry must decide how to monetize the 5G user experience. The improved network performance will enable most the value creation in 5G. New types of customers and partnerships may also characterize the business context of 5G.

As an example of further economic research of 5G, we have formulated a planning problem for a mobile network operator. Based on how the reviewed literature treats the quality aspect of the 5G connection experience as an important factor for the subscribers, we show how this can be included in an optimization model for network planning. In the resulting portfolio theory model, a mobile network operator seeks to maximize their total profit, by choosing an optimal portfolio of available connections to offer to its customers, to provide the necessary connections for the network traffic related to the customers' service demands, while also meeting their quality demands. The optimal prices to charge for each service provided to each market segment of customers are also decided, in addition to the decision of whether the operator should invest in additional network capacity or not.

There are several shortcomings in this thesis. First, the reviewed literature is limited to research on 5G explicitly. As 5G first and foremost will be an evolution of 4G, rather than an entirely new communication system, many of the economic issues of 4G will be relevant for 5G as well, but not all of them are referred to in the 5G literature yet. As one of the main objectives was to explain how 5G will be different from 4G, economic research on 4G networks had to be left out of the scope of this thesis. After discovering how narrow the field of 5G economics currently is, it should be recommended to rather continue economic research on 4G into 5G, than using only publications regarding 5G technology as a starting point. The fact that 5G is still on a early stage is a clear shortcoming for economic research. This is likely to change once the first set of standards are released. Another shortcoming is that we have not been able to perform a numerical analysis of the planning model described in chapter 4. This should be performed by someone who has more competency on solving large optimization problems than the author of this thesis.

5.2 Recommendations for Future Research on 5G Economics

As described above, a more thorough economic review would be obtained if research on the economic aspects of 4G was collated with what is known about 5G so far. Before technologies to be included in 5G are finally selected, techno-economic analysis in the same manner as shown by Nikolikj and Janevski (2015) could be performed for the promising technology candidates, to assess the profitability of the technologies before they are selected. Equally, it will be similarly important to assess the costs of the selected technologies together after the first set of 5G standards are released, to understand the size of the total future costs associated with network development, and plan how these costs should be financed. When 5G is commercially deployed, empirical data will be available, which means that further economic research can be based on these data. Then it will also be possible to compare predicted costs and actual costs, to check the accuracy of old estimates. This research may also be compared to the costs of 4G networks, to study whether there is a gain in terms of cost reductions.

Regarding the revenue aspect, it will be important to understand how the increased network performance affects the users' willingness to pay. How much is 5G connectivity worth for the users, and how should the new user experience be monetized? The mobile network operators should explore this question, to develop a pricing strategy that the users will accept. The socioeconomic aspect of wireless connectivity may also be subject for further research. One research question example is which value 5G will have for socioeconomic development in different places in the world. Regarding portfolio theory for telecommunications, numerical analysis of the model presented in chapter 4 could be performed. Another suggestion for application of portfolio theory for 5G, is the proposed intelligent QoS Management, where the best wireless connection for a given service is chosen based on available QoS data and restrictions for personal costs.

Bibliography

3GPP, 2017. Release 15. Retrieved March 23rd, 2017.
URL `http://www.3gpp.org/release-15`

Agyapong, P., Iwamura, M., Staehle, D., Kiess, W., Benjebbour, A., 2014. Design considerations for a 5g network architecture. IEEE Communications Magazine 52 (11), 65–75.

Bangerter, B., Talwar, S., Arefi, R., Stewart, K., 2014. Networks and devices for the 5g era. IEEE Communications Magazine 52 (2), 90–96.

Bhushan, N., Li, J., Malladi, D., Gilmore, R., Brenner, D., Damnjanovic, A., Sukhavasi, R., Patel, C., Geirhofer, S., 2014. Network densification: the dominant theme for wireless evolution into 5g. IEEE Communications Magazine 52 (2), 82–89.

Carlton, A., February 2017. 5g reality check: Where is 3gpp on standardization? Retrieved March 18th, 2017.
URL `http://www.networkworld.com/article/3174138/mobile-wireless/5g-reality-check-where-is-3gpp-on-\standardization.html`

Chávez-Santiago, R., Szydelko, M., Kliks, A., Foukalas, F., Haddad, Y., Nolan, K. E., Kelly, M. Y., Masonta, M. T., Balasingham, I., 2015. 5g: The convergence of wireless communications. Wireless Personal Communications 83 (3), 1617–1642.

Chen, S., Zhao, J., 2014. The requirements, challenges, and technologies for 5g of terrestrial mobile telecommunication. IEEE Communications Magazine 52 (5), 36–43.

Dai, R., Tang, S.-l., 2009. Analysis on vertical integration of telecom operators. ICMECG International Conference on Management of e-Commerce and e-Government.

Demestichas, P., Georgakopoulos, A., Karvounas, D., Tsagkaris, K., Stavroulaki, V., Lu, J., Xiong, C., Yao, J., 2013. 5g on the horizon: Key challenges for the radio-access network. IEEE Vehicular Technology Magazine 8 (3), 47–53.

Denazis, S., Hadi Salim, J., Meyer, D., Koufopavlou, O., 2015. Software-defined networking (sdn): Layers and architecture terminology.

Gaivoronski, A., Nesse, P., Østerbo, O., Lønsethagen, H., 2016. Risk-balanced dimensioning and pricing of end-to-end differentiated services. European Journal of Operational Research 254 (2), 644–655.

Gohil, A., Modi, H., Patel, S. K., 2013. 5g technology of mobile communication: A survey. International Conference on Intelligent Systems and Signal Processing, 288–292.

Hong, X., Wang, J., Wang, C.-X., Shi, J., 2014. Cognitive radio in 5g: A perspective on energy-spectral efficiency trade-off. IEEE Communications Magazine 52 (7), 46–53.

Hossain, S., 2013. 5g wireless communication systems. American Journal of Engineering Research 02 (10), 344–353.

Janevski, T., 2009. 5g mobile phone concept. Institute of Electrical and Electronics Engineers, 1–2.

Javaid, M. A., 2013. 5g technologies: Fundamental shift in mobile networking philosophy. SSRN Electronic Journal.

Li, Q. C., Niu, H., Papathanassiou, A. T., Wu, G., 2014. 5g network capacity: Key elements and technologies. IEEE Vehicular Technology Magazine 9 (1), 71–78.

Liang, C., Yu, F. R., 2015. Wireless network virtualization: A survey, some research issues and challenges. IEEE Communications Surveys Tutorials 17 (1), 358–380.

Lucas, E., February 2017. In south korea, the race is on for olympics 5g next year. Retrieved May 3rd, 2017.
URL `https://phys.org/news/2017-02-south-korea-olympics-5g\-year.html`

Maillé, P., Simon, G., Tuffin, B., 2016. Vertical integration of cdn and network operator: Model and analysis. IEEE Mascots.

Maillé, P., Tuffin, B., 2014. Telecommunication Network Economics: From Theory to Applications. Cambridge University Press, Cambridge, UK.

Marek, S., November 2016. Koreas kt says it will launch 5g by 2019 will verizon follow? Rettrieved May 3rd, 2017.
URL `https://www.sdxcentral.com/articles/news/koreas-kt-says-will-launch-5g-2019-will-verizon-follow/2016/11/`

NGMN, 2014. Ngmn alliance: 5g white paper - executive version.

Nikolikj, V., Janevski, T., 2015. Cost-effectiveness assessment of 5g systems with cooperative radio resource sharing. Telfor Journal 7 (2), 68–74.

Nokia, 2014. White paper: 5g use cases and requirements.

Nokia, 2016. White paper: 5g for mission critical communication: Achieve ultra-reliability and virtual zero latency.

Olsson, M., Cavdar, C., Frenger, P., Tombaz, S., Jantti, R., 2013. 5green: Towards green 5g mobile networks. WiMob 2013, The 9th IEEE International Conference on Wireless and Mobile Computing Networking and Communications: GROWN International Workshop on Green Optimized Wireless Networks, 1–5.

Osseiran, A., Boccardi, F., Braun, V., Kusume, K., Marsch, P., Maternia, M., Queseth, O., Schellmann, M., Schotten, H., Taoka, H. e. a., 2014. Scenarios for 5g mobile and wireless communications: the vision of the metis project. IEEE Communications Magazine 52 (5), 26–35.

Patel, S., Chauhan, M., Kapadiya, K., 2012. 5g: Future mobile technology-vision 2020. International Journal of Computer Applications 54 (17), 6–10.

Patil, S., Patil, V., Bhat, P., 2012. A review on 5g technology. International Journal of Engineering and Innovative Technology 1 (1), 26–30.

Roh, W., Seol, J.-Y., Park, J., Lee, B., Lee, J., Kim, Y., Cho, J., Cheun, K., Aryanfar, F., 2014. Millimeter-wave beamforming as an enabling technology for 5g cellular communications: theoretical feasibility and prototype results. IEEE Communications Magazine 52 (2), 106–113.

Rost, P., Bernardos, C. J., De Domenico, A., Girolamo, M. D., Lalam, M., Maeder, A., Sabella, D., Wbben, D., 2014. Cloud technologies for flexible 5g radio access networks. IEEE Communications Magazine 52 (5), 68–76.

Sapakal, Reshma S.Kadam, S. S., 2013. 5g mobile technology. International Journal of Advanced Research in Computer Engineering Technology 2 (2), 568–571.

Seehusen, J., March 2013. Moderne bygg dreper mobildekningen. Retrieved April 15th, 2017.
URL `https://www.tu.no/artikler/moderne-bygg-dreper-\mobildekningen/234252`

Sharma, P., 2013. Evolution of mobile wireless communication networks - 1g to 5g as well as future prospective of next generation communication network. International Journal of Computer Science and Mobile Computing 2 (8), 47–53.

Tehrani, M. N., Uysal, M., Yanikomeroglu, H., 2014. Device-to-device communication in 5g cellular networks: Challenges, solutions and future directions. IEEE Communications Magazine 52 (5), 86–92.

Vollen, D.-R. Z., July 2016. Første kommersielle 5g-nett i korea. Retrieved May 3rd, 2017.
URL `http://www.cw.no/artikkel/5g/forste-kommersielle-5g-\nett-korea`

Wang, C.-X., Haider, F., Gao, X., You, X.-H., Yang, Y., Yuan, D., Aggoune, H., Haas, H., Fletcher, S., Hepsaydir, E., 2014. Cellular architecture and key technologies for 5g wireless communication networks. IEEE Communications Magazine 52 (2), 122–130.

Yazici, V., Kozat, U., Sunay, M., 2014. A new control plane for 5g network architecture with a case study on unified handoff, mobility, and routing management. IEEE Communications Magazine 52 (11), 76–85.

Zenios, S. A., 2008. Practical Financial Optimization, 1st Edition. Blackwell.