Institutionen för systemteknik
Department of Electrical Engineering

Examensarbete

A Study on Random Access Performance in Next Generation Mobile Network Systems

Examensarbete utfört i Kommunikationssystem
vid Tekniska högskolan vid Linköpings universitet
av

Magnus Thalén

LiTH-ISY-EX--15/4850--SE

Linköping 2015

Department of Electrical Engineering
Linköpings universitet
SE-581 83 Linköping, Sweden

Linköpings tekniska högskola
Linköpings universitet
581 83 Linköping

A Study on Random Access Performance in Next Generation Mobile Network Systems

Examensarbete utfört i Kommunikationssystem
vid Tekniska högskolan vid Linköpings universitet
av

Magnus Thalén

LiTH-ISY-EX--15/4850--SE

Handledare: **Antonios Pitarokoilis**
ISY, Linköpings universitet
Pål Frenger
Ericsson AB

Examinator: **Danyo Danev**
ISY, Linköpings universitet

Linköping, 2 juli 2015

Avdelning, Institution
Division, Department

Communication Systems
Department of Electrical Engineering
SE-581 83 Linköping

Datum
Date

2015-07-02

Språk
Language

☐ Svenska/Swedish
☒ Engelska/English
☐ ____________

Rapporttyp
Report category

☐ Licentiatavhandling
☒ Examensarbete
☐ C-uppsats
☐ D-uppsats
☐ Övrig rapport
☐ ____________

ISBN
—

ISRN
LiTH-ISY-EX--15/4850--SE

Serietitel och serienummer
Title of series, numbering

ISSN
—

URL för elektronisk version

http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-120551

Titel
Title

En studie i random access prestanda i nästa generations mobila nätverkssystem
A Study on Random Access Performance in Next Generation Mobile Network Systems

Författare
Author

Magnus Thalén

Sammanfattning
Abstract

The next generation of mobile telecommunication, 5G, will be specified in the near future. One of the proposed changes relative to the previous generation, 4G, is the inclusion of a new system control plane (SCP). The purpose of the SCP is to improve system scalability, forward compatibility, peak performance and to enable a higher degree of support for advanced antenna techniques. This is done by logically separating data transmitted explicitly from and to the user, the dynamic transmissions, from the broadcasted transmissions that remain constant regardless of user activity, the static transmissions, and by then redesigning the static part to make it more lean. This is expected to have several positive effects such as considerably more free resources, resulting in energy savings and potentially increased usage of MIMO. Knowing what effect the SCP has upon aspects such as random access is of importance when designing the solution that will go into the standard.

Simulations show that there is potential in the inclusion of the new SCP. While the simulated 5G candidate systems that include the SCP have an increased delay when running the random access procedure, some aspects of the procedure have been improved. The main differences relative to the simulated 4G system are the performance of the first message in the procedure, which increased, and the performance of the second message in the procedure, which decreased. The deficiencies found in the handling of the second message, however, can be alleviated by using a more proper algorithm and channel design than what was used in this study.

Nyckelord
Keywords

5G, LTE, System control plane, Random access

Sammanfattning

Nästa generationes mobila kommunikation, 5G, kommer specificeras inom en snar framtid. En av de föreslagna förändringarna relativt den tidigare generationen, 4G, är tillägget av ett nytt systemkontrollplan. Syftet med detta systemkontrollplan är att förbättra systemets skalbarhet, framåtkompabilitet, maxprestanda och att möjliggöra en högre grad av användande av avancerade anntenntekniker. Detta görs genom att logiskt separera den data som explicit sänds till och från användare, dynamisk transmission, från den data som sänds oavsett om det finns någon aktiv användare eller ej, statisk transmission, och genom att sedan förändra den statiska delen så att den blir smalare. Detta förväntas ha flera positiva effekter så som avsevärt mer oanvända resurser vilket resulterar i energibesparingar och potentiellt en högra grad av MIMO användade. Att veta vilka effekter systemkontrollplanet har på aspekter som random access-proceduren är av stor vikt när man ska designa den lösning som ska bli del av standarden.

Simulationer visar att det finns potential i tillägget av systemkontrollplanet. De simulerade 5G kandidatsystemen som innehåller systemkontrollplanet har en ökad fördröjning vid användande av random access-proceduren, men vissa aspekter av proceduren har förbättrats. De huvudsakliga skillnaderna relativt det simulerade 4G systemet är prestandan av det första meddelandet i proceduren, som förbättrades, och det andra meddelandet i proceduren, som försämrades. Bristerna i hanteringen av det andra meddelandet kan dock åtgärdas genom en mer lämplig algoritm och kanaldesign än vad som användes i denna studie.

Abstract

The next generation of mobile telecommunication, 5G, will be specified in the near future. One of the proposed changes relative to the previous generation, 4G, is the inclusion of a new system control plane (SCP). The purpose of the SCP is to improve system scalability, forward compatibility, peak performance and to enable a higher degree of support for advanced antenna techniques. This is done by logically separating data transmitted explicitly from and to the user, the dynamic transmissions, from the broadcasted transmissions that remain constant regardless of user activity, the static transmissions, and by then redesigning the static part to make it more lean. This is expected to have several positive effects such as considerably more free resources, resulting in energy savings and potentially increased usage of MIMO. Knowing what effect the SCP has upon aspects such as random access is of importance when designing the solution that will go into the standard.

Simulations show that there is potential in the inclusion of the new SCP. While the simulated 5G candidate systems that include the SCP have an increased delay when running the random access procedure, some aspects of the procedure have been improved. The main differences relative to the simulated 4G system are the performance of the first message in the procedure, which increased, and the performance of the second message in the procedure, which decreased. The deficiencies found in the handling of the second message, however, can be alleviated by using a more proper algorithm and channel design than what was used in this study.

Acknowledgments

First and foremost I would like to thank my supervisor at Ericsson AB, Pål Frenger, and my supervisor at Linköping University, Antonios Pitarokoilis, for their constant support throughout these past six months. It is thanks to you that I have been able to move forward despite facing many difficulties along the way.

I also wish to thank everyone else who has been involved in my thesis for creating a great environment that has made this a very positive experience for me. Your presence has been invaluable and has helped me greatly in realising this thesis.

Linköping, June 2015
Magnus Thalén

Contents

Notation

Abbreviations

Abbreviation	Definition
3GPP	3rd Generation Partnership Project
bps	bits per second
CDF	Cumulative Distribution Function
dB	deciBel
dBW	deciBel Watt
EPS	Evolved Packet System
ISD	Inter Site Distance
LTE	Long Term Evolution
MAC	Medium-Access Control
MBSFN	Multi Broadcast Single Frequency Network
MIMO	Multiple Input Multiple Output
PDCCH	Physical Downlink Control Channel
PDCP	Packet Data Convergence Protocol
PDSCH	Physical Downlink Shared Channel
PHY	Physical Layer
PRACH	Physical Random-Access Channel
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Shared Channel
RAN	Radio-Access Network
RAR	Random Access Response
RLC	Radio-Link Control
SCP	System Control Plane

1 Introduction

For the last few decades a new generation of mobile communication systems has been developed with a periodicity of about ten years. Next in line is the fifth generation, 5G, which is currently in the process of being specified. Due to the ever increasing demands combined with new research being implemented the expectations are greater than ever. 5G is expected to deliver rates up to 10Gbps for up to a hundred times as many devices as today, with a latency of less than 1ms. Furthermore a 90 percent reduction in energy usage as well as virtually complete availability and worldwide coverage is expected [14].

In this thesis we aim to investigate a subset of changes that may potentially be implemented in the new standard. These modifications might fundamentally alter the design of contemporary communication systems. More specifically we are interested in how these changes affect the procedure used when mobile users perform the initial procedure to connect to the cellular network, i.e. the random access procedure.

1.1 Motivation

Due to their close relation to the telecommunication field, Ericsson as well as other major players have a large interest in participating in the ongoing standardization effort. Considering that 5G is intended to be out on the market beyond year 2020 [13] this is the time to get involved. It is believed that 5G will need to undergo fundamental changes, relative to previous generations, in order to reach the high expectations on data rates and latencies [8].

The predecessor of 5G, 4G, was a considerable upgrade relative to previous systems. 4G contains functions such as spectrum flexibility, multi-antenna transmission and intra-cell interference control [7]. All of this has made 4G a relevant technology, and considering the longevity of even earlier generations it will continue to be so for many years to come.

All wireless transmissions make use of electromagnetic waves, meaning that the physical nature of the signals are electromagnetic waves. The nature of the electromagnetic waves is such that an antenna receiving two different signals at the same time, where the signals are of the same frequency, will be unable to tell the signals apart. The receiver equipped with that antenna will only be able to perceive the sum of the signals.

In order to avoid excessive interference between signals 4G schedules all signals in a certain service area on different time and frequency resources. This improves the probability of successful reception of transmitted messages, but it also imposes some constraints as it further limits the total number of transmissions that can be done in a certain time window.

The increasing amount of time-frequency resources spent on overhead in 4G has become an increasingly large issue. As result of this there will of course be less resources left to use for transmission of other data, but another concern is the impact it has on future development.

The main task of wireless systems such as 4G is to transfer information from transmitter to receiver. In order to accomplish this it is necessary to also transmit certain control information. The control information, which for instance consists of transmission schedule and system settings, can be considered to be overhead as it is only transmitted as a means to make the system work.

One important kind of control signal is the reference signal that each base station transmits in regular intervals. The reference signals are necessary because they allow the users to gain a measurement of the channel between user and base station. As all users need this information the base stations broadcast the reference signals, meaning that it is sent in a way that allows all users in the vicinity to receive it.

While there is a need of broadcasting reference signals it is not necessarily needed to constantly spend as many resources on this as is done in 4G. This can be considered to be an improvement area as this causes the system to be hard to extend, and attempting to put in even more functionality would likely become an increasingly difficult and complicated process. Instead it would probably be beneficial to start over from a clean slate and reduce the amount of overhead clogging the system. Doing so would also make the system more energy efficient and reduce interference. As 5G is expected to evolve and change considerably compared to 4G, now is a good opportunity to implement this change.

Long Term Evolution (LTE), which is synonymous to 4G, and similar systems provide coverage to the users through strategically placed base stations. Base stations are essentially towers equipped with antennas, enabling them to transmit and receive electromagnetic waves. The base stations are generally evenly distributed as a single base station can provide coverage to an area around it. In places with more traffic the base stations are more densely deployed as a single base station is limited in the amount of traffic it can handle.

In LTE the term eNodeB is used to describe a base station. Each eNodeB consists of one or more cells. Each cell is an independent logical entity that serves a certain area around the eNodeB. All communication to and from all users in this area is handled by the designated cell.

When a user wants to connect to the system the user will try to perform random access to the cell that governs the area that the user is located in. As the name implies random access is the procedure when a user, that the system has no prior knowledge of, suddenly attempts to establish a contact with the system. In order for the user to be able to do this it needs to have received a reference signal and system information from the cell. In LTE the reference signals are unique to each cell, meaning that a reference signal obtained from one cell is only useful when trying to communicate with that specific cell. This results in the user being able to differentiate between cells, the user will realize if it moves into another cell as that cell would then transmit a different reference signal. The opposite of this would be a system where all reference signals were equal. The user would then have no way to distinguish between different cells. The desirable and flexible approach would be to make both of these alternatives possible. This would allow the system to choose whichever reference signal setup that it finds most suitable at the time.

Enabling the user to make use of several cells in the random access procedure, as opposed to only one cell in LTE, has been identified as a promising change. If this was implemented, and all cells in an area broadcasted the same reference signals, it would reduce interference between cells and thus allow for a more dense placement of eNodeBs without increasing system overhead.

An addition to 5G that has been suggested is the inclusion of a new system control plane (SCP) [4][5][6]. At its core the SCP is a subset of all transmitted data in the system. The SCP represents the data that is necessary for basic system functionality, such as initial access to the network. This includes signals sent even when all users are idle, or when there are no users in the system, as well as signals the random access and paging procedures. This includes signals such as reference signals and some broadcasted system information, but it does not include data that is explicitly requested by the user. The purpose of this system control plane is specifically to reduce the amount of signals that are unnecessarily broadcasted while also logically separating the signals that are part of the SCP from the other signals, making it possible to configure the system so that users will be unaware of specific cells during the initial access to the system. As the problems are inherent to the nature of the previous generations it would not be simple to solve through a new release of the LTE-standard. Instead it is convenient to include it as a more integrated change for the upcoming 5G.

The inclusion of the system control plane would change many parts of the system. The focus of this thesis, however, is specifically how it affects the random access procedures. Before the upcoming standardization effort it is important to fully understand the consequences of different suggestions. Having data that support your suggestions is key. Likewise it is important to be aware of any potential weaknesses of your suggestions, which is why it is crucial to gain insight into what effects the SCP has on random access, among other things.

1.2 Purpose

The aim of this thesis is to find to what extent the SCP is positive or negative to the random access procedures of the upcoming 5G system. To be able to make such a judgement we will need to look at different metrics that determine the effectiveness of the random access.

When performing the random access procedure a user will send and receive a certain set of messages in accordance with a protocol. The resulting signal to noise ratios (SNR) at the receiver of these messages as well as energy efficiency will need to be considered. The total time it takes for a user to connect to the system and the amount of concurrent users the system can serve is also of interest.

Depending on the achieved result it may be necessary to include additional changes to the random access procedure in 5G. Otherwise the inclusion of the SCP may cause the performance of the this procedure to deteriorate below acceptable levels. It is therefore necessary to gather data of performance of the random access functions during different circumstances and compare it to LTE. LTE is considered as a baseline because it is the closest predecessor of 5G and is thus the closest to 5G in both performance and structure.

If the final results are positive it becomes more likely that the SCP, as described in this report will, be a part of 5G. While the SCP is considered to be a good change overall it is not certain that it will not negatively impact some parts of the system, making it necessary to investigate its effects on procedures such as random access.

1.3 Problem Formulation

Researchers have identified the inclusion of a new system control plane to be a promising change for the upcoming 5G [4][5][6]. While it is known that this would have benefits to some aspects of the system compared to the current LTE-standard, it is also possible that it will incur some drawbacks. The random access procedure is one of the areas that will be directly affected by this change. Therefore it is necessary to investigate in detail how metrics such as the delay and power consumption of the procedure are affected and how the procedure can be improved as a result of the SCP. The system's capacity in terms of random access user throughput should also be considered. This thesis aims to answer the following:

- How is the power control for the random access procedure affected by the SCP?
- How does the SCP affect the random access procedure in terms of delay compared to the 4G baseline?
- How can the random access procedure that uses the SCP be improved and how do these changes affect the delay and user throughput of the procedure?

The first question refers to the power that users estimate that they will need to use in order to successfully transmit the preamble to its intended receiver. The result will differ from 4G as the SCP enables the user to combine reference signals from several different cells.

The SCP is certain to affect the performance of the random access procedure, though it is not known whether the effect will be positive or negative. The purpose of the second question is to gain information on this matter.

In the event that the SCP negatively impacts the random access procedure it may be necessary to find ways to decrease this effect. The third question refers to any potential changes that will improve the procedure. In this case it is mainly the performance in terms of total delay and user throughput that is considered.

1.4 Limitations

When analysing the capabilities of the system there are different parameters that can be changed. Each possible configuration is likely to yield quite different results. It is not possible to investigate all of the possible circumstances relevant to the random access procedures, so this thesis is limited to a smaller set of representative cases. It is of interest to see how performance metrics, such as delay, are affected by changes in some key variables. Except for these few variables, essentially everything will be held constant. The variable parameters are:

- **Inter site distance**, the distance between neighbouring eNodeBs.
- **Number of preambles**, the number of preambles that users may choose from when transmitting the first message of the random access procedure. Each user chooses one.

- **Intensity**, the number of users/s that are added to the system and that the system is tasked to service.

5G will of course consist of many changes when compared to LTE, but the only change considered in this report is the inclusion of the SCP. Therefore the 5G test system discussed in this report is actually close to LTE in functionality. Most other changes that are likely to be included in 5G will probably not significantly impact the random access procedure and are therefore not included in the 5G test system used in this report.

1.5 Thesis Outline

This report is divided into an introductory part that presents some theory and background, and a result part that covers the execution and achieved results.

- **Chapter 1** gives an overview of the problem and the suggested solution, and presents the problem formulation.
- **Chapter 2** introduces the concept of wireless transmission and some related technologies and phenomena.
- **Chapter 3** presents the typical structure of cellular networks and provides information about the cellular networks system named 4G. Some information about its structure and protocols are described. Additionally the random access procedures used in 4G are described in more detail.
- **Chapter 4** gives and introduction to the SCP. The issues that have prompted the inclusion of the SCP into future cellular network systems are presented. Furthermore the actual changes that constitute the SCP and the expected benefits are described.
- **Chapter 5** details the scenario as well as parameters used in the simulator during the different stages of the investigation.
- **Chapter 6** presents the results.
- **Chapter 7** discusses noteworthy results as well as possible weaknesses of the used method. The impact the thesis might have on society is briefly mentioned.
- **Chapter 8** states what can be concluded from the results and attempts to answer the questions posed in the problem formulation. Suggestions of what additional research could be made in order to answer these questions are also made.

2

Wireless Transmission Overview

Wireless mobile systems make use of many different kinds of techniques to acquire sufficient quality of signal transmission. These techniques are essential for the level of modern cellular services. Further exploitation of these techniques is therefore something that is desirable in upcoming generations as well as in new releases of current mobile telephony systems. This is thus something that should also be kept in mind when analysing the topic of this report, the performance of the random access procedure in a future system. For this reason this chapter introduces the most relevant techniques and phenomena in order to give a framework of understanding that is helpful when weighing the options of 4G and 5G against each other.

This chapter briefly introduces some wireless transmission basics, fading and diversity.

2.1 Basics

Typically when talking about wireless transmissions we are interested in the communication between two units, a transmitter and a receiver. The transmitter is the unit that transmits the signal while the receiver is the unit that receives the signal. The transmitter and the receiver do not necessarily have any special properties that differentiates them, rather they are only defined by the direction the signal is going in. The signal always originates at the transmitter. The task of the receiver is to detect the signal and interpret it appropriately.

The physical entities that constitute the signals are virtually always electromagnetic waves. Electromagnetic waves travel at the speed of light and have many different names depending on the frequency of the wave, and by extension the application in which it is used. For example, visible light, radio waves and microwaves are all instances of electromagnetic waves. It is often convenient to model the electromagnetic waves as sinusoidal waves. The general equation

$$x(t) = a \sin t, \tag{2.1}$$

where x is the electromagnetic wave, a is the amplitude and t is time, can be used as a simplified model of an electromagnetic wave.

An important property of electromagnetic waves is that they decrease in power the further they travel. This shows through the amplitude which decreases and eventually becomes so small that the wave is no longer measurable at the receiver. The path loss is a metric that describes the attenuation of the signal strength. In free space conditions, where there are no objects obstructing the signal, the path loss can be calculated as

$$P = \left(\frac{4\pi d}{\lambda}\right)^2, \tag{2.2}$$

where P is the path loss, d is the distance between the transmitter and the receiver and λ is the wavelength of the carrier of the transmitted signal. Because of this fundamental constraint the distance between the transmitter and the receiver is a limiting factor in wireless communication. Communication is only possible if the transmitter and the receiver are not separated by too large of a distance as in that case it will be impossible for the receiver to detect the transmitted signal due to its low energy.

If the only thing obstructing wireless communication was the path loss it might still be possible to communicate regardless of how small the amplitude of the transmitted signals are, as long as it is larger than zero. In practice however, there exist several other obstacles that prevent this. The perhaps most prevalent problem hindering this is the ubiquitous noise. Noise is a general term that is usually used to describe all kinds of distortion of signals, except for the interference from other signals, and can have a number of causes. The most common kind of noise is that which is called white noise. White noise, which is a synonym to thermal noise, exists in all kinds of circuitry and also afflicts all electromagnetic waves that travel through the air. The white noise has a constant power spectral density, meaning that it is equally strong at all frequencies, and is generally modelled to have a Gaussian distribution.

If a signal at the transmitter can be described by (2.1) then the same signal at the receiver can in some simple cases be described as

$$y(t) = b \sin(t + \phi) + n(t), \tag{2.3}$$

where $y(t)$ is the received signal, b is the amplitude of the received signal, t is the time, ϕ is the phase shift relative to the transmitted signal and $n(t)$ is the white noise. The receiving amplitude b can be calculated as

$$b = \frac{a}{\sqrt{P}}, \tag{2.4}$$

where P is the path loss according to (2.2) and a is the amplitude of the sent signal at the transmitter. The noise $n(t)$ on the other hand is not affected by the distance. This means that when a signal modelled as (2.3) has travelled far enough the noise portion of the signal will be significantly larger than the intended message, making it impossible for the receiver to make any sense of the signal. Figure 2.1 shows a sine wave before passing through a channel as well as what happens when the signal is attenuated and has noise added to it.

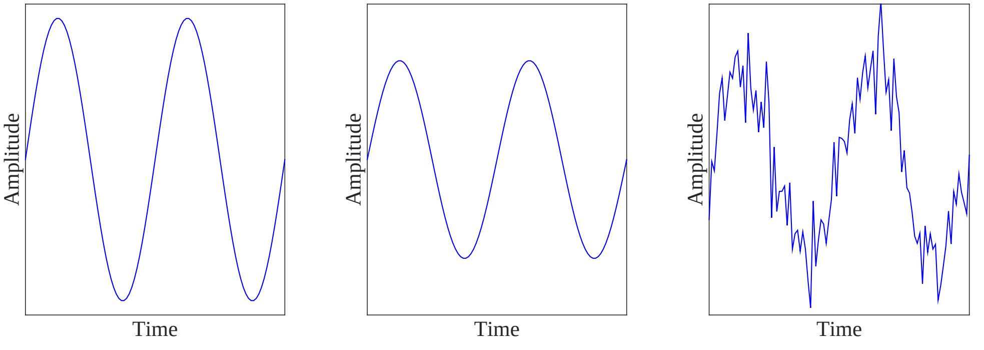

Figure 2.1: *The effects on a sine wave when passing through a channel that attenuates the signal and applies additive white Gaussian noise.*

When discussing the difference between the signal at the transmitter and the same signal at the receiver it is common to make use of something called the channel. The channel is essentially a filter between the transmitter and the receiver through which the signal propagates. It describes exactly how the signal is affected by travelling through the air. The channel is dependent on the positions of the transmitter and the receiver, and the environment. Assuming that the environment changes, which is always the case in real scenarios, it is also dependent on the time. The channel is not, however, dependent on the signal that travels through it, all signals are treated the same way. A general relation between the transmitted and the received signal is

$$y(t) = h(t) * x(t) + n(t), \tag{2.5}$$

where $y(t)$ is the received signal, $h(t)$ is a mathematical representation of the channel, $x(t)$ is the transmitted signal, $n(t)$ is the noise, t is the time and $*$ denotes convolution. The majority of wireless communication likely occurs close to the surface of the earth. The average call between two mobile phones, for example, will usually have the two phones at ground level. An effect of this is that the transmitted signals are obstructed by objects in the environment. For example buildings and mountains may very well block the path of the signal. Electromagnetic waves can not always pass through house walls, and certainly not mountains. What they can do is reflect and refract off objects. For this reason the transmitted signal will typically take several paths and bounce off different objects before arriving at the receiver. This is called multipath propagation.

Figure 2.2 shows the concept of multipath propagation. Two of the lines displayed in the graph correspond to the two signals that arrive at the receiver after having travelled along two different paths. The third line, the signal which the receiver preceivs, is simply the sum of the two other signals.

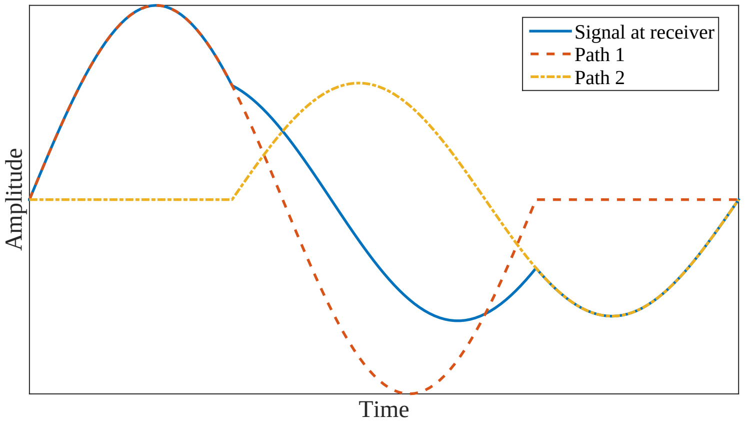

Figure 2.2: *An example of multipath propagation without noise.*

2.2 Pre- and post-processing

Except for the actual transmission of the signal there are also some steps that, to some degree, need to be taken to prepare the signal for the distortive effects of the channel. This includes source coding, channel coding and modulating the message. The result is a compressed, error protected message that is ready to be mapped onto a waveform. After receiving the signal the receiver needs to reverse the operations in order to retrieve the original message. This typical flow is showed in Figure 2.3.

2.2.1 Source coding

The purpose of source coding is to compress and reduce the size of the target information. This is not something unique to wireless transmission but is rather done in very many different contexts. Many different source coding schemes exist and they yield a different degree of compression depending on the information that is to be compressed.

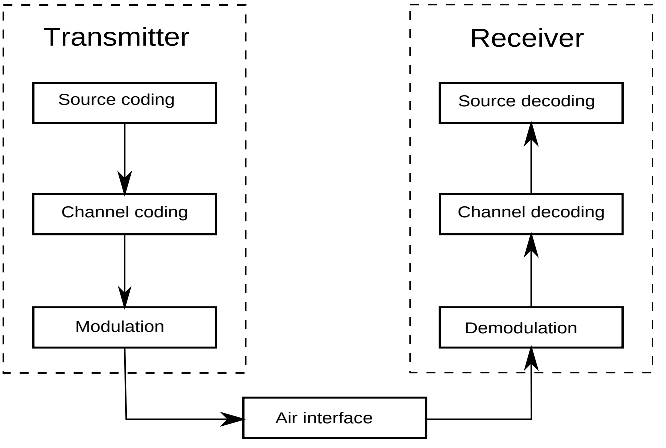

Figure 2.3: *A common order of operations for information that is to be transmitted from a transmitter to a receiver.*

In wireless transmission the source coding is beneficial as it allows the transmitter to actually transmit shorter bit sequences while losing little or no information. This comes at a computational cost both at the transmitter and at the receiver, but this is acceptable as it typically is the actual wireless transmission that is the bottleneck and not the pre- or post-processing.

2.2.2 Channel coding

Channel coding is included in order to counter errors that may appear in the receiver due to distortion of the transmitted signal. When the receiver receives the signal it will first demodulate it, resulting in a series of bits. If not for the error correction it would be impossible to tell whether the information had been corrupted or not. With error correction, however, it may be possible to detect and possibly even correct a number of errors. In a practical situation this allows the receiver to either correct the errors or at least discard it and request a retransmission of the message if needed.

The error correction operates on a set of bits and works by adding redundancy. The redundant bits are used to create a structure that allows the receiver to spot errors if they exist.

2.2.3 Modulation

The modulation is used in order to convert the message consisting of a series of bits into a waveform that can be transmitted over the air. The modulation essentially maps the bits to different looking sinusoids. The resulting signal that will then be transmitted is all the sinusoids appended in the same order as the bits that made up the message.

2.3 Fading

If a transmission between transmitter and receiver takes place in an area with different obstacles spread out, as opposed to in a free space, the signal will propagate towards the receiver along several paths as described in Section 2.1. In the end this causes the receiver to receive the summation of several distorted copies of the signal. Due to the distortion it is not certain that the received signal will look anything like the original transmitted signal. Signal properties such as the amplitude, phase and angle of arrival are subject to variations due to the surrounding environment [12]. One effect of this is that the power of the resulting signal at the receiver may at times be enhanced and at times reduced.

The fading phenomenon is traditionally divided into two parts depending on their rate of change: large-scale fading and small-scale fading. Large-scale fading is the fading component caused by large objects obstructing the signal and large movements. Small-scale fading on the other hand, which is also called Rayleigh fading [12], corresponds to the significant effects that can be caused by movements of the receiver as little as half a wavelength of the transmitted signal. The name Rayleigh fading comes from the fact that the signal amplitude is distributed according to a Rayleigh distribution [12].

2.4 Diversity

Diversity is the method of providing more instances of the same information to a receiver in order to reduce the negative effect of fading. Fading, which is caused by multipath propagation, can randomly both improve and worsen the quality of a signal at the receiver. What happens when some kind of diversity is included is that the receiver, through for instance multipath propagation, receives several copies of the same information. Even if the receptions of each single instance is no better than before it becomes much more likely that at least one of the instances is of sufficient quality for the receiver to interpret correctly. Several different ways to provide diversity exists.

2.4.1 Spatial Diversity

Spatial diversity is simply the method of attaining diversity by placing transmitters and receivers at different positions. Between each existing transmitter and receiver there will be at least one path. Therefore, as long as there is more than just one transmitter and one receiver there will be a gain in diversity. It is assumed that all transmitters send the same signal and that the receivers have a way of synchronising their information. However, this diversity gain can only be said to happen under a certain constraint. Only when the additional transmitters are spaced sufficiently far apart from the original transmitter, or when the additional receiver is spaced far enough apart from the original receiver. If, for instance, two receivers are located at almost the exact same position it is likely that the signals transmitted from the transmitters will propagate along the same paths. Such a case will not give additional diversity. The distance needed in order to gain diversity varies, but is at least 0.5λ, where λ is the wavelength of the transmitted signal [9]. A metric of how similar two units are in this sense is correlation. When the distance between two units increase the correlation will generally decrease.

One practical way of creating spatial diversity in cellular systems is by using several different system nodes to communicate with a user when sending a message. This method may be especially relevant to this thesis as the system control plane may allow this behaviour for certain procedures.

2.4.2 Delay Diversity

Delay diversity makes use of time differences in order to gain diversity. If a channel is time dispersive the different propagation paths will cause different delays. In such a case the receiver will receive several instances of the transmitted signal, although they are probably corrupted and interfere with each other. This is an example of delay diversity. Note that this is possible even in the case where there is only a single transmitter and a single receiver. It relies on the quality of the channel. Even if the channel is not time dispersive it is possible to create artificial time dispersion by transmitting the same signal from several antennas with small time delays in between [9]. If the antennas are placed apart this can also be seen as spatial diversity. The time dispersion and delay diversity can, however, also be achieved this way even when antennas are placed right next to each other, with less space between them than what is required for spatial diversity.

3

LTE Overview

Long-Term Evolution (LTE) is a mobile telephony system that was first publically available during 2009 [2]. It is a system consisting of logical units called eNodeBs and mobile users such as mobile phones. A traditional implementation of the eNodeB is a base station. An eNodeB consists of one or more cells, where a cell is also a logical construct. A single cell represents the interface that is used for transmission towards mobile phones in a specific area. This area that the logical cell governs is what is usually indicated when the word cell is mentioned. The eNodeBs are placed in such a way that they cover the area which they are supposed to provide coverage to. A typical, albeit idealistic, deployment of eNodeBs that is common for mobile telephony systems, is to arrange them in a hexagonal grid such that each eNodeB is located at the point where three hexagons intersect, see Figure 3.1.

Mobile users will typically communicate wirelessly with one of the cells of the eNodeB that is adjacent to whatever hexagon the user happens to be in at the moment. To accomplish this different techniques are used for transmission and reception of the signals that contain the requested data.

This chapter will give a short overview of LTE as a whole and will additionally go into more detail in the parts of LTE that are especially relevant to this thesis: the random access procedures as well as the parts of LTE necessary for random access.

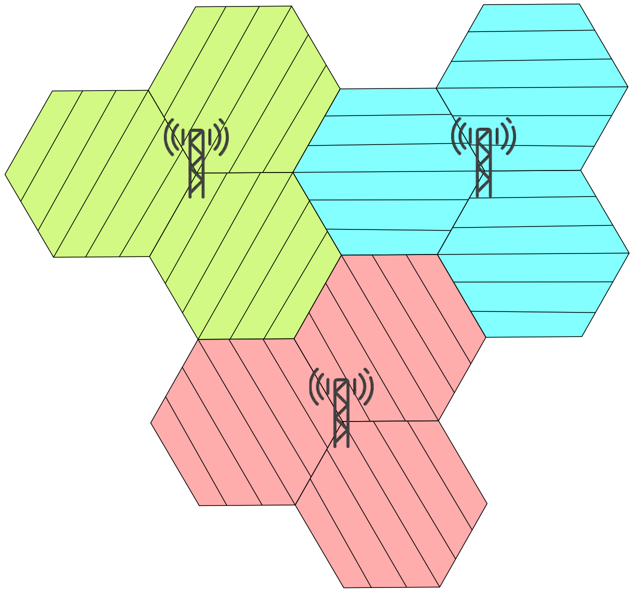

Figure 3.1: *A hexagon deployment using three cells per eNodeB. The hexagons with the same pattern are the areas that are governed by the cells that belong to the closest eNodeB.*

3.1 Cellular Networks Infrastructure

The most basic component of today's cellular network systems can be considered to be the user. The users are typically very spread out and most of them are likely to move at some point. This causes the need for a widespread coverage. The purpose of the system is to provide services, such as the ability to call other users or browse the Internet, and these services should preferably be available at all times.

The high requirements on availability require the deployment of certain infrastructure. The basic unit that provides the coverage is the eNodeB. The eNodeB is a stationary unit that on one hand wirelessly communicates with the surrounding users, and on the other hand communicates, typically through wires, with the underlying system. The main wireless part of cellular network systems is the part between the user and the eNodeB. The communication between eNodeBs are typically done using a wired approach, or sometimes through dedicated radio links that do not interfere with other communication.

It would be ideal if each station provided as much coverage as possible, meaning that each eNodeB should be able to pick up wireless transmissions sent from an area as large as possible. Thus, if carefully positioned it would be possible to deploy less eNodeBs without affecting system performance, allowing the operators to save money. A simple way of improving the eNodeBs ability to pick up signals is to give the eNodeBs antennas as clear a view as possible of the surroundings. As mentioned in Chapter 2, physical objects have the ability to obstruct the signals, causing a degradation in signal quality. The placement of eNodeBs on high buildings is advantageous as there will on average be much fewer objects that obstruct the view between the eNodeB and the different nearby users.

In addition to this it is of course also important to place the eNodeBs in such a way that the coverage they provide do not overlap unnecessarily. What limits the coverage is, except for obstructing objects, the distance between the eNodeBs and users. Assuming that measures have been taken to remove or lessen the issue of obstructing objects, the provided coverage can in some cases roughly be modelled as a circle with the eNodeB in the middle. A drawback of using circles to model the coverage area is, however, that it is very unwieldy to cover a large area completely with only circles without at the same time causing a lot of overlap between them. A more simple way of fulfilling both demands of complete coverage and as little overlap as possible is to use hexagons as building blocks instead of circles.

Commonly the eNodeBs employ several sets of antennas placed around the eNodeB, facing different directions. In such a case the antennas will not transmit in all directions, but they rather transmit their respective signals in an area spanned by an angle of 60 to 120 degrees, depending on the scenario. A way to structure the different antennas is to group all antennas facing a certain direction into one group. A common division of antennas is to divide them into three groups where each group of antennas transmits their respective signals to a third of the surrounding area, corresponding to a width of 120 degrees. In such a case each group of antennas could be mapped to the cell serving the corresponding area. A typical placement of eNodeBs consisting of three cells are shown in Figure 3.1. The result of this division is that the different groups all transmit their signals to essentially non-overlapping areas.

3.2 LTE Structure

LTE is at the topmost level divided into two parts: the Radio-Access Network (RAN) and the Evolved Packet System (EPS). The RAN takes care of all activities that are needed for a typical radio transmission between eNodeB and user such as scheduling, encoding of the signal, adding error correction and so on. The EPS, on the other hand, handles all other functionality that is needed to connect all entities in the network and make it run properly such as access to Internet, connection between eNodeBs as well as other necessary back end functions. It is primarily the RAN that is of interest in this thesis due to the random access being a part of it.

The RAN can be divided into several different protocol layers, as shown by Figure 3.2, where a protocol layer contains functions necessary for transmission and reception of messages. Adjacent layers also provide interfaces for each other, allowing them to call on each other in order to accomplish their tasks. These layers range from the topmost layer that performs the initial processing of any intended message, to the physical layer which interfaces the physical antenna components. The layers fulfil different tasks and provide proper building blocks for the system so that the interfaces in between the layers are effectively designed. The different layers are the Packet Data Convergence Protocol (PDCP), the Radio-Link Control (RLC), the Medium-Access Control (MAC) and the Physical Layer (PHY).

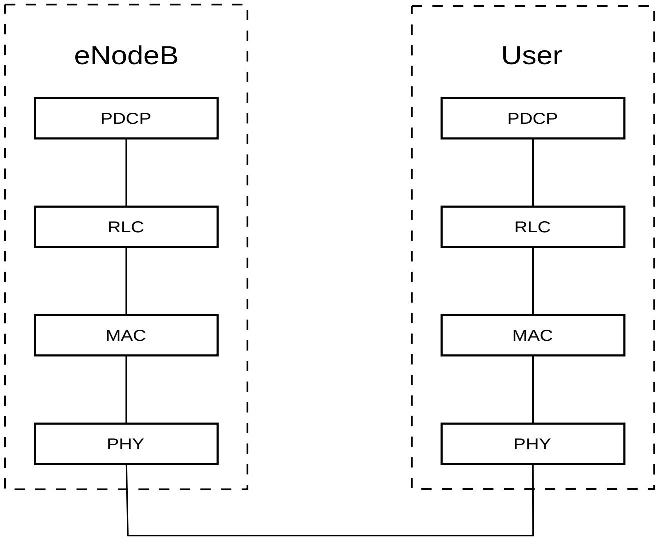

Figure 3.2: *The different protocol layers of the RAN. Users and eNodeBs communicate with each other through the physical layer.*

For each transmission there is a transmitter and one or more receivers. The different protocol layers exist in both the transmitter and the receiver but the related functions differ somewhat. Some functions, such as scheduling, are only performed for the network while the user simply adapts. The most common case, however, is that each function in a certain protocol layer at the transmitter has a mirroring function in the same protocol layer in the receiver. This is the case for functions such as encoding and modulation of signals which are mirrored by corresponding decoding and demodulation. It should also be mentioned that the behaviour of the layers in some cases depend on whether the message that is to be transmitted is on the downlink or on the uplink. A message is sent through the downlink if the message originated from an eNodeB and is sent to a user, and vice versa for the uplink. The difference is due to the fact that the demands on the preprocessing are different in the two cases. When sending from a user, on the uplink, the user has likely already received instructions from the eNodeB on when and how to send it.

Messages that are to be sent using the RAN are in the form of packets called IP-packets. An IP-packet consists of a header and a payload part. The header is simply some necessary meta data describing the message, and the payload is the actual data being sent. Before and after passing through the air interface this message is altered according to the instructions of the different protocol layers.

The first purpose of the PDCP is to compress the header of the packet in order to reduce the size of the message. The second purpose is to solve the privacy issue that is apparent when transmitting data over the air. After all, the transmitted signal will typically propagate with similar attenuation in all directions covered by the transmitting cell. To solve this problem the message is encrypted so that only the designated receiver is able to decipher it.

Upon receiving the compressed data blocks from the PDCP the RLC has the task of combining these blocks into larger or smaller blocks when necessary. The new data blocks are created by combining the incoming data blocks and splitting them where needed in order for the resulting data block to be of a specific length. The desired length can vary and depends on the data rate used, for high rates it is more efficient with large blocks whereas lower rates require smaller blocks. On the receiver side the RLC should also ensure that there are no errors in the messages that will be passed to the layer above it, the PDCP. If it detects any errors it is able to request a retransmission.

The MAC defines two sets of channels, the logical channels and the transport channels. Both of these kinds of channels are abstract in the sense that no signals are sent through the air; instead the channels function as interfaces to the surrounding protocol layers. The logical channels act as an interface between the MAC and the RLC and are defined by the type of information that is to pass through the channel. The transport channels instead act as an interface between the MAC and the PHY protocol layers and are defined by how the data on the channel is to be transmitted. A task of the MAC is to multiplex the channels so that they are correctly connected. Additionally the MAC takes care of all the scheduling. The scheduling determines what is to be transmitted in what resource block, where a resource block is a certain frequency during a certain time period. All scheduling in the cells belonging to an eNodeB, both for the downlink and for the uplink, is performed in this eNodeB, and the necessary information is then transmitted in advance to the affected users. Similarly to the RLC layer, the MAC also has the ability to request retransmissions in the cases of erroneous detections of signals sent through the air.

The physical layer, as the lowest layer, is responsible for, among other things, coding, modulation and mapping between transport channels and physical channels. LTE defines a scheduling entity called a frame, and a frame in turn consists of ten subframes. A subframe divides the available frequencies into resource blocks of 180kHz that each spans 1ms. The whole frame thus covers a time period of 10ms.

Each physical channel corresponds to a set of these blocks, all data belonging to a certain physical channel may only be transmitted during these times and at the corresponding frequencies. Following the end of each 10ms frame there is another frame, repeating into infinity.

Some of the available physical channels are the Physical Downlink Shared Channel (PDSCH), the Physical Downlink Control Channel (PDCCH), the Physical Uplink Shared Channel (PUSCH), the Physical Uplink Control Channel (PUCCH) and the Physical Random-Access Channel (PRACH). The PRACH, PDSCH and PUSCH in particular are used in the random access procedure.

3.3 Random Access in LTE

The random access functionality in LTE has the purpose of establishing a connection between a user and a cell. This is needed in the cases when there has been no prior communication between the cell and the user, since in that case the user and cell have no way of deciding a common schedule through which they can communicate. There are several situations where this is relevant, some examples are when a user's equipment has just been turned on and needs to connect to the system, when a user needs to perform a handover operation as it is about to leave a cell and enter another and thus needs to connect to the new cell, and when an already existing connection is lost due to a previous failure in communication.

An inherent issue with the random access procedure is that the user needs to acquire necessary information such as what frequencies to use and when to transmit its random access messages. In order to solve this, the cell broadcasts the required information which allows the user start the procedure.

The first part of the required information is received from a system information block (SIB). Several SIBs exist and they are transmitted on predetermined frequencies with certain time periods. The user simply has to repeatedly try to receive this message and it will eventually succeed in doing so. The second SIB in the set of SIBs provides information to the user of at which frequency and during what times, relative to the reception of the received downlink synch signal, the cell can receive the initial random access message.

A known and predetermined reference signal, often called a pilot, is also required and is therefore transmitted regularly by the cells. Since the user already knows the form and shape of the pilot the only knowledge gained by receiving this signal is how the channel has affected the signal. The channel refers to the radio link between the cell and the user. The user compares the received signal power with the power that it knows was used when transmitting the signal to attain an estimate of the channels path loss. This information is later used when the user transmits its own initial random access message to calculate the needed transmission power.

Two different kinds of random access procedures exist: contention-based and contention-free. As the name implies the contention-based scheme is what is used when several users simultaneously try to connect to the same cell using the same resources, this is the standard approach that is considered in this thesis. The contention-free scheme avoids collision between users by assigning a unique preamble, which is a kind of message used during the first step of the random access, to the user prior to the otherwise first step of the random access procedure. This is only possible when the cell somehow has information about the user beforehand, meaning that it is only possible in some use cases such as handover, and not in the initial connection case where the system has no information about the user whatsoever. The cost of doing a contention-free operation is the need for additional preambles; the system only has a certain amount of preambles, of which a majority is used for the contention-based access, so the use of the contention-free access is limited. The gain, on the other hand, is that it is possible to avoid the random delay caused by the risk of collision, when several users inadvertently use the same preamble at the same time causing the signals to collide and be indistinguishable at the receiver cell, in the contention-based scheme [11].

Figure 3.3 shows the set of messages sent in the contention-based procedure. The UE and eNodeB send a total of four messages, assuming that no step fails in which case the UE will have to start over. It is the UE that initiates the procedure.

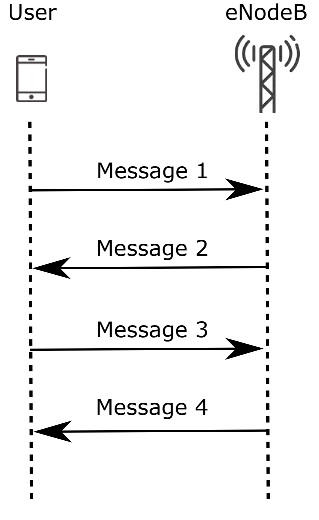

Figure 3.3: *The different messages sent during the contention-based random access procedure. The UE initiates the process by transmitting a preamble, after which the eNodeB and UE keep responding with consecutive messages until a total of four messages have been sent.*

3.3.1 Message 1

The message that initiates the random access process, assuming a contention-based scheme, is sent from the user and is called the random access preamble. The PRACH is used for this transmission. The purpose of the message is to inform the cell that this user is trying to access the network and to enable the cell to prepare the transmission of the second message.

The user chooses the preamble from a set of at most 64 preambles, possibly a bit less as the number of preambles used for contention-free access is subtracted from this number. While the function that determines what preamble to use is deterministic the input parameters to this function, such as the amount of data the user intends to transmit, can be considered to be randomly distributed [9]. This is equivalent to randomly choosing a preamble. Having decided upon a certain preamble the user then chooses the transmission power according to the equation

$$P_t = \min\{P_{\text{max}}, P_{\text{initialTarget}} \times P \times P_{\text{ramp}} \times \Delta\}, \tag{3.1}$$

where P_t is the transmission power, P_{max} is the maximum allowed power, $P_{\text{initialTarget}}$ is the baseline power value, P is the estimated path loss, P_{ramp} is a ramping factor and Δ is a constant value set by the system. If the user does not receive the expected answer it then keeps retransmitting a preamble but with a higher power, P_{ramp} corresponds to this gradual increase in power and is increased according to a set protocol with each failed transmission.

As the preamble that the user chooses is random there is a possibility that other users choose the same preamble. Since the preambles do not contain any user information a certain preamble will look the same no matter who sends it. If two or more users send the same preamble the result will therefore be that the cell receives the approximate sum of the two signals but perceives it as a single signal. The cell will then proceed as usual with message two which will likely lead to a conflict that needs to be resolved. Therefore it is not desirable that different users pick the same preamble at the same time.

3.3.2 Message 2

The second message in the random access procedure is the Random Access Response (RAR). The second message is used to inform the user of its relative timing error so that it will be able to synchronize with the cell, scheduling information on when and on what frequency the user may transmit the third message of the process, and a temporary identification for the user. The preamble signature is also included in the RAR so that the user can know whether this specific RAR is directed towards itself or another user. The transmission is done on the PDSCH [9].

In the event that there was a collision in the first step, that is if two users sent the same preamble at the same time, the cell will still only transmit a single RAR as it cannot distinguish between the two signals. Both users are then likely to successfully decode the incoming RAR and proceed to the third step with the same scheduling grant and identification.

The contention-free variant will stop at this step as the remaining steps are superfluous when it is already known that there was no collision.

3.3.3 Message 3

This message which is sent from the user through the PUSCH is the first of two steps in ensuring that there is no contention. The cell identifies the transmitter by the temporary identification assigned to the user during the second message. Additionally a kind of identification that is unique to the user is also part of this message as it is needed in order for the cell to be able to actually tell any contending users apart.

3.3.4 Message 4

The fourth and final message is sent on the PDSCH and contains information about which one user was the intended receiver of the grant given in the second message. It does this by choosing one out of possibly several unique identities received during step 3. If there was no contention then the single user will be identified as the chosen one, and if there was contention one user will be chosen. In both cases the chosen user will receive a message containing their unique identification. Any contending users will upon receiving the same message realize that they were not the intended target of the message and will start over from the beginning of the procedure.

4

System Control Plane

The concept of the system control plane (SCP) was developed in order to solve some problems that have been identified in LTE and to improve the logical structure of the upcoming 5G system. The lack of logical separation between the different kinds of signals in LTE causes a decreased ability to make use of the multiple input multiple output (MIMO) capabilities existing in the system. This is because some broadcasted system information has certain constraints in how the antenna resources are to be used, and these constraints are then also unnecessarily imposed on all other kinds of signals as well. Another issue with LTE is that with time the amount of undefined resource elements has decreased as more and more of the resources are used for new features. Furthermore, the transmission of all mandatory reference signals prevent the power consuming components in the base station from entering sleep mode. The lack of undefined resource elements also makes it overall more difficult to extend the system with future changes.

This chapter gives some background on the SCP and describes the problems, solution, motivations and the expected effects in more detail.

4.1 Definition and Proposed Implementation

The system control plane (SCP) is defined by the set of signals and procedures that are necessary for basic system functionality. This includes, but is not limited to, the the random access and paging procedures, as well as the signals necessary to execute these procedures. Random access is the procedure of a user making initial contact with a cell, and paging is the procedure of a cell locating and sending a message to a user.

Given this definition a SCP could be said to exist in LTE as well. Such a SCP would not be as useful, however, due to constraints specific to LTE. LTE for instance specifies that system information and data should be transmitted using the same antennas, meaning that you would not be able to logically separate the signals of the SCP from the other signals anyway. This prevents the realization of a logical separation of different kinds of transmissions, which is a desired property of the SCP.

Primarily we want the SCP to have two properties, the signals and procedures belonging to it should have an ultra-lean design, and as mentioned previously there should be a logical separation of transmissions belonging to the SCP from all other transmissions. As long as these properties exist, and the definition is adhered to, any implementation would be acceptable.

The implementation that we consider in this thesis contains two different signals, the access information table (AIT) and the system signature index (SSI). The AIT is a table that provides necessary system information to the user, including the information that is necessary for executing the random access procedure, and the SSI is an index to this table which indicates which part of the table is relevant. The SSI also fills the function of giving the users a signal that can be measured in order to calculate the power needed for transmitting the preamble in the random access procedure, something that in LTE was done by cell specific reference signals.

As for the procedures, we are only interested in the implementation of the random access procedure in this thesis. As mentioned earlier, the procedure here uses the SSI instead of cell specific reference signals to measure the power to use when transmitting the random access preamble. All other changes to the procedure, when compared to LTE, varies between simulations and are explained in Chapter 5. These changes consist of different ways of making use of several cells, both when receiving the first message of the procedure and when transmitting the second message of the procedure.

4.2 Static Signals

In LTE there are many different signals that are sent. The amount of different defined data channels described in Section 3.2 alone indicates that there are many types of information that passes between the mobile users and the eNodeBs. When analysing the potential areas of improvement for 5G, researchers have observed that a particular subset of the transmitted signals are responsible for a very large part of the total number of transmissions [5]. This subset consists of all the signals that must be sent regardless of user activity and system state. These signals are called static.

The static subset of the signals are used for the most basic system functionality and include broadcasted system information such as the AIT and the SSI. The

AIT and SSI are always being broadcasted with a set periodicity. The system can not know when new users will appear so it always has to provide the information necessary for the users to be able to initiate random access.

4.3 Dynamic Signals

The set of dynamic signals is the complement of all static signals. The dynamic signals are only sent under certain conditions, when there is an explicit need for it. For example signals that contain data that a user has requested are dynamic.

The SCP also includes dynamic signals. Both the random access procedure and the paging procedure are only executed under certain conditions, when a user needs to gain access to the network and when the network needs to inform a user of something. For this reason both of these procedures consist of dynamic signals.

4.4 Problems in LTE

In LTE there is a scheduler that allocates signals to the time-frequency resources according to the needs of the system. The scheduler assigns resources to the dynamic signals as they appear. The resulting schedule can be seen as a grid of squares, where each square signifies a specific transmission at a certain frequency and time.

The way LTE is designed ensures that the reference signals, described in Section 3.3, are sent as if the system is always operating at maximum capacity, meaning that signals are transmitted all the time and essentially all time-frequency resources are in use. This does of course guarantee a minimum of throughput at all times and simplifies the design. But considering the scarcity of the situations where the system is under heavy load and suffers from lack of time-frequency resources, it can also be viewed as a waste of resources. A consequence of this frequent use of reference signals is the inability for the cell to turn off and initiate sleeping mode. Since the time it takes to turn on the cell again when sleeping already takes up a large part of the time between reference signal transmissions there is not much left for actual sleep, which results in potentially more energy than necessary being used.

Figure 4.1 [10] shows how many empty subframes are used for increasingly high traffic scenarios. The figure represents scenarios of a country wide network, including both rural and urban areas, that are subject to different loads. Going from a scenario where there are no users present at all, the no traffic-scenario, to a scenario with a large amount of traffic, the high traffic-scenario, the number of empty subframes has only decreased by 19%. At first this appears to be a good result, considering that the more free resources we have the better. What causes these numbers, however, is that even when there are no users present in the system there is already a significant amount of subframes that are not empty. So many of the subframes are already used that adding large amounts of traffic only makes a marginal difference in the number of empty subframes. This indicates that much energy is spent in LTE even when there are no users present in the system.

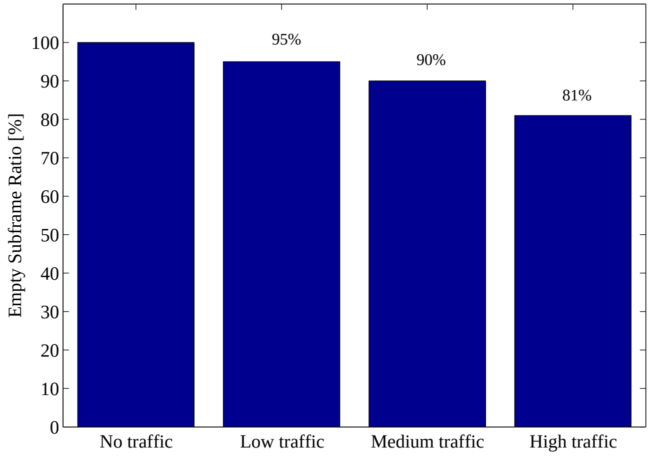

Figure 4.1: *The ratio of empty subframes relative to the total number of subframes for increasingly high traffic scenarios [10].*

Figure 4.2 [10] shows how much energy is spent on static and dynamic transmissions respectively. The scenario is the same as in Figure 4.1, a country wide network with varying amounts of traffic. When there is no traffic 100% of the energy is spent on static transmissions, whereas the high traffic scenario still has more than 90% of the energy spent on the static transmissions. This further supports what Figure 4.1 indicated; even in high traffic scenarios a very large amount of the total energy is spent on static transmissions, of which reference signals are a very large part. The surprising distribution of energy suggests that it would be worthwhile to improve upon how static signals are handled in LTE.

In LTE it is also hard to truly utilize massive MIMO due to the very frequent reference signals. To start with, it is hard to use MIMO for different frequencies at the same time, meaning that having fewer reference signals to transmit is beneficial. It is also troublesome to be able to handle both signals that are received weakly, such as broadcasted signals, and signals that are received strongly, such as those that are sent using for MIMO. This is so because it puts larger demands on the hardware used in the user equipment. While possible it can also be difficult to have all antennas coordinate when you take into account the time it takes

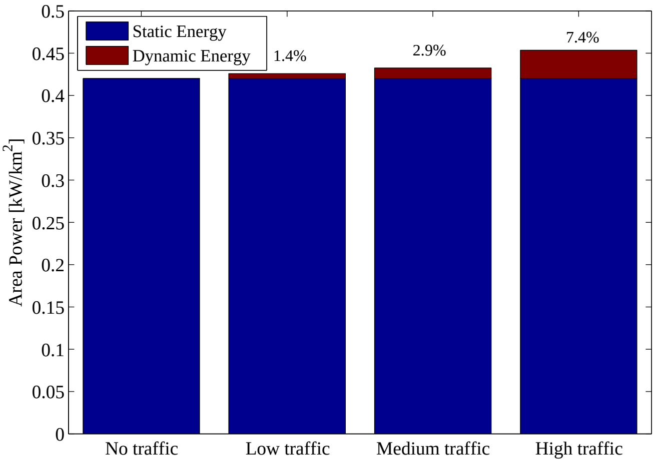

Figure 4.2: *The amount of energy used by the static (blue) and dynamic (red) for increasingly high traffic scenarios [10].*

to change modes and cooperate. In the case of analogue beamforming it is easier to adjust the antennas properly if there is more uninterrupted time than what is currently available due to the frequent reference signals.

Another issue that has become prevalent is the cost of the increased interference between cells that is caused by having more dense networks. As the frequency spectrum is used for many more things than in the past it is convenient to attempt to use higher frequencies due to the lack of free low frequency bands. As shown by (2.2) however, using higher frequencies cause a higher path loss. Therefore to achieve the same coverage we need to make the system denser in the sense that the eNodeBs should be placed less sparsely. This also increases the interference forcing more signals to be retransmitted, thus leading to increased overhead. Another reason for wanting to make the system denser is to increase performance. With more cells there would of course be less users per cell, leading to a lower load on each cell. The increased cost could have been avoided if the overhead did not scale with the increased system eNodeB density.

Yet another issue is the the drawbacks that come as a result of using more advanced transmission techniques. One example of an underutilized technique is antenna tilting. Antenna tilting can be used at an eNodeB to redirect the antennas towards temporary hotspots and increase performance at that location. However, a side effect of this is that the coverage changes. It is easy to imagine the user dissatisfaction if some users were to lose coverage at critical times despite being at a place where there previously was coverage, and by extension it is not hard to see why the operators are unwilling to make use of this technique. Relative to the gain in maximum throughput, the loss of coverage would incur too large of a cost. It would of course be a different matter if there was a way to somehow remove the drawback.

There is also another issue with having a very large part of the resources allocated even when no dynamic data is being transmitted. That is because it leaves very little space for future needs. It is expected that there will be many releases of 5G after the initial one so, not having any resources left to use for future inventions and additions is certainly a drawback. In LTE there is a feature called Multi Broadcast Single Frequency Network (MBSFN) that was originally intended different kinds of media such as TV. In the end the resources allocated to this feature were used for a lot of different and originally unintended things. This supports the fact that having free resources is important for forward compatibility, and free resources is something that LTE does not have in great abundance.

4.5 An Ultra-Lean and Logically Separate SCP for 5G

The SCP is a change that is to benefit the system by making a more fundamental change in the system's structure. Due to the difficulty of making large changes in already existing systems it is easier to include it in a new system rather than trying to accommodate for it in the existing LTE. In particular the SCP is expected to counter the problems mentioned in Section 4.4.

One fundamental difference with the inclusion of the SCP is the logical division of area covering signals, that are necessary for basic system functionality, and user specific signals. The logical division in itself does not change the content of the two signal sets in any way. Instead it only ensures that the signals from the two sets can use two disjoint sets of resources. This division of resources is not necessarily fixed, the main point is rather that it should be possible to configure the system so that the static signals and the dynamic signals can operate as they wish without obstructing the transmission of the other set of signals.

Further the SCP will change both the contents and the schedule of the static set of signals. For instance, instead of sending reference signals every 0.25ms the SCP will decrease this periodicity by several times. With the SCP the static set of signals will only contain AITs tables and SSIs. The AIT could be transmitted more rarely, perhaps only once every several seconds, as the data it contains is valid for large areas and is not likely to change. The system signature index should on the other hand be transmitted a bit more frequently as different indexes might be transmitted from adjacent nodes.

The idea is to remove a large part of the allocated resources and only broadcast what is necessary to gain access to the system. All other things will be handled on a need basis. While the SSIs can be measured and used to estimate for instance the path loss, it is only intended to be used for the random access. Even though the SSIs might be useful to other functions as well these other functions will need to make use of their own dedicated and dynamic measurement signals. This is in order to ensure that the system does not transmit any more SSIs than necessary, which in turn brings down the amount of static signals that are transmitted.

Furthermore the SCP will allow nodes that are placed near each other to transmit the same SSI. This is different from LTE in which adjacent cells need to strictly send different reference signals and system information. Because of this the users do not need to differentiate between cells as much during the initial access phase. Instead, it might be that all nearby cells transmitted the same SSI. In such a scenario the user would in fact not be able to tell how many cells it received signals from. This is because the signals are sent with the same timing from the cells and are then superimposed upon each other in the air.

4.6 Benefits of the SCP

By reducing the amount of resources allocated to static transmissions the energy consumption of the eNodeBs can be significantly reduced. Considering how a large fraction of the energy in LTE is spent on the static signals, as shown in Figure 4.2, this is potentially a large gain relative to the effort.

When adjacent cells can send the same SSI these will add constructively at the receiver, rather than interfere with each other. In addition the reduced amount of static signals will also reduce the interference for other, dynamic, signals. This will to an extent counter the effect of increased interference caused by making the networks more dense. Less interference is very useful as interference increases the likelihood that transmissions fail. Failed transmissions will eventually need to be retransmitted resulting in increased overhead.

Additionally, logically separating the static and dynamic transmissions allows for increased usage of more dynamic techniques such as the antenna tilting mentioned in Section 4.4. Since the static and dynamic transmissions do not share antenna resources it is possible to utilize the technique for the dynamic transmissions only without affecting how the antennas used for the static transmissions are used. This way all users are still guaranteed the basic functions, such as the random access, as they are provided by the static transmissions, even though the coverage of the dynamic transmission may be temporarily decreased.

A strength of this new solution is also its flexibility. As has already mentioned it is possible to have all cells transmit the same SSI, but any configuration is possible. If found necessary it is even possible to emulate LTE by making all adjacent cells transmit different SSIs, which would be equivalent to how LTE systems have adjacent cells transmit different reference signals.

5

Simulator Setup

The presented results were generated based on two different methods. The first was to utilize an Ericsson-internal simulator to calculate path loss values between the cells and randomly placed users. This data was then used in MATLAB to calculate the expected results for some interesting scenarios. The second method was to run the simulator using a similar setup and log the different results of the participating users.

This chapter describes the two processes in detail and provides information about the setups of the simulations.

5.1 Path Loss Simulations

In order to generate the first results the simulator was used to generate a map consisting of nodes, cells and users. 1000 users were placed randomly on the map and their positions recorded. Then the path losses between each pair consisting of one cell and one user were generated. The gathered data was used to calculate the expected results of some interesting situations.

The inter site distance varied between simulations while all other parameters were held constant. Relevant parameters are presented in table 5.1.

The deployment is such that the nodes are placed in a hexagonal network, approximately as shown in Figure 5.1. The deployment supports wraparound, meaning that all cells will be treated as if they were in the middle of the map. Nodes on the edge of the map will simply wrap around to the opposite side of the map to find their neighbour.

Table 5.1: Data describing the deployment used in the path loss simulations.

Parameter	Value(s)
`Number of nodes`	19
`Number of cells per node`	3
`ISD`	{500,1732,2500,5000,9000,15000}m
`Carrier frequency`	2GHz
`Bandwidth`	5MHz

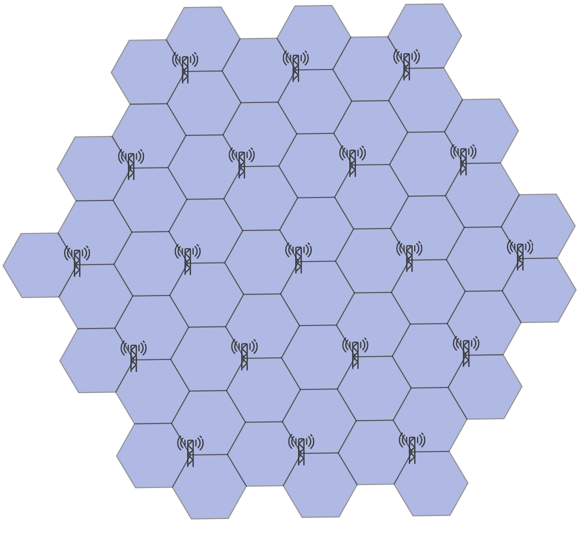

Figure 5.1: The deployment used in the path loss simulations, 19 nodes and 57 cells in total.

The simulator differs between urban and rural scenarios. In the urban scenarios the indoor users suffer a path loss penalty, the path loss increases more quickly as the distance increases and the multiple paths are calculated differently. The present simulation uses an urban scenario defined by the 3rd Generation Partnership Project (3GPP) [3]. All users are indoors in this simulation.

For the 5G test systems two different ways to receive the first message of the random access procedure, coordinated and uncoordinated reception. The perhaps most straightforward method would be to have each cell individually attempt to decode the message, exactly the same way as is done in 4G but for several cells simultaneously. As long as one cell is able to decode it then the reception can be considered to be a success. This is because cells are able to synchronize information without the need for wireless communication, and while a communication cost still exist it should be relatively cheap and can be ignored in this context. Therefore, if at least one cell successfully decodes the preamble the system can proceed to the next step, the transmission of the random access response, in an appropriate manner. The probability of error in this case is given by the equation

$$P_e = \prod_{i \in I} f\left(\frac{p}{P_i N_0}\right), \tag{5.1}$$

where P_e is the error probability, P_i is the path loss between the user and the ith cell, N_0 is the thermal noise, I is the set of all cells, p is the power used to transmit the random access preamble and f is a function that takes SNR as input and returns an error probability. We call this uncoordinated reception of the random access preamble.

The opposite of this method would be to combine the received signals prior to the actual decoding attempt. The summation of the signals would have a larger SNR than each of the component signals, causing the decoding of the combined signal to have a larger chance of success. However, only one such chance exist as it uses all of the available information. The error probability for this method can be described by the equation

$$P_e = f\left(\sum_{i \in I} \frac{p}{P_i N_0}\right), \tag{5.2}$$

where P_e is the error probability,P_i is the path loss between the user and the ith cell, N_0 is the thermal noise, I is the set of all cells, p is the power used to transmit the random access preamble and f is a function that takes SNR as input and returns an error probability. We call this coordinated reception of the random access preamble.

5.1.1 Metrics

The different metrics that are calculated in MATLAB are:

- The power used by users to transmit the random access preamble according to (3.1).
- The SNR that cells received the preamble with. The SNR is defined as

$$\frac{\int_{-\infty}^{\infty} |x(t)|^2 \, dt}{\int_{-\infty}^{\infty} |n(t)|^2 \, dt}, \tag{5.3}$$

  where $x(t)$ is the signal component of the received signal, $n(t)$ is the noise component of the received signal and t is the time.

- The expected error probabilities for decoding the preamble in two different ways. The first is by having each individual cell independently decode the preamble that they received, uncoordinated reception. The second is by having a single cell decode the sum of all received preambles, coordinated reception. This metric is defined as

$$\frac{a}{a+b}, \tag{5.4}$$

  where a is the number of preambles that were successfully decoded by the system and b is the number of preambles that were not successfully decoded. In the case of uncoordinated reception the system is considered to have successfully decoded the preamble if at least one cell decoded it.

5.2 Performance Simulations

The second set of results consists of simulations of users that are trying to access the system through the contention based random access procedure. After a user succeeds it also transmits a block of eight bits before it is considered done. The user is then terminated. The only present interference is caused by transmissions that belong to instances of the random access procedure that have been initiated by other users. that which is caused by the other users and cells that send messages belonging to the random access procedure.

The simulations start with zero initial users, and users are continuously created according to the parameter intensity. The intensity states how many users per second are created. The time between the creation of each user is uniform, and the users are also stationary. The deployment is hexagon-shaped with wraparound. In total there are 7 eNodeBs with 3 cells in each, approximately as shown by Figure 5.2.

These simulations are all done in the 3GPP typical urban scenario [3].

Of the two ways of decoding the preamble described in Section 5.1.1 only the uncoordinated reception is simulated in the performance simulations. Instead two 5G candidate systems that differ in the way the random access response (RAR) is

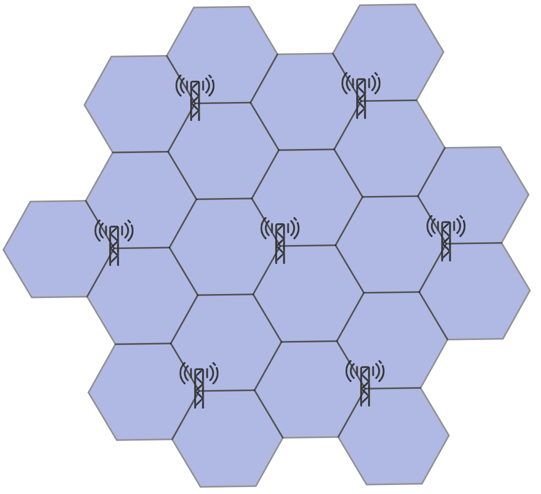

Figure 5.2: *The deployment used in the performance simulations, 7 nodes and 21 cells in total.*

Table 5.2: *Data describing the deployment used in the performance simulations.*

Parameter	Value(s)
Number of nodes	7
Number of cells per node	3
ISD	{500,1732,2500,5000,9000}m
Time duration	5s
Carrier frequency	2GHz
Bandwidth	5MHz
Number of preambles	{64,128,256}

transmitted will be simulated. In one of the 5G candidate systems all cells that received a preamble will also send a RAR for this preamble, we call this uncoordinated transmission of the RAR. The other candidate system coordinates the transmission. Of all cells that receive the same preamble at the same time only one cell, the one with the best channel quality towards the transmitter of the preamble, will send a RAR. We call this coordinated reception of the RAR.

The parameters that are primarily considered during the performance simulations are the intensity, the number of available random access preambles and which candidate 5G system that was in place. Relevant parameters are shown in table 5.2.

5.2.1 Metrics

Only small changes are done in the parameter values but different metrics are measured. These metrics are:

- The time it takes for a user, after the it has been created, to perform the random access procedure and send an eight bits large message.
- The power used for transmitting the preamble according to (3.1).
- The number of cells that detected the preamble sent by a user.

6

Results

Several different setups have been used when gathering the data. Most of them are very similar in that they make use of a hexagonal deployment and wraparound. The inter site distance (ISD), the number of preambles in use and which system is being simulated are the most common varying parameters.

This chapter presents the acquired results. Either by using the data gained from simple simulations to do post processing in MATLAB or through the more advanced performance simulations.

6.1 Post Processing Results using Path Loss Estimates from Simulations

The first set of results were produced through calculations in MATLAB using data generated in the simulator. The data consists of path loss estimates between cells and randomly placed users. These results are interesting by themselves, but they were also used in order to identify interesting areas of investigation for the second part of the results, the performance simulations. The simulations in this section used the setup described in Section 5.1

6.1.1 Message 1: Random Access Preamble - Transmission

The contention-based random access procedure is always initiated by a user transmitting the first of a chain of messages, the random access preamble. The power used when transmitting this message is determined by equation (3.1) which depends primarily on the estimated path gain between user and eNodeB. The path gain is estimated using reference signals.

In 4G each user listens for reference signals from only one cell, the reference signal is cell-specific, and thereafter only communicates with that cell. Meanwhile in the 5G test system all the cells will synchronously transmit the same reference signal and the users will receive the sum of these but will perceive it as just one signal. This will cause the 5G test system users to sometimes use a different power than the 4G users would use since the perceived path loss is different in the two cases. Furthermore, several cells will be able to receive the sent preamble in the 5G test case whereas only one cell will be receptive of the preamble of a single user in 4G.

A comparison between the power used in 4G and the new experimental system can be found in Figure 6.1. The figure shows a cumulative distribution function (CDF) of the power that randomly placed users calculated that they should use to transmit the random access preamble. The graphs include several pairs of differently styled red and black lines, each pair corresponding to a certain ISD. The red lines correspond to the 5G test system and the black lines correspond to the 4G system.

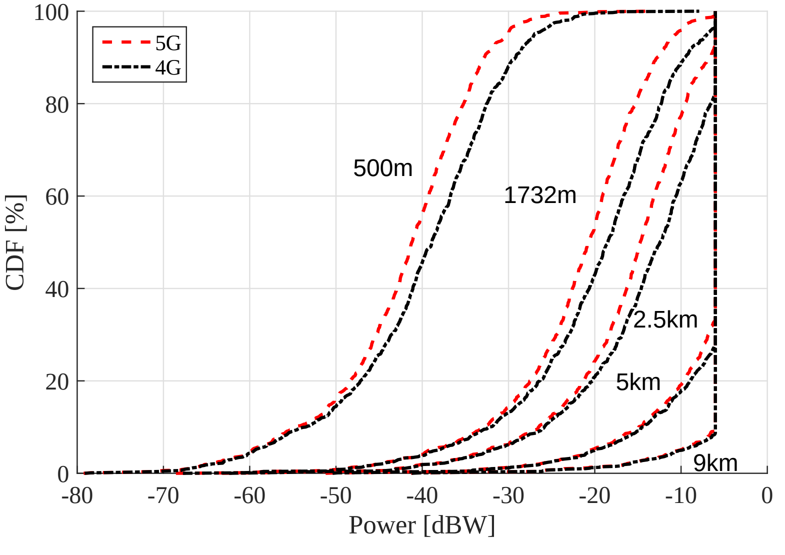

Figure 6.1: *A CDF of the power a user is expected to use when transmitting the random access preamble according to* (3.1)*. ISDs from 500m to 9km. Red lines show the 5G test system while black lines show the 4G system.*

It can be seen that in all cases the red line is further to the left than the corresponding black line, at least for large parts of the lines, which signifies that in general the 5G test systems' users recognized that it was suitable to use a lower amount of power in the 5G test system. On average the 5G candidate system's users used about 2 dB less power than the users of the 4G system. Both the users in the 5G test system and the users in the 4G reference system uses (3.1) to calculate the power to use. For the different randomly placed users in the two systems the only variable that actually differed between different users was the path loss. Therefore if less power was used in the 5G test system then it means that the 5G test system's users overall experienced a smaller path loss in the received reference signals. Considering that the reference signals are transmitted with the same power in both the test and reference system this result is the effect of the 5G test system users being able to combine the reference signals from more than one cell.

None of the lines, each corresponding to a single simulation for either the 5G test system or the 4G system, reach a power higher than −5 dBW. This is because the users at that point can not use more power as they are limited to this value in the simulations.

Except for users at −5 dBW there are cases when the 5G system and the 4G system performs similarly and cases where they do not. The intervals where the two systems show similar performance correspond to the cases when a user was located very close to an eNodeB. In this case all other eNodeBs are about as far away as they can get, meaning that the reference signals transmitted from those far away eNodeBs have much less of an impact. The opposite, when there are large differences between the red and the black lines, correspond to the cases where users in the 5G test system were placed in locations that were about equally far from the closest eNodeBs. In such a case the contribution of the reference signals in the user is not dominated by the reference signal from a single eNodeB, but rather divided fairly equally between two or more eNodeBs.

If the users use power according to (3.1) when transmitting the preamble, as they have been doing here, there will be a difference in how many of the preambles that would be successfully received. This is both due to the observed fact that the 5G test system has the users use less power and due to the 5G test system having several eNodeBs attempt to receive the message. These two effects counter each other to an extent. The next subsection examines which of these two effects it is that dominates the other.

6.1.2 Message 1: Random Access Preamble - Reception

Figure 6.2 presents a CDF of the SNRs at the different receiver cells when a single user transmits a random access preamble, the ISD is 1732m. Each line in the graph contains a single point that corresponds to a single user. In the 4G system there is only one cell that is listening for a preamble from the user, for which reason the graph only contains one line for the 4G system. In the 5G test system, however, there is no set limit for how many cells are listening for a preamble from the same user. In this graph only the three cells with the best SNRs are shown, as the remainder of the cells typically perceived the preamble with a very low SNR, meaning that they are very unlikely to be able to decode the message.

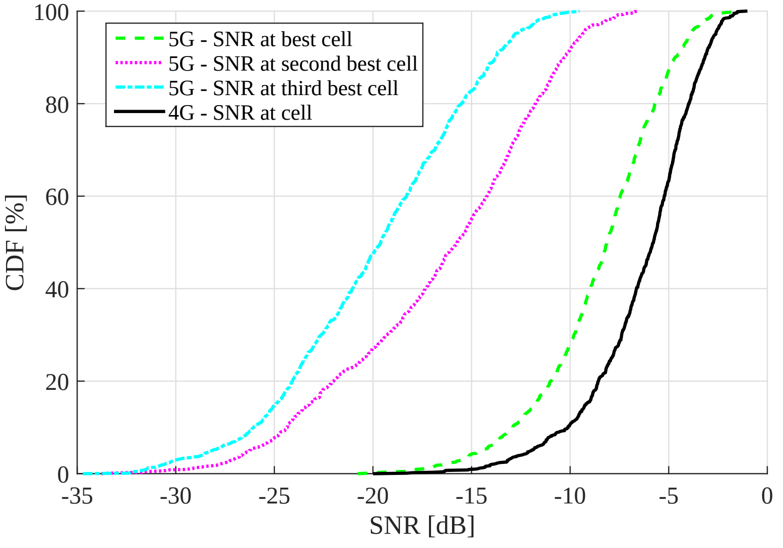

Figure 6.2: *A CDF of the SNR of the preamble in different cells for the 5G and 4G systems with fading. The black line shows the SNR at the single receiving cell in the 4G system. The green, purple and blue lines show the SNR at he best, second best and third best cells in the 5G test system. ISD is 1732m.*

At a distance of 1732m almost all users estimate that they will be able to transmit their preamble in such a way that the receiver can receive it with a SNR of -5 dB, meaning that almost no users were forced to use the maximum allowed power during the transmission of the preamble. Because of this the results for the 4G system, represented by the rightmost black line, are centred around -5 dB. If the SNR calculation did not take the fast fading into account almost all cells of the 4G system would have exactly -5 dB, as shown in Figure 6.3 that except for the lack of fading uses the same simulation setup. With fast fading however, as Figure 6.2 shows, the results are spread out with the median result being at about -6 dB.

In the graph with fading, Figure 6.2, it can be seen that the green line, corresponding to the SNRs at the best cell for the 5G test system, has approximately the same shape as the black line. The difference between the two is that the green line has a SNR value that is about two dB lower at the same percentile. An observation is

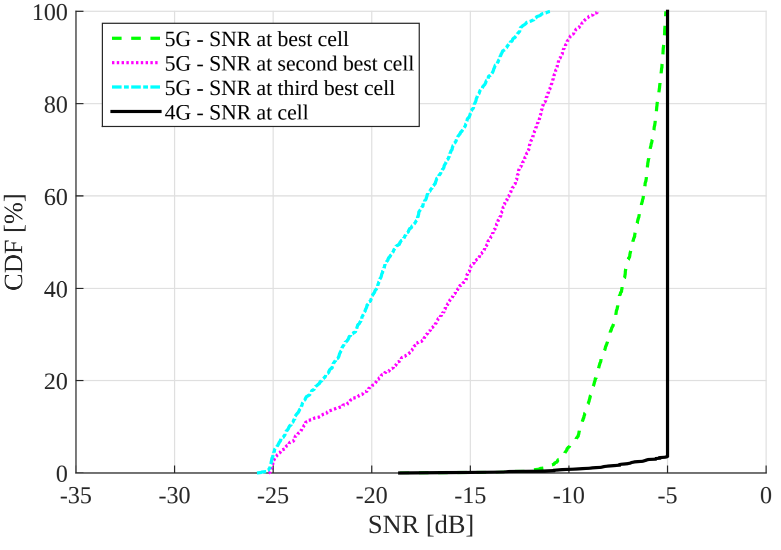

Figure 6.3: *A CDF of the SNR of the preamble in different cells for the 5G and 4G systems without fading. The black line shows the SNR at the single receiving cell in the 4G system. The green, purple and blue lines show the SNR at the best, second best and third best cells in the 5G test system. ISD is 1732m.*

thus that in the large majority of the cases the cell in the 4G system receives the preamble with a higher SNR than any single cell in the 5G test system does. The second best and the thirds best cell in the 5G test system appear to have SNRs about ten and fifteen dB smaller than the single receiving cell in the 4G system, respectively.

For the cells in the 5G test system it should be noted that values at the same percentiles are not necessarily comparable. For instance, if the best cell in the 5G test system receives the preamble with a SNR of −5 dB then the second and third best cells are likely to have received the same preamble with SNRs around −25 dB. That is, high values in the best cell are likely to correspond to low values in the rest of the cells. This is because a cell will only have a high SNR if it is close to the user, and if one cell is close to the users the other cells will be further away from the user than normal, resulting in lower SNRs at these other cells. The same is also true for the inverse. If the best cell receives the preamble with a low SNR, relative to the typical SNR at the best cell, the other cells are more likely to have received the preamble with relatively high SNRs.

Overall, no single cell in the 5G test system is able to gain the same SNR as the cell in the 4G system. By somehow combining the results from the different cells in the 5G test system, however, it is possible to improve the systems ability to decode the preamble.

Figure 6.4 shows a histogram of the number of cells that successfully decoded a user's preamble, in this case the ISD was 500m. Whether a cell is able to decode the preamble or not is dependent on the SNR of the received signal at the receiver.

In the 5G test system there were about 50% less cases where no cell was able to detect the preamble compared to the 4G system. This is a not insignificant amount considering that a failure here will force the user to retransmit the preamble, slowing down the procedure. In the large majority of the cases in the 5G test system, the number of successful cells varied between zero and five.

Figure 6.5 shows the mapping from SNR to error probability that was used to calculate the results in Figure 6.4. Having calculated the SNRs, the numbers in Figure 6.4 were generated using the probabilities gained through the mapping. The same mapping is used when generating the remainder of the results in this section.

In the 4G system it is quite straightforward how to process the received preamble. Either the cell is able to interpret it correctly, in which case the procedure continues to the next step, or it can not. If the cell fails the user will simply have to start over and retransmit the preamble. For the 5G test system it is not so simple however. Several cells will have received the preamble and in order to achieve acceptable results the received signals at the different cells will need to be combined somehow. The two ways to do this that are used here, coordinated

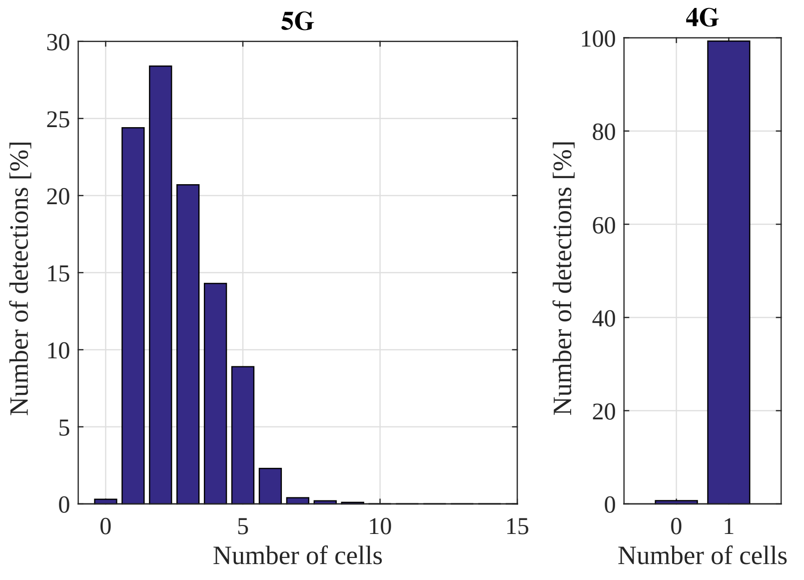

Figure 6.4: *Histograms of the number of cells that successfully decoded each preamble for the 5G test system (left) and for the 4G system (right). ISD is 500m.*

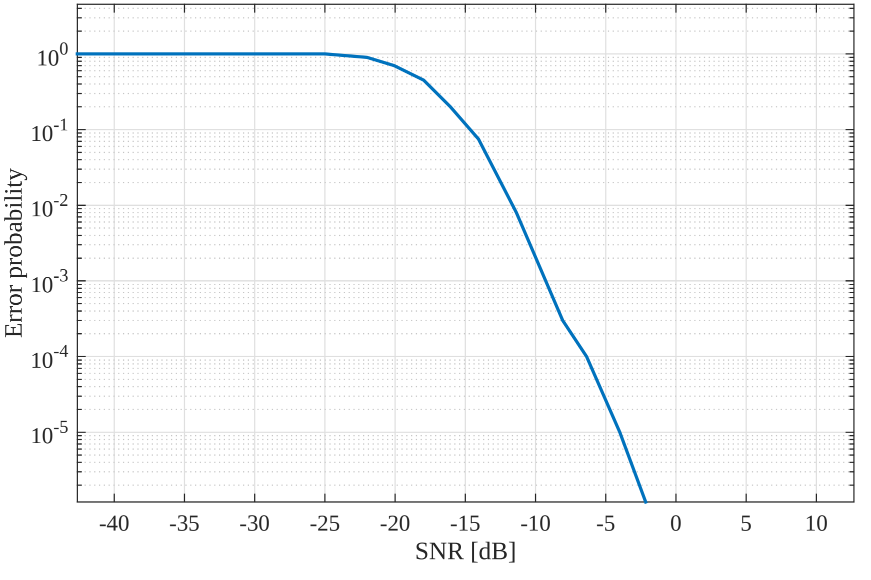

Figure 6.5: *A mapping of SNR to error probability used to generate the results in Figure 6.4, Figure 6.6 and Figure 6.7.*

and uncoordinated reception, are explained in Section 5.1.

Which method would yield the best results is hard to predict due to the irregular nature of the relation between the signal to noise and interference ratio (SINR) and error probability. Figure 6.6 shows a CDF of the estimated decoding error probabilities for different randomly placed users for the 5G test system and the 4G system. This figure represents the case of uncoordinated reception. For ISDs at 5km and below this method gives lower error probabilities for the 5G test system than for the 4G system. For ISDs at and larger than 9km the results appear to be equal.

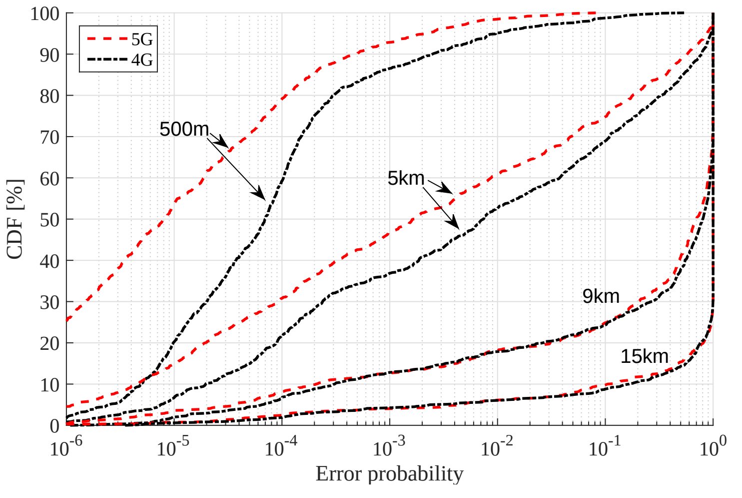

Figure 6.6: *The error probabilities for detections of the preamble using uncoordinated detection for different ISDs. The 5G test system and the 4G system are represented by red lines and black lines respectively.*

Similarly, but for the case of coordinated reception the situation is shown in Figure 6.7. For ISDs between 5km and 9 km this method shows a slightly better performance compared to both the uncoordinated method and the 4G system. For other ISDs the coordinated and uncoordinated methods performs similarly. The uncoordinated method has a larger amount of users with extremely small error probabilities compared to the coordinated method. This is not important, however; as long as the error probability is below some fairly small value such as 0.1% there is almost no gain from further reducing it. Therefore the coordinated method can overall be said to perform slightly better than both the uncoordinated method and the 4G reference system.

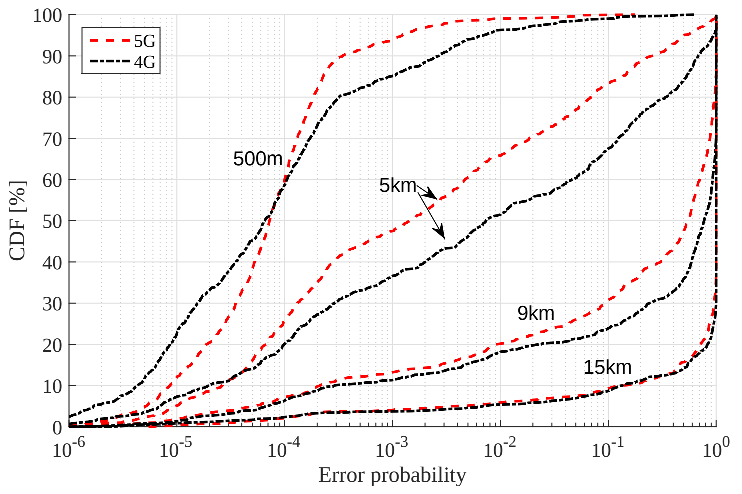

Figure 6.7: *The error probabilities for detections of the preamble using coordinated detection for different ISDs. The 5G test system and the 4G system are represented by red lines and black lines respectively.*

6.2 Random Access Delay Simulations

The second set of the results, the random access delay simulations, were created by running more extensive simulations while logging a desired set of parameters. These results give a clearer view of how the 5G test system compares to 4G as a whole with respect to the random access procedure. The simulations in this section used the setup described in Section 5.2.

6.2.1 Message 1: Random Access Preamble

Figure 6.8 shows a CDF of the power in decibels of watts (dBW) that was used by randomly placed users. This simulation made use of a deployment of 7 eNodeBs with each consisting of 3 cells. This is different from setup used for the MATLAB calculations, showed in Figure 6.8, which had 19 eNodeBs with 3 cells each. The powers used by the users show the same distribution in both cases for all simulated ISDs, but all users in the performance simulation make use of about 3 dB more power then in the MATLAB case. A part of this is caused by the different amount of cells. The more cells there are the more reference signals are being transmitted which cause the users to receive a stronger signal. As a result they will transmit the preamble with less power. It is likely, however, that this difference is also caused by small differences in the parameter settings.

Figure 6.9 shows the frequency of the amount of receiving cells for the preamble.

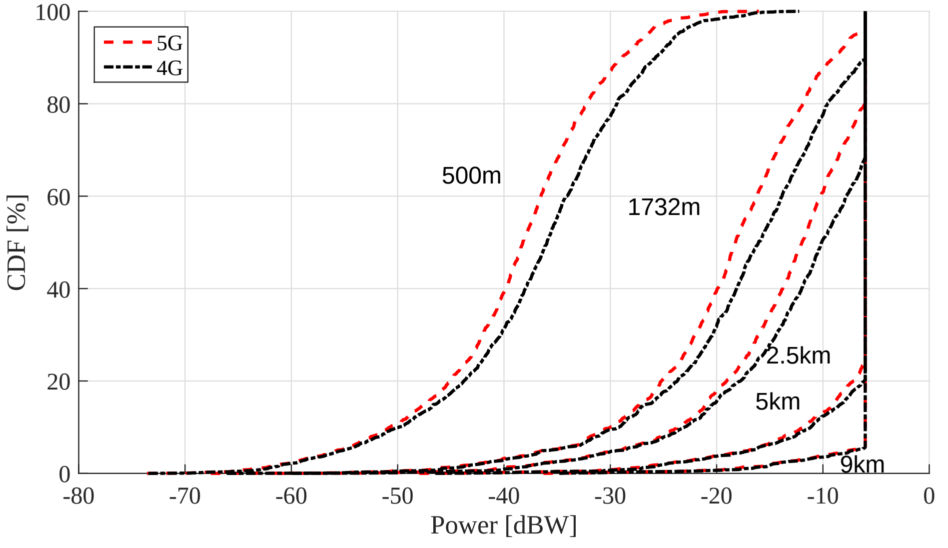

Figure 6.8: *A CDF of the power used when a user transmits the random access preamble according to* (3.1). *ISDs from 500m to 9km. Red lines show the 5G test system while black lines show the 4G system.*

In this simulation the users were created with an intensity of 9000 users/s, but contrary to what is stated in Section 5.2 only for a time period of approximately 200ms. For this reason the results represent the case when the system is not yet completely congested, meaning that the interference between signals is comparatively lower. For a majority of the users there were between 1 and 5 cells that managed to receive and decode the preamble. This is a rather large difference compared to results from the MATLAB calculations. The cause of this is the observed difference in transmission power. As more power is used the preamble will reach further, allowing more cells to successfully decode it.

6.2.2 Performance Simulations of Complete Random Access Procedure

The simulations in this section have all been run using uncoordinated reception of the random access preamble. Considering that the results of the uncoordinated and coordinated reception was quite similar the uncoordinated method was chosen as preferable due to its smaller complexity. The smaller complexity is due to no communication being needed in the actual reception and decoding of the signal.

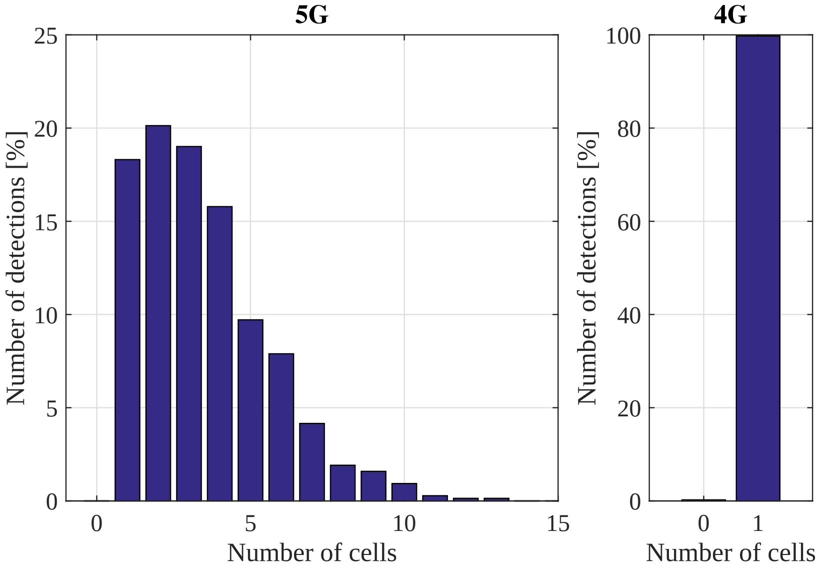

Figure 6.9: *The number of cells that detected the preamble for the 5G candidate systems (left) and the 4G system (right). ISD is 500m.*

The transmission of the random access response (RAR) can, just like the reception of the preamble, be done in several different ways. In the following simulations two alternatives have been considered. The first, uncoordinated transmission of the RAR, makes all the cells that successfully decoded the preamble also transmit a RAR. The second method, coordinated transmission of the RAR, makes all the cells coordinate and the system chooses the cell which it believes has the highest probability of successful transmission of the RAR, based on the estimated path loss.

Figure 6.10 shows the time it took for the 90th percentile of the users, with respect to time taken, to successfully complete the random access procedure and send the message of 8 bits. For all intensities the 4G case show a superior performance compared to the two 5G candidate systems. Already at intensities of 15000 users/s both of the 5G candidate systems have delays more than twice as large as the 4G system for the users at the 90th percentile. At the relatively low intensity of 5000 users/s the 4G system and the 5G candidate system with uncoordinated transmission of the RAR perform similarly, while the 5G candidate system with coordinated transmission of the RAR has an additional delay of 240ms for users at the 90th percentile.

While it is hard to tell what level of performance can be considered acceptable, the results for the 5G candidate systems seem unsatisfactory. A way to improve them might be to increase the number of preambles that the systems can make

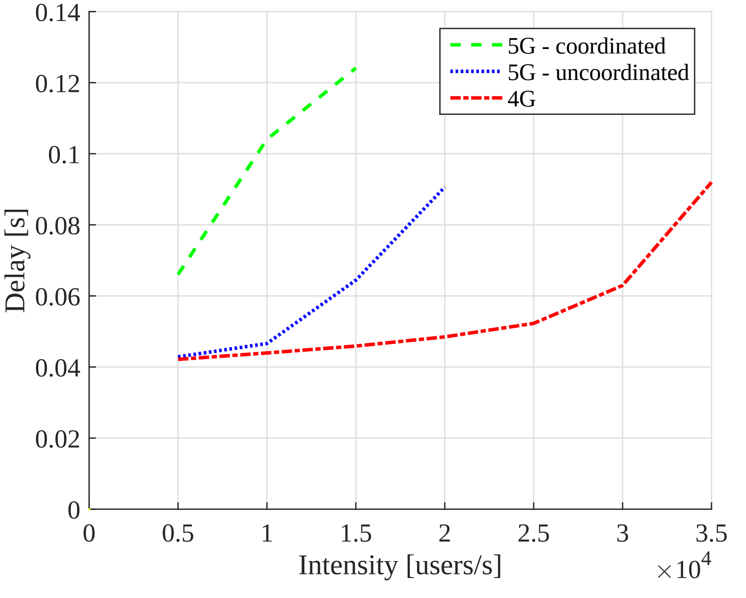

Figure 6.10: *The delay of completing the random access procedure and send a message of eight bits for the users at 90th percentile at different intensities. The three lines correspond to 5G with coordinated transmission of RAR (green), 5G with uncoordinated transmission of RAR (blue) and 4G (red). ISD is 500m.*

use of. In the 5G candidate systems the users are not restricted to a single cell during the random access procedure, and therefore it is likely that they will interfere with users that are physically located in other cells by sending the same preambles. This scenario would not occur in 4G.

In 4G each cell has a set of 64 preambles that it makes use of. The adjacent cells, however, make use of different sets of preambles. The result is that users from adjacent cells essentially never transmit the same preamble, avoiding any confusion. However, while no single cell makes use of more than 64 preambles the total number of different preambles in a larger area is larger than 64. Therefore it might actually be more fair to also increase the number of preambles that the 5G candidate systems, that does not have the cell distinction for random access, can use. Depending on how many nodes in the system transmit the same system signature index (SSI) the number of preambles in the 5G candidate systems could be increased by several times and still be smaller than the total number of preambles in a 4G system of the same size. This is especially true for the 5G candidate systems that are investigated here, as all cells transmit the same SSI, leading to the complete lack of cell distinction during the random access procedure.

Figure 6.11 shows the total time taken for users to complete the random access procedure and send the package of eight bits. In this case 4G is compared to 5G candidate systems using the uncoordinated method for transmission of the RAR, with a varying number of preambles. As the number of preambles increase the performance of the 5G candidate system approaches that of the 4G system. At an intensity of 20000 users/s the 5G candidate systems with 64, 128 and 256 preambles have about 100%, 40% and 20% more delay than the 4G system respectively. At intensities lower than 10000 users/s there is essentially no difference between the systems however.

Figure 6.12 shows the same scenario with the same setup with the exception that the coordinated transmission method is used for the RAR in the 5G test system instead. At an intensity of 15000 users/s the 5G candidate systems with 64, 128 and 256 preambles have about 150%, 100% and 30% more delay than the 4G system respectively.

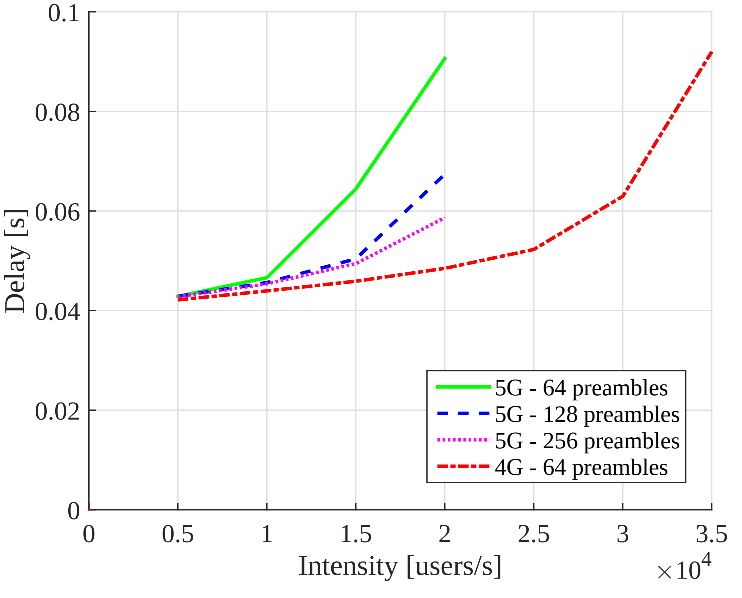

Figure 6.11: *The delay of completing the random access procedure and send a message of eight bits for the users at 90th percentile at different intensities. 5G candidates with uncoordinated transmission of RAR with 64 to 256 preambles are compared to 4G with 64 preambles. ISD is 500m.*

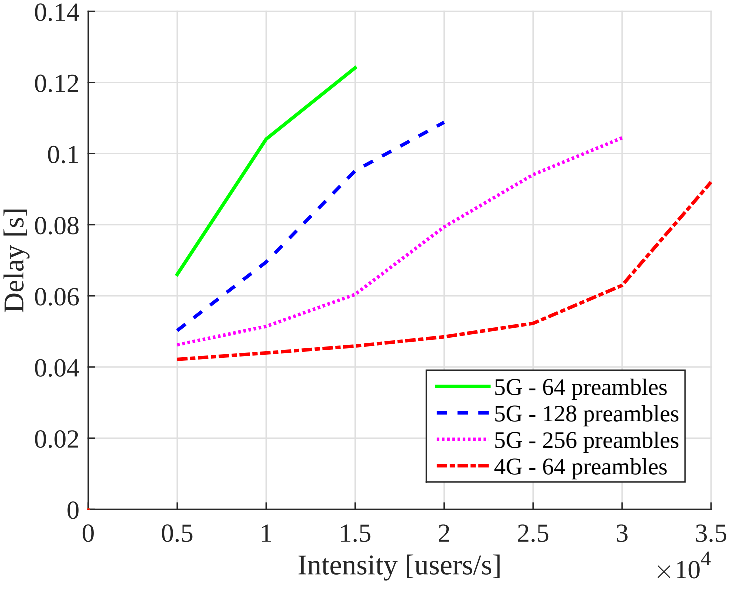

Figure 6.12: *The delay of completing the random access procedure and send a message of eight bits for the users at 90th percentile at different intensities. 5G candidates with coordinated transmission of RAR with 64 to 256 preambles are compared to 4G with 64 preambles. ISD is 500m.*

7 Discussion

This chapter discusses the presented results, the method and some related societal aspects.

7.1 Results

The results were acquired through two different methods. Although there is an overlap in the data which they present they do not completely agree with each other. There is for instance a difference in the power used for transmitting the preamble, as shown by Figure 6.1 and Figure 6.8. The general trend appears to be that the results acquired through the complete simulations of the random access procedure resulted in smaller delays, but at the expense of additional power usage. This difference might be a result of a simplified implementation of the simulator, but it could also be caused by errors in the implementations of the calculations done in MATLAB. For this reason the results acquired through the supposedly theoretically sound calculations should be regarded with some doubt. The comparisons are, however, relative in each case, and the conclusions should therefore still be valid.

An interesting fact is that the calculations done in MATLAB seem to indicate that the 5G candidate systems are better than the 4G system in most aspects, while the performance simulations seem to show the opposite. The MATLAB calculations that indicate that the 5G candidate systems have a higher ability to receive the first message in the random access message, the preamble, seem believable. After all, the 5G candidate systems make use of more signals and have diversity due to the larger number of usable reference signals and cells that are available to a single user. The performance simulations do not necessarily contradict this, but they do indicate that the 4G system requires less time for users to complete the random access procedure. Therefore it might be that it is the part of the random access procedure that has not been explicitly looked at that causes this reversal in performance.

It is mainly the third and fourth message of the random access procedure that have not been specifically investigated. The reason for this is that we have modelled the transmission and reception for these messages in the same way in both the 5G candidate systems and in the 4G system. In both cases the user will transmit the third message to a single cell, and this same cell will alone respond with the fourth and final message. A possibility is that the interference caused by the additional amount of random access responses (RAR) transmitted in some of the 5G candidate systems could cause this by increasing the failure rate of the subsequent messages. Comparing the delays of the 5G candidate systems with coordinated and uncoordinated transmission of the RAR this does not appear to be the case, however. It is actually the system with uncoordinated transmission, the system that transmits relatively many RARs, that takes less time to complete the random access procedure.

The most plausible reason for the performance degradation in the 5G test systems therefore appears to be the second message. As for the candidate system with coordinated transmission of the RAR the degradation is likely due to how the system handles collisions. Using the coordinated algorithm the whole system will only transmit a single RAR for each received preamble at a certain time instance, regardless of how many collisions there are. This behaviour is really only suitable for very small systems that are not much larger than a cell. For larger systems large amounts of received preambles are likely to be dismissed, even when the colliding users are in completely different areas, forcing many users to unnecessarily restart the process. The uncoordinated method does not suffer from this deficiency as RARs will be transmitted from all cells that received a preamble. Instead, the problem here might be connected to the sheer number of RARs being transmitted. Each RAR requires some control information to be transmitted on the physical downlink control channel (PDCCH), and so the large number of RARs might cause this channel to be congested.

Something that, not surprisingly, appears to improve the performance of all 5G candidate systems is the usage of a larger amount of preambles. The more preambles there are the less likely it is that there will be collisions during the contention resolution of the random access procedure. There is a limit to how many preambles can be feasibly used however, as it becomes harder to tell the preambles apart the more of them there are. At some point the cost will outweigh the gains, so a moderate amount of preambles is likely better.

7.2 Method

The method used in this study was essentially to first make calculations of some interesting situations, acquiring some more theoretical results, and then run simulations of users performing the complete random access procedure. Hints of what to look for in the complete simulations were gained through the results of the calculations.

Mainly the procedure as a whole and the first message, and to some extent the second message, have been subject to investigation. While it is necessary to investigate the system in terms of some metric such as total delay of the random access procedure in order to determine which system performs better, there are weaknesses to this approach. At this point we are for instance able to claim that the 4G system requires less time for the random access procedure. This is not very helpful with regards to finding measurements to improve the procedure in the 5G candidate systems. Had we put more focus on investigating the success rate of the individual messages, and the causes of the success rates, we would have been better equipped to suggest changes.

The whole investigation has been heavily reliant on an Ericsson-internal simulator. This simulator contains large amounts of code that is difficult to interpret. Constructing a new simulator only for the purpose of testing the random access procedure is not reasonable considering the huge amount of work necessary, but a result of using the already existing simulator is that the acquired results are less trustworthy. It is reasonable to assume that the implementation of the 5G candidate systems contains at least a few bugs. The original simulator is also likely to contain some small bugs that might affect the result. This makes it more difficult to replicate the result for those who do not have access to the same simulator.

Most of the sources that have been referred to in this report are either primary sources, for instance [8] and [12], or are related to reputable sources, such as [4], [5] and [6] that are results of the well known EARTH project [1] that involves reputable parties such as Ericsson. For this reason the sources are overall deemed trustworthy.

7.3 Societal Aspects

Having a quick way for users to access the system is a necessity for 5G, so a reasonably quick random access procedure is necessary if the system control plane (SCP) is to be included. This study investigates the issue of the random access procedure with the SCP in place, and might thus contributes to the inclusion of the SCP.

As noted in Chapter 4, one expected improvement that the inclusion of the SCP will yield is the potentially large reduction in power usage. This is arguably beneficial for society as a whole, considering that this would make the wireless systems cheaper to operate which by extension makes the technology more accessible.

8 Conclusions

This chapter summarizes and draws conclusions about the results, relating the results to the problem formulation. This chapter also states what was omitted from the investigation due to lack of time, and suggests what could be done in order to further investigate the topic.

8.1 Summary

The power used for transmitting the preamble is larger for the 4G reference system than for the 5G candidate systems. The difference varies between 0 and 3 dBW with an average of about 2 dBW. All users in the 5G candidate systems investigated transmit the preamble with equal power.

It has been observed that the reception of the preamble is equally good or better in the 5G candidate systems compared to the 4G reference system. Calculations proved that both of the systems with coordinated and uncoordinated reception of the preamble perform better than the 4G reference system. For inter site distances (ISD) of 5km and smaller both 5G candidate systems with coordinated and uncoordinated reception of the preamble have a larger probability of successfully decoding the preamble than the 4G reference system. This shows that the error caused by using (3.1) in the 5G candidate systems is smaller than the gain of using several cells to receive the preamble. For larger ISDs the differences between the candidate systems and the reference system decrease.

The delay of performing all the steps in the random access procedure is, however, larger for the 5G candidate systems than for the 4G reference system, assuming an identical deployment. At an intensity of 15000 users/s the 5G candidate systems using uncoordinated and coordinated transmission of the random access response (RAR) are roughly 50% and 200% slower than the 4G reference system respectively. It should be noted though, that use cases with low load but high coverages are important, and in these the 5G candidate systems perform better. This is because it is the first message which usually determines the coverage of the random access process.

It appears that it is the second message in the random access procedure that has a negative impact on the 5G candidate systems. In all investigated candidate 5G systems the number of collisions increase as the system now compares all preambles, as opposed to only comparing the preambles that were received by the same cell.

The candidate 5G system using coordinated transmission of RAR has its performance degraded due to its inefficient way of handling collisions. In the event of a collision the system transmits only a single RAR. If the contending users are located near each other this is the correct decision, but since the 5G candidate systems compares all preambles received in the whole system, which spans a large area, it is usually very inefficient. When users who are far apart transmit the same preamble you want to transmit a RAR to both of these users.

The drop in performance of the candidate 5G system using uncoordinated transmission of RAR is likely caused by the large number of RARs that are transmitted. This candidate system transmits a RAR at all cells that received a specific preamble, instead of just choosing one cell as in the case of coordinated RAR transmission. The RAR is transmitted on the physical downlink shared channel (PDSCH) and therefore additionally requires certain control information to be sent on the physical downlink control channel (PDCCH). The PDCCH is a bottleneck in this case as its capacity is too small due to it not being designed to be able to handle this scenario.

A way to improve the performance of the 5G candidate systems is to increase the number of preambles available. Since the 5G candidate systems originally used one set of 64 preambles, as opposed to the 4G reference system that used different sets of 64 preambles for adjacent cells, this is not an unfair countermeasure.

Increasing the number of preambles for the 5G system to 256 from 64 reduces the delay at the intensity of 15000 users/s by roughly 30% for the candidate system using uncoordinated transmission of RAR and 50% for the candidate system using coordinated transmission of RAR.

8.2 Future Research

Given more time there are certain areas that should have been investigated more thoroughly. Some parts of the procedure were not specifically targeted, or were mostly subject to investigation in the MATLAB calculations. In order to verify the results it would have been useful to investigate these further.

The largest difference between the test and reference systems lies in the two first messages of the random access procedure, and thus it would have been interesting to measure more relevant metrics regarding this in the simulations of the complete procedure. Specifically how frequently users transmitted the same preamble, and how often this collision can be ignored due to there being a large distance between the contending users. As the users are spread out over a larger area they will not necessarily interfere with each other when transmitting the same preamble, but since the system is unable to detect these false alarms all collisions are expensive regardless. The main difficulty with this is to actually find a metric that accurately describes this.

Finding ways other than those used in the coordinated and uncoordinated RAR transmission to choose how many and from where to send RARs would be useful. One possible way of doing this would be to make use of spatial properties of the received preamble, such as angle of arrival. If, for instance, two different cells received the same preamble but noticed that the origin of these preambles could not be the same user due to an incongruity in the angles of arrival, the system would know that at least two users transmitted the same preamble. The system is then aware that it should therefore transmit at least two RARs.

In order to create more relevant results more effort could have been put into creating relevant scenarios. The scenarios with increasingly large intensities do not really reflect any common real scenario, instead it might have been interesting to for instance simulate bursts of users suddenly trying to connect to the system. This would simulate situations such as when a train enters an area, upon which all passengers will try to connect to the network at the same time.

Bibliography

[1] ICT EARTH. URL https://www.ict-earth.eu/. Accessed: 2015-05-25. Cited on page 57.

[2] LTE wikipedia. URL https://en.wikipedia.org/wiki/LTE_(telecommunication). Cited on page 15.

[3] 3GPP TR 25.943 V9.0.0. URL http://www.etsi.org/deliver/etsi_TR/125900_125999/125943/09.00.00_60/tr_125943v090000p.pdf. Cited on pages 35 and 36.

[4] Most promising tracks of green network technologies. *EARTH deliverable D3.1*, 2010. URL https://bscw.ictearth.eu/pub/bscw.cgi/d31509/EARTH_WP3_D3.1.pdf. Cited on pages 3, 4, and 57.

[5] Green network technologies. *EARTH deliverable D3.2*, 2011. URL https://bscw.ictearth.eu/pub/bscw.cgi/d70460/EARTH_WP3_D3.2.pdf. Cited on pages 3, 4, 26, and 57.

[6] Final report on green network technologies. *EARTH deliverable D3.3*, 2012. URL https://bscw.ictearth.eu/pub/bscw.cgi/d70472/EARTH_WP3_D3.3.pdf. Cited on pages 3, 4, and 57.

[7] David Astély, Erik Dahlman, Anders Furuskär, Ylva Jading, Magnus Lindström, and Stefan Parkvall. Lte: The evolution of mobile broadband. *IEEE Communications Magazine*, 2009. Cited on page 2.

[8] Federico Boccardi, Robert W. Heath Jr., Angel Lozano, Thomas L. Marzetta, and Petar Popovski. Five Disruptive Technology Directions for 5G. *IEEE Communications Magazine*, 2014. Cited on pages 1 and 57.

[9] Erik Dahlman, Stefan Parkvall, and Johan Sköld. *4G LTE/LTE-Advanced for Mobile Broadband*. Elsevier Ltd., 2011. Cited on pages 13 and 22.

[10] Pål Frenger and Mårten Ericson. Assessment of Alternatives for Reducing Energy Consumption in Multi-RAT Scenarios. IEEE VTC-Spring-2014. Cited on pages 27, 28, and 29.

[11] Stefania Sesia, Issam Toufik, and Matthew Baker. *The UMTS Long Term Evolution*. John Wiley & Sons Ltd., second edition, 2011. Cited on page 21.

[12] Bernard Sklar. Rayleigh fading channels in mobile digital communication systems part 1: Characterization. *IEEE Communications Magazine*, 1997. Cited on pages 12 and 57.

[13] Cheng-Xiang Wang, Fourat Haider, Xiqi Gao, Xiao-Hu You, Yang Yang, Dongfeng Yuan, Hadi M. Aggoune, Harald Haas, Simon Fletcher, and Erol Hepsaydir. 5G Wireless Communication Systems: Prospects and Challenges. *IEEE Communications Magazine*, pages 122–130, 2014. Cited on page 1.

[14] Dan Warren and Calum Dewar. Understanding 5G: Perspectives on future technological advancements in mobile. 2014. URL `https://gsmaintelligence.com/research/?file=141208-5g.pdf&download`. Cited on page 1.