2008:037 CIV

MASTER'S THESIS

Public Safety Communication over 3GPP LTE

Stefan Sundkvist

Luleå University of Technology

MSc Programmes in Engineering
Computer Science and Engineering
Department of Computer Science and Electrical Engineering
Division of Signal Processing

2008:037 CIV - ISSN: 1402-1617 - ISRN: LTU-EX--08/037--SE

Public safety communication over 3GPP LTE

Stefan Sundkvist

February 7, 2008

Abstract

The 3rd Generation Partnership Project (3GPP) is ready to present a new cellular network standard; Long Term Evolution (LTE). LTE improves communication for public safety users like police, medical, and rescue workers in cellular networks. This creates a need to examine whether LTE can satisfy public safety users requirements.

This thesis has a focus on group communication with a voice service. Using the LTE cellular network simulator we find the cell capacity and examine the performance at the capacity limit. We show that LTE can serve large groups for communication and by examining different voice activities we show there is only a small effect on capacity.

With a study of optimization techniques we show that combining a low codec bit-rate with two bundled speech frames per IP packet gives a significant increase in capacity. By improving the scheduler algorithm by increasing the priority to old packets in the sending queue, we increase the capacity further. Based on the results from the evaluation we determine that LTE can satisfy the public safety users requirements, making LTE an option for public safety communication.

Acknowledgements

I would like to give a special thanks to Mats Folke and Peter de Bruin at Ericsson Research for their insight and expertise during this thesis. Not to forget anyone I would also like to give a big thank you to everyone at Ericsson Research for their support that have been crucial for the realization of this thesis. My examiner Magnus Lundberg Nordenvaad should have my last thanks for giving me the opportunity to do this project at Luleå University of Technology.

Contents

Terminology

Abbreviation[1]	**Explanation**
3GPP	3rd Generation Partnership Project
AMR	Adaptive Multi Rate codec
CDF	Cumulative Distribution Function
CQI	Channel Quality Indicator
DTX	Discontinuous Transmission
DVB	Digital Video Broadcasting
eNodeB	LTE base station
FTP	File Transfer Protocol
HARQ	Hybrid Automatic Repeat reQuest
HSDPA	High Speed Downlink Packet Access
IMS	IP Multimedia Subsystem
IP	Internet Protocol
LTE	UTRAN Long Term Evolution
MOS	Mean Opinion Score
OFDM	Orthogonal Frequency Division Multiplexing
OMA	Open Mobile Alliance
PAPR	Peak to Average Power Ratio
PoC	Push To Talk over Cellular
PTT	Push To Talk
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RAN	Radio Access Network
RB	Resource Block
RLC	Radio Link Control protocol
ROHC	Robust Header Compression
RTP	Real Time Protocol

[1] A complete list of 3GPP abbreviations is available in [7].

SC-FDMA	Single Carrier Frequency Division Multiple Access
SID	Silence Insertion Descriptor
TETRA	Terrestrial Trunked Radio
TTI	Transmission Time Interval
UMTS	Universal Mobile Telecommunications System
UTRA	UMTS Terrestrial Radio Access
UTRAN	UMTS Terrestrial Radio Access Network
VAF	Voice Activity Factor
VoIP	Voice over IP
WiMAX	Worldwide interoperability for Microwave Access

Chapter 1

Introduction

Public safety users like police, firemen, and rescue workers needs to have a good and reliable way to communicate with their command central and other public safety users in the field, with special requirements for high network coverage, high system capacity, and low latency. Today most of these users have a dedicated public safety networks to communicate, for example, digital systems like RAKEL [18], P25 [19, 20], or some old analog system that use radio-broadcast techniques for basic group communication. Compared to commercial cellular networks (e.g., GSM or UMTS), public safety users have a high cost per subscriber in a dedicated network because of a limited number of subscribers bearing the cost of the network.

An evolution in commercial cellular networks has resulted in high spectrum efficiency, reduced latency, and improved prioritization of users. Cellular networks have in the last couple of years become an option for public safety users to reduce the cost per subscriber. Advanced services like video and high-speed data transfers are made available to public safety users taking advantage of the higher bit-rates.

In 2004, 3GPP launched a Study Item titled *Evolved UTRA and UTRAN* [4] and resulted in the start for a next generation cellular network with the *Long Term Evolution* (LTE). Providing improved coverage, higher system capacity, and significantly reduced latency [10]. The high capacity and low latency should enable 3GPP LTE to meet the requirements of the public safety organizations.

To enable a voice service in broadcast networks using half-duplex communication, with a single user speaking at the same time, the user finds a button on their user equipment to request the floor, switching from being a passive listener to an active speaker. We often call this technique *Push to Talk* (PTT) when talking about broadcast or multicast radio networks.

A cellular network (e.g. GSM, UMTS, or 3GPP LTE), has the PTT feature artificially by using unicast communication techniques, and almost instantaneously sets up a connection to all users in a predefined group with a press of a button. *Open Mobile Alliance* (OMA) defined the PTT feature for cellular networks in the *Push to talk over Cellular* (PoC) [21] standard.

The use of unicast techniques for group communication in the 3GPP LTE network instead of multicast or broadcast techniques may enable group communication for public safety users. Enable public safety users to take advantage of a low cost per subscriber and high bit-rates.

1.1 Objectives

The objective of this thesis is to evaluate the use of unicast techniques for group communication in 3GPP LTE, with public safety users in the same network as commercial users would be in. The scenario we use is built upon a large sporting event or accident where public safety users spread out in a confined area; we will evaluate network and speech performance at the capacity limit. There is also an evaluation of a few techniques to improve the capacity in the network.

1.2 Delimitations

This study concentrates on user capacity and speech performance in a 3GPP LTE network regarding group communication. Related issues with call setup are beyond the scope of this thesis, as are other multimedia services suited for group communication. The simulations only considers the downlink, as there is only a single user speaking simultaneously in the uplink resulting in the downlink being the limiting factor.

Chapter 2

Technical background

2.1 3GPP Long Term Evolution

The standardization work for 3GPP Long-term Evolution (LTE) started in November 2004 at the 3GPP RAN Evolution Workshop in Toronto, Canada [1]. Some of the requirements decided were:

- Reduced cost per bit
- More services at lower cost with better user experience
- Flexibility of use of existing and new frequency bands
- Simplified architecture and seamless mobility with legacy systems

Later in December 2004 at the 3GPP RAN #26 meeting 3GPP approved the description of the study item on Evolved UTRA and Evolved UTRAN [2]. The study item for the requirements evolved UTRA and evolved UTRAN [3] was put forward to develop a framework for an evolution of 3GPP radio-access technology to a high-data-rate, low-latency, and a packet optimized radio access technology. Some of the targets to clarify were:

- Increased peak data rate e.g. 100 Mbps (downlink) and 50 Mbps (uplink)
- Improved spectrum efficiency
- Possibility for a radio-access network latency below 10 ms
- Reduced control-plane latency (e.g. call setup) to less than 100 ms
- Scalable bandwidth of 20 MHz, 15 MHz, 10 MHz, 5 MHz and also less than 5 MHz

In May and June of 2007 it was recognized that LTE, as currently specified meets all and exceeds many of the LTE performance targets specified in 3GPP TR 25.912 [4].

LTE uses Orthogonal Frequency-Division Multiplexing (OFDM), for the downlink (i.e. from the base station to the terminal), OFDM is a well-established technology. And is also used with other wireless technologies like WiMAX [14] and broadcast technologies like DVB [13].

OFDM transmission uses a large amount of sub-carriers. A straightforward multi-carrier extension to a narrowband transmission scheme would lead to only a few bands. In a 20 MHz spectrum, we would need guard-bands to prevent the different sub-carriers to interfere with each other. The bands with orthogonal sub-carriers, can be much closer to each other, resulting in a larger number of sub-carriers.

In the frequency domain in LTE the OFDM sub-carrier spacing, Δf, is 15 kHz. The time domain divides with Δf giving the OFDM symbol duration time $1/\Delta f$ + cyclic prefix. The cyclic prefix will prevent interference between sub-carriers even in a time-dispersive channel. This creates a two-dimensional grid with frequency on one axis and time on the other as seen in figure 2.1. Each element in the grid is a resource element. One such element contains one OFDM symbol that can be coded with quadrature phase shift keying (QPSK) or quadrature amplitude modulation with 16 or 64 levels (i.e. 16QAM or 64QAM), giving the highest modulation 64QAM, 6 bits per resource element.

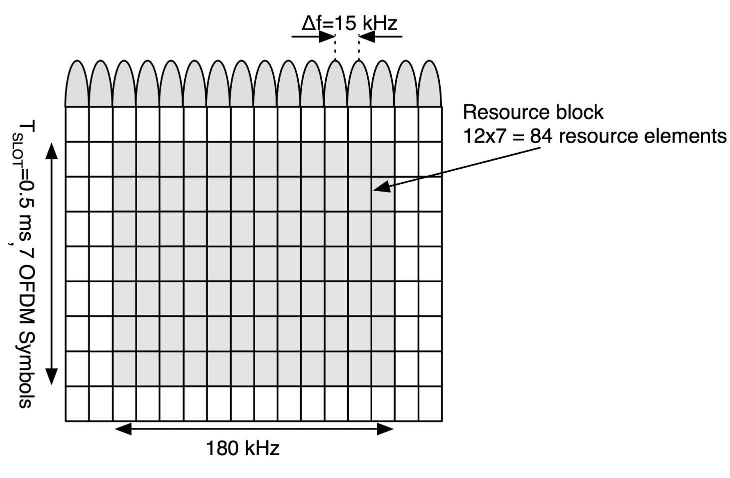

Figure 2.1: LTE physical downlink channel based on OFDM

The resource elements are grouped into resource blocks consisting of 12 subcarriers with a total of 180 kHz in the frequency domain and one T_{slot} = 0.5 ms in time domain that contains 7 OFDM symbols. With the highest modulation, 64QAM, a resource element contains 6 bits. Given a resource block with a total of 84 resource elements, it will be able to carry up to 504 bits with 64QAM. Two resource blocks are then grouped together in the time domain creating a 1 ms Transmission Time Interval (TTI).

For each resource block in LTE, a scheduling mechanism decides in both time and frequency domain how many RBs each terminal will get. A more in-depth description on LTE Scheduling can be found in chapter 2.2.

For the uplink (i.e. from the terminal to the base station) LTE uses a Single Carrier Frequency Division Multiple Access (SC-FDMA) transmission technology. The decision to use SC-FDMA instead of OFDM in the uplink is caused by a drawback in OFDM with a high *Peak to Average Power Ratio* (PAPR) that makes the power amplifiers inefficient. This causes the battery in the terminal to drain faster. With SC-FDMA this is resolved by grouping resource blocks in a more efficient way. This increases the coverage and battery time in the terminal.

A deeper introduction in LTE can be found in the book "3G Evolution: HSPA and LTE for Mobile Broadband" [11].

2.2 Scheduling of resources in LTE

A modern scheduler should differentiate between the Quality of Service (QoS) classes available, optimize each service to its specific needs and maximize the use of the radio resources. It should also be easy to the operators to include new services into the network.

The scheduler used in 3GPP LTE makes a decision about what resources should be scheduled to each subscriber in each cell. The decision is done at every TTI (1 ms) by choosing which modulation (e.g. 64QAM) and coding scheme (e.g. turbo codes) to use in each resource block. If multiple antennas are used the scheduler must also select which antenna configuration to use.

In the current UMTS QoS architecture, there are four types of classes, conversational, streaming, interactive, and a background traffic class. With these four basic classes, the network can differentiate between real-time conversational VoIP sessions that need time priority and an FTP background session where highest possible bit-rate is more important. Giving these services has a very different set of requirements. There should also have functionality to separate different groups of users with different priority. As a public safety user can have a higher priority then a civilian citizen without any emergency call waiting to get trough.

The scheduler exploits channel variations. This is done by transmitting data to terminals with advantageous channel conditions. This is the same idea used in other schedulers (e.g. HSDPA) as well. As explained earlier, LTE uses OFDM which enables both time and frequency domain scheduling. This gives the opportunity to use both the time and frequency domain when deciding what terminal should receive resources. This can be compared to High Speed Download Packet Access (HSDPA) which only can use the time domain variations. This is especially advantageous to LTE where the use of frequency domain scheduling gives traffic with low bit-rate but with low delay constraints (e.g. VoIP session) a higher probability to be sent. Whereas the time domain changes slowly, exploiting the frequency domain gives a higher probability to send the data within the time constraints set to the service.

To make a good choice for modulation and coding scheme the scheduler needs to know the channel conditions. In the downlink, a reference signal is sent to the terminals that estimate the performance of the channel and sends a channel quality report back to the eNodeB as a Channel Quality Indicator (CQI). The CQI gives the eNodeB information of the current channel conditions in the terminals, and information about spatial multiplexing if using multiple antenna techniques where the scheduler takes advantage of spatial diversity as well.

2.3 Public safety communication

Traditionally public safety organizations use radio broadcast systems to communicate with each other. These systems often only have voice communication, and can only be used in a limited range from the dispatcher. In a larger fire, where firemen from different departments must cooperate there might be problems where their communication equipment is not compatible with the equipment from the other department. This problem was identified in several studies [16, 22] in the USA, where different departments has been found delivering messages in the same way the old Greeks would, on foot. Sending public safety users to deliver the message, because their systems would not talk to each other.

There are several projects that want to unite the communication between different organizations like fire departments, rescue workers, police, and other first responder units. Among them, there are some digital systems like terrestrial trunked radio (TETRA) [25] and the Project 25 (P25) [19]. Both systems where constructed to solve the issues with old analogue systems. One of the problems that are unsolved is the high cost for subscribers in the network, which is caused by a small number of users in the system to bear the cost of the network infrastructure. Earlier studies on push to talk over cellular and professional mobile radio [8] conclude that requirements of low delays and availability are not met in UMTS networks. However, most of the issues in the report have been sorted out in LTE, making it a candidate for a public safety communication network.

There are several special requirements for public safety users, here are some that has been identified in the PoC and PMR study [8]:

- The response times requirement for call setup time is between 0.3 to 1 second. The most cited number for average usage is 0.5 s.
- Full radio coverage for the served area, even in exceptional conditions.
- Radio capacity should accommodate all first responders on the site, even in larger a larger accident or sporting event where several departments might want to communicate.
- Good voice quality enabling the listener to identify the speaker, even with large amounts of background noise.

These requirements should be considered as a baseline case. There are also different requirements depending on the services.

Services today mostly consist of voice services. The larger bandwidth in LTE gives new abilities for high bit-rate services like video and transfers of data (e.g. images, sensor data) from the emergency site to the control center.

2.4 One-to-many communication in LTE

Earlier we have seen that regular public safety user's use broadcasted radio transmissions for group communication between public safety users and communicate between terminals without the use of a network. However, in LTE and other cellular technology this communication needs to go through the network instead of using broadcast radio transmissions between terminals.

This voice service has often been marketed as a Push to Talk service, where the user push a button to request to the floor to speak. In cellular networks is it known as, Push to Talk over Cellular (PoC) and standardized by the Open Mobile Alliance (OMA). The commercial application of PoC let the subscribers create small talk groups, where they can make group calls by pressing a button on their PoC enabled phone.

From a technical point of view, PoC is related to other dial-up services like Voice over IP (VoIP) that is an one-to-one service, except that a PoC service has a call already setup and does not need to dial-up each user in the group. Both of them use a Session Initiation Protocol (SIP) and Real Time Protocol (RTP) for session initiation and to relay voice frames. In the VoIP case it is only one-to-one communication, however, it is not hard to extend this scenario to an one-to-many communication, where the speech frames are duplicated and sent to each of the recipients.

However, the PoC standard maintained by OMA is to a degree implementation specific. In LTE it is possible to implement a successful PoC service through the IP Multimedia Subsystem (IMS) from 3GPP. However for public safety users who have other demands on the service than regular users there might be a need for a modified or optimized IMS solution.

2.5 User behavior

For an ordinary one-to-one session, a typical voice communication is above 50% speech activity. To describe differences between an one-to-one session and an one-to-many session with public safety users we need to describe some common concepts.

A session in an one-to-one scenario is the same as the duration of a phone call, where each active speech period is called a talk-spurt. In old circuit switched systems a session uses the same bit-rate during the whole session. However, with packet switched communication the channel is divided into small blocks of data and using discontinuous transmission (DTX) the packets is only transmitted during a talk-spurt, and so reducing the total amount of traffic.

In the duration between to talk-spurts, there are two parts, one containing speech and the other that is quiet. The ratio between these gives the Voice Activity Factor (VAF), and measures how many people talk in a session. In the one-to-one scenario a voice activity of 0% means that there is no talk at all, and a voice activity of 100% is when there is talk always in both directions. A 50% voice activity is therefore, a session where both users talk but not at the same time. We can define VAF in an equation where we have the talk time

t_{speech}, and the quiet time t_{quiet}, and VAF is defined as

$$\mathrm{VAF} = \frac{t_{speech}}{t_{speech} + t_{quiet}}. \tag{2.1}$$

For one-to-one communication, the users are dispersed in time reducing the total amount of traffic. However, in one-to-many communication with public safety users the talk-spurts become synchronized as the group gets the speech at the same time resulting in higher peak traffic for the whole channel with the same amount of users. Also, the traffic during a quiet period is reduced to nothing, as no one transmits any traffic in quiet periods with DTX.

Chapter 3

Cellular network simulator

The event-driven simulator models an end-to-end cellular, which use an event object to describe a specific event. These objects are then inserted in an event queue sorted in time, then executed in sequential order, either directly or later at a specified time.

For the simulator, all parts in the simulator are made as a module that easily can be exchanged. Each module represents one part of the communication chain 3.1, for example, traffic models, channel coding, communication links, user behavior and many other types of events exist in a network. A module consists of several variables to configure a simulation, for a specific scenario. Each scenario is specified in a configuration file, which contains parameters for each of the modules used, and which modules to use. These may be timer parameters, delays for encoders, or decoders, antenna fading parameters, user behaviors, and more. To help with the complexity of an end-to-end cellular network simulator, most of these parameters have been given default values. This brings the complexity when configuring a simulation down to a more manageable level. When only needing to change parameter settings that differ from the default situation, the configuration is kept to a minimum.

Another useful part you need when simulating an end-to-end cellular network is variables to collect data during the simulation. For example, you might want to know how many packets have been lost or the delay of each packet. For this purpose, the simulator has a separate framework to configure variables that collects data from within the simulation. However, in a simulation there is a large amount of data to collect and often there is only a need for statistics about the data. One common thing to collect is histograms of e.g. delays or the number of packets that have been sent. These data are later collected for further post-processing, producing graphs and other statistics.

3.1 Simulation design

In this section, there is an outline of the simulator setup used. The cellular simulator used is as mentioned built with modules which describe each part of the communication chain. In figure 3.2 you see a simplified view of the simulator.

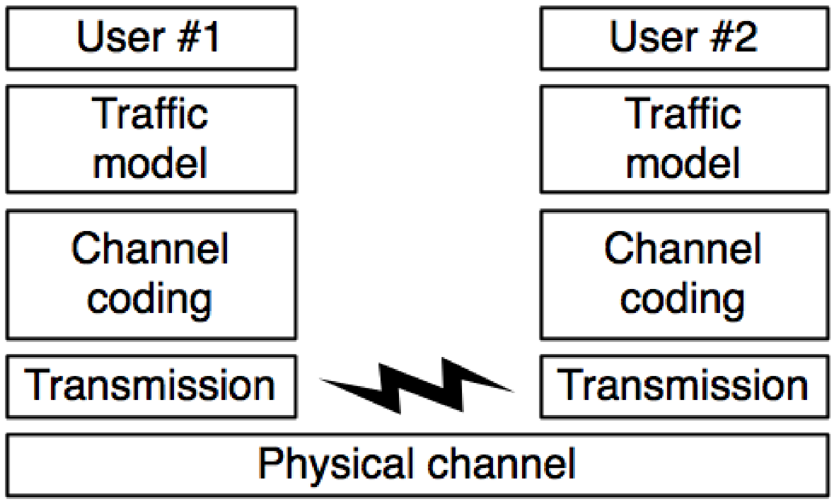

Figure 3.1: The chain of communication between users

In figure 3.2, the simulation consists of a core simulation class and will handle the event queue making sure every event gets executed in the correct order. It is also responsible for reading the configuration and setting up what modules to use in the simulation. These classes are later added as children to our parent simulation class, describing each module in the communication chain.

The simulations in this thesis are dealing with the LTE network, and therefore have a common LTE network class, used for all LTE simulations. The LTE simulator class sets default values and creates a default simulator. This default simulator is all that is needed to run a simulation. Some child modules that will be used by all instances of the simulator class are the radio module, eNodeB site module, and a traffic module. The traffic module is of the most importance here, as this is where we define how the users should behave in the cellular network. The eNodeB site module defines the type of site that should be used in the simulation. For instance, we will use a regular hexagon type cell, but there are other patterns that may be used, which may be suitable for other types of scenarios. The radio module controls the bandwidth and radio spectrum to be used.

The users are created in a user generator class. The users are considered as users with one-to-one sessions, as they have many parameters in common. The difference is in the behavior of their transmission. An one-to-one user sets up a call to the recipient user and uses the connection for a specified time until it closes the connection and finishes the call (i.e. hangs up). However in one public safety scenario we want one-to-many communication. This means that the call setup is made differently. Exactly how this is done is out of the scope of this thesis and thus not a part of this report. One part that is included is the fact that in one-to-many communication, as the name implies; only one person is speaking while the other is passive listeners. Resulting in small amounts

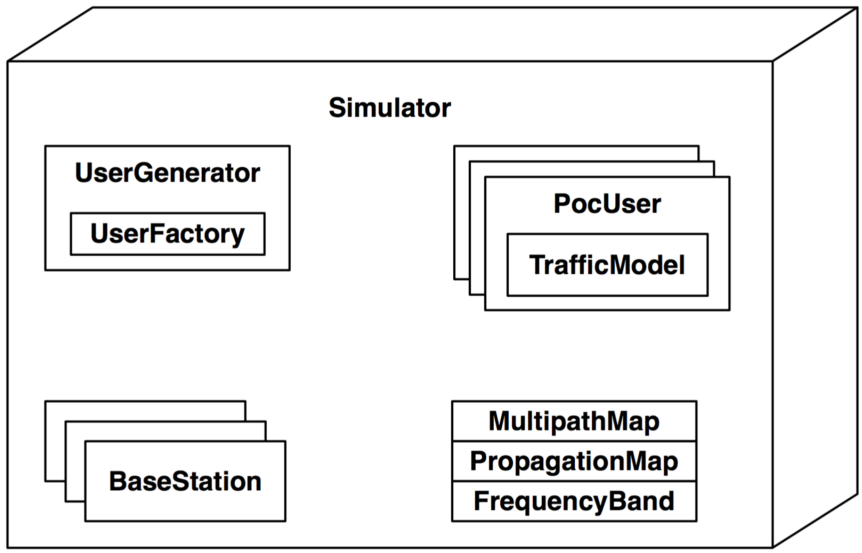

Figure 3.2: A simulator with classes at a specific time

of traffic in the uplink compared to the downlink, making the downlink the limiting factor. Not only do they receive everything in the downlink, the traffic is also synchronous (i.e. sent at the same time instance). This creates some traffic patterns that needs further study, and is the main focus of this thesis.

Each user in the simulation is placed within our network and consisted of 3 sites, containing 3 cells each, and can be seen in figure 3.3. The users are placed at random positions if not stated otherwise. On their initial spot, they use a random walk simulation to illustrate movement and give a more lifelike travel pattern. The walking speed is 3 km/h. In these simulations, the users walk in a straight line and wrap around borders. So when the user walks past the border, it appears on the opposite side of the cell.

The scheduler class makes a model of the scheduler that should be used in the simulation. As described in section 2.2, it schedules the traffic events using an additive scheduler, setting a higher weight to packets that need a higher priority. For example, one can set an extra high priority to packets that have been in the sending queue for a specified amount of time.

Each user is also able to set specific network settings for their connection. One of these modules that are in all these simulations is Robust Header Compression (ROHC) [9]. This module compresses the headers in each packet to about 3 bytes, reducing the total amount of traffic.

3.2 LTE Simulation settings

As discussed before 3GPP LTE uses OFDM that is configured with these downlink parameters, uplink is beyond the scope of this thesis.

These are only some of the vast number of settings available with LTE, which have a very flexible configuration. However, these are some of the standard

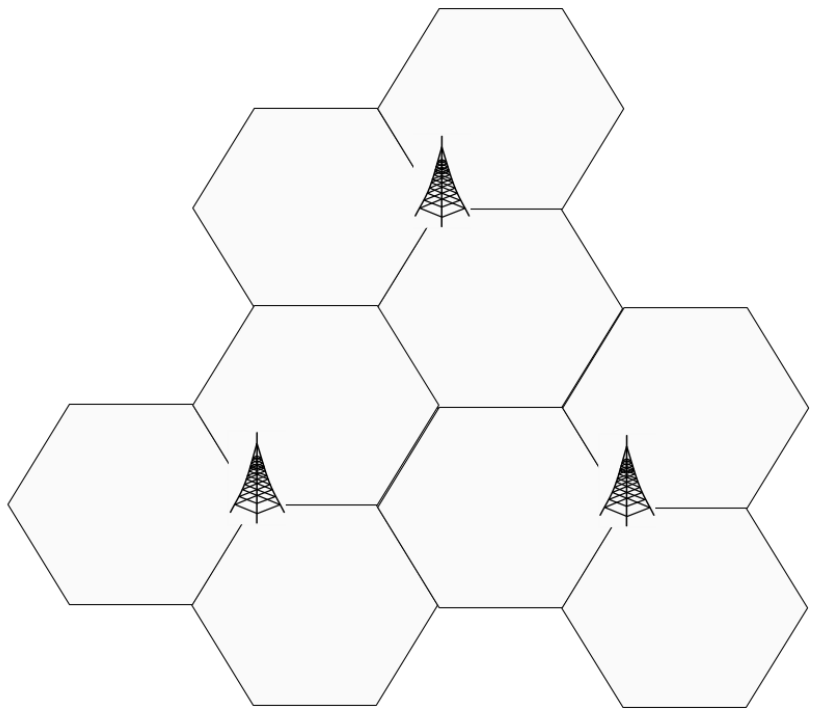

Figure 3.3: Cell network used in simulations

Parameter name	**Parameter value**
Subbands	25
Symbols per subband	144
Transmission Time Interval	1 ms
Bandwidth	5 MHz
Carrier frequency	2 GHz

Table 3.1: General simulation settings

setting believed to be used in the upcoming 3GPP LTE standardization that is on its way. Some of the settings like subbands are dependent on other settings. For example, the number of subbands used is related to the bandwidth. As each subband is given a spectrum of 180 kHz, and we have a 10 % guardband giving 4.5 MHz of useful bandwidth. This is also what we get when multiplying the subband spectrum by 25. Other bandwidths may give other numbers of subbands. For the symbols per subbands, they are here given a smaller number then in reality to model signaling overhead.

3.3 Scheduler overview

The scheduler in LTE is an additive scheduler in the time domain, meaning that different schedulers may be added freely of each other by adding weights for each specific service, enabling the operator a flexible environment to make network optimizations. There are two kinds of weights, one for the frequency domain that is completely dependent on the CQI for the channel, choosing the best resources possible for each user. The other weight is in the time domain. These weights might measure a vast number of parameters within the network, however considering the scope of this thesis the focus is on these weights.

- Fixed weight
- Delay weight
- Age weight

The *fixed weight* is a basic weight that can be used with all types of traffic to give a bonus weight. This is used here because we need to give our public safety users a higher priority then other users. This is achieved by giving a higher fixed weight to public safety users, higher than other users have with all their scheduling weights combined.

Delay weight uses a function to specify the increase in weight. This function may however be any type of function; for example, nonlinear, and exponential functions. The x-axis in defined with delay weight as the time in milliseconds from the last time a resource unit was delivered from that queue.

Age weight is related to delay weight. It uses the same concept with functions to map a measurement onto an axis of scheduling weight. The weight itself is also similar as delay weight. However instead of looking at the delay of the last scheduled resource unit, the age weight looks on the age of the last packet in the queue. This will later be used to help older packets by giving them a higher priority.

Chapter 4

Simulation scenario

4.1 Public-safety users

In the event of a large accident or a sporting event where large amount of public safety users might be located in a small area. This gives a starting point for the simulations. All simulations are based on this scenario with a large concentration of public safety users, distributed over the area. Each user is equipped with a terminal that will receive group calls from the groups he happens to listen to, most commonly a dispatcher or a co-worker in the field.

A public safety user will start the simulation with a pre-initialized session, i.e., the call setup has already been taken care of before the simulation. The user is considered to have an access session open during the simulation, as we consider only the connection level of the simulation. On the connection level, the talk-spurts are synchronized between the users acting as passive listeners, as they will receive them simultaneously. This will simulate an one-to-many communication scenario used in all simulations.

4.2 Traffic model description

To develop a model that is used in the coming simulations there is a need for some clear definitions on the talk activity in our network. This especially for group communications where the definition differs on some points compared to regular voice calls. In [23] five levels of speech activity are defined. The highest level is the access session level. At this level, we have all users that are online in the network. For our public safety usage this level is always on, as most users will have their radio communication equipment on at all times. Below that level is the application level. This level is the time between a public safety user pushes the push to talk button until it's released. In our simulations, we simulate one of these sessions for different group sizes. The third level is the connection level, here we can identify talk-spurts that are simulated. The levels below are the packet and burst levels. These levels are not the core focus of this thesis; however, they are considered as they might affect some of the results. A graphical description of these levels can be found in figure 4.1.

In the connection level, we have many talk-spurts. Before we state the

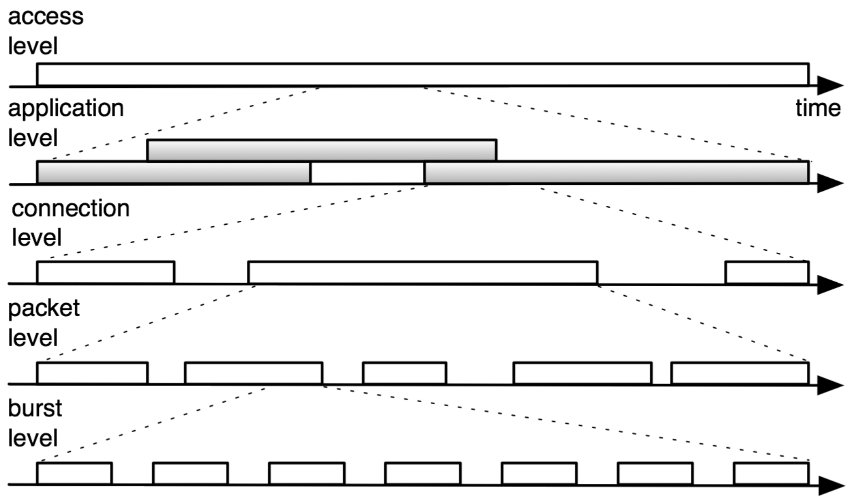

Figure 4.1: Different activity levels

timings of these, the definition of the talk-spurt needs to be well defined. The arrival times of a talk-spurt can be seen as a non-negative stochastic sequence $\{A_n\}_{n=1}^{\infty}$, defined as $A_n = T_{n+1} - T_n$. This definition is the same as stated in [15,23] and a graphical representation can be found in figure 4.2. Other than arrival time there is also a duration time $\{D_n\}$ for each talk-spurt. Because only one speaker is allowed simultaneously within a group A_n needs to be greater then D_n.

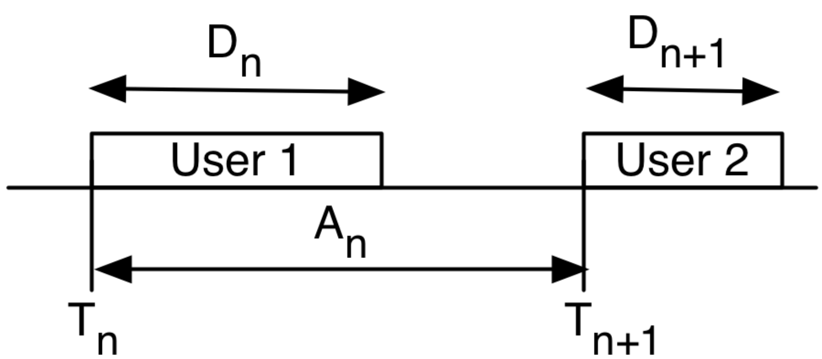

Figure 4.2: The talk model used in simulations

4.3 Talk-spurt timing in real world scenario

This section discusses talk-spurt duration, and arrival time that has been seen in studies of public safety users like [23], carried out in Aachen in the TETRA network. These show a speech activity in the network that can be expected in other networks with public safety users like our cellular network studying 3GPP LTE. The measurements needed to calculate the Voice Activity Factor (VAF) is the mean values of the arrival time and duration time so that:

$$\text{VAF} = \frac{E[D]}{E[A]} \tag{4.1}$$

In the TETRA trial study [23] they state that the arrival timing $E[A] = 11.6$ s and $E[D] = 5$ s. This would mean a voice activity factor of 43%. However, these numbers have a high variance that is a result of the large variations occurring during the day. On busy hours, the voice activity factor might get closer to 90% and other off hours as low as 10%.

The simulation reflects the findings in the TETRA [23] study. However, in a simulation you need to consider factors like simulation time and the computational resources available. This can be done by trying to limit the duration of the talk-spurts (i.e. D). In the simulations three duration times where used, set at voice activity factor 10%, 50% and 90%. The arrival time is set to 3.4 s in all cases. This was a compromise between smaller talk-spurt timings and while being large enough to be statistically sound.

Chapter 5

Simulations

5.1 Speech Quality

When simulating voice services, speech quality is one of the hardest things to measure and still one of the most important. Speech quality is commonly measured with listening tests and is therefore, subjective. However, with only objective measurements available, listening tests is not an option. The indirect approach to measure speech quality is therefore to measure the parameters that are known to decrease the speech quality, like packet loss. Several studies have linked packet loss to speech quality, e.g. [12,17], giving a tool to measure speech quality objectively.

In an one-to-one scenario, it is common to have a requirement for voice communication and subsequently good speech quality if it has less than 2% packet loss, 1% lost packets in each direction. However, in the one-to-many scenario considered in this thesis there is only traffic in the one direction, so consequently, the requirement is set to 1% packet loss [5,6].

The capacity limit for the cell is related to packet loss, and has been decided to be when 95% of the users has less than 1% packet loss. In the simulation this is archived by increasing the load in the cell until it reaches capacity limit, where we measure jitter and resource usage. Also for public safety users we are interested in the intra group latency defined in section 5.1.1.

5.1.1 Intra group latency

In this thesis, group level performance is measured in intra group latency, defined as the time for the speech to propagate to all users within the group. This because, bursty traffic as one-to-many communication is, causes a large load on the network and can't transmit to all users immediately. A large latency could be perceived by users as a decrease in speech quality. For example, a large delay between two listening users close together can degrade the service quality if they can differentiate between the users incoming speech.

The evaluation of intra group latency is performed at the capacity limit and presented as a Cumulative Distribution Function (CDF) and the measured when 95% of the users has started to receive a new talk-spurt.

5.1.2 Jitter

For the individual users within the group, with real-time requirements, the speech quality is degraded if the delay between each packet is too large for the algorithms in the user equipment, resulting in a lost speech frame. Measuring this delay helps the manufacturer to build the user equipment, avoiding speech quality degrading because of jitter.

In the simulation environment this delay is measured as jitter, or the variance in the network delay between each packet. For the simulations, this time is defined as the time that 95% of the user's stays within during the talk-spurts.

5.1.3 Resource allocation

To be able to maintain a good network the operator of the network need to know the network resource used to be able to plan and design the network. This is done by measuring the network resources used in each cell. As stated earlier a network has several resource blocks which are transmitted at each TTI. Measuring the number of users able to transmit at each TTI would then measure how much of the resources are being used at each TTI.

The simulations define this as the number of users being able to transmit per TTI. This gives operators an idea how many resources the service requires for a specific load, here it is measured at capacity limit.

5.2 Simulation setup

In section 5.1 there is a detailed description of the components that the speech quality depends on. However, in a simulation environment, there are also many network aspects to consider when measuring the speech quality and network performance. So for a good measurement the performance of each parameter should be identified and analyzed in a simulation.

The simulations are divided in two sets, where the first set look at the performance of the traffic scenario without any optimization of the traffic. Looking at speech quality and network performance at the capacity limit, giving an idea of the network performance, when the cell load is low that is more probable in a realistic scenario. Also, the relative comparison to a conventional traffic scenario with point to point communication, analogues with a regular telephone call, shows how much capacity is lost when performing point to multi-point communication.

The second set of simulations looks at the possibilities to increase the capacity limit in the cell by optimizations. Depending on what optimization performed it will decrease the performance on some speech quality parameters, either by lower sound quality or increased delays in the network, while trying to stay within the limits specified for public safety users. To get an idea on what optimization is worth pursuing they will be given a relative comparison, looking at the performance gain and the effect on the whole system.

5.3 Performance without optimization

Using three simulations, to evaluate the capacity limit, each showing different aspects and described below. Important aspects are groups in the network and the evaluation of different speech activity factors. During the day, the voice activity experience large variations, and an increase in speech activity could cause congestion because of an increase in traffic. The effect of different types of speech activity is therefore, important.

The network performance with different number of groups, give some insight in how the performance changes from a best case scenario to the worst case. The worst case would be having only a single group, and dividing the group until the best case resulting in the one-to-one traffic model. Using this as a model results in a relation between the number of groups and the capacity in the cell.

However, first a reference comparison between the capacity with one-to-one traffic and one-to-many traffic, showing in percentage the capacity decrease. Also look at the speech quality and network parameters in the new scenario at capacity limit.

5.3.1 Reference comparison between one-to-one and one-to-many traffic

The setup used for comparison between the one-to-one and one-to-many traffic models tries to give a reference with the rest of the simulations. There are several design decisions made in the construction of the simulation, these to compare between the normal telephone call model using one-to-one traffic and the public safety model where the traffic is one-to-many.

The reference scenario use a 30 seconds session length, where the arrival time between each talk-spurt set to 3.4 seconds on average. The voice activity factor, defining how much speech traffic there is in every arrival time-frame, set to 50% meaning that there are on average 1.7 seconds long talk-spurts and therefore in average 1.7 seconds long periods without any traffic.

For an estimation of the capacity limit, the simulator sets a static network load in a 30 second session. After a session without reaching the capacity limit, we rerun the simulation with a higher load. Other parameters are as specified in chapter 4, and a description of the scheduling is in section 5.4.1.

At capacity limit we evaluate intra group latency, jitter, and resource usage giving an idea of the performance, and how it performs considering the requirements of public safety users.

5.3.2 Voice activity effects on capacity

One aspect of our scenario to simulate is voice activity, and the effect it has on the capacity. In this scenario, the voice activity is in the first simulation raised to 90% and in the second simulation lowered to 10%. This results in lower total traffic in the network. However, this will indicate if the capacity depends on voice activity in the one-to-many public-safety scenario.

The analysis for system measurements made at capacity limit, shows how they effect the performance and if they are within the requirements for public safety usage. The data analyzed is using a voice activity at 50% to compare with the reference simulation. The table 5.1 shows the voice activity settings in use.

Other simulation settings are the same as in the reference simulation above, using 30 second long sessions and increasing the cell capacity until the capacity limit of the cell.

Voice activity factor	10%	50%	90%
Arrival time	3.4 s	3.4 s	3.4 s
Duration time	0.4 s	1.7 s	3.0 s

Table 5.1: Voice activity factor settings

5.3.3 Multiple groups effect on capacity

In the reference scenario, there is one single large group that receives the talk-spurt whichmay be an extreme situation. In day-to-day traffic it is more common with smaller groups communicating.

These simulations use the same setup as the reference simulation, changing only one of the parameters. The scenario divides a large group into several groups, doubling the number of groups each time. Going from two to eight groups, and interpolated between to have a capacity graph from one group up to the best case that is the one-to-one scenario.

5.4 Performance with optimization

Optimization analysis shows what optimization is worth the cost in a cellular system. Many tricks can increase in the cell capacity. However in this thesis we consider only three of them, excluding other types of optimizations.

5.4.1 Improved scheduling

Scheduling in LTE as discussed in chapter 2.2, exploits both the time and frequency domain to take advantage of the fluctuation in the physical channel. In previous simulations, scenarios the scheduling has been done with a round-robin like scheduler. A round-robin scheduler is the most simple of schedulers and is constructed so that a delay weight is added almost immediately as the packets come in to the queue. This gives a completely fair scheduling where every user is given the same amount of time.

With a real-time system like group voice communication, we have time constraints, an old packet will be considered lost if older than a specific time. In a round-robin scheduler all packets have the same amount of time, and no prioritization of older packets that might be lost because of their age. This can easily be fixed by giving old packets a higher priority, by adding an age weight

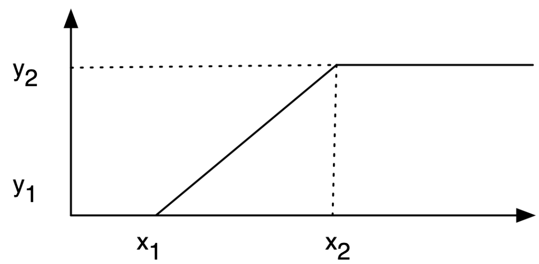

Figure 5.1: Linear function used in weight calculation

to the scheduler that increases the priority of packets that have been in the queue for a time. In figure 5.1 we see a description of the function used for the scheduler settings. In table 5.4.1 we see the settings used in these simulations, in the reference scenario only the Delay weight is used, in the more advanced scheduler also the age weight is used.

Weight type	x_1	x_2	y_1	y_2
Delay weight	0	2	0	200
Age weight	45	65	0	200

Table 5.2: Scheduler settings with age weight added.

5.4.2 Decreasing the codec bit-rate

Earlier the codec bit-rate has been set to 12.2 kbit/s and is the highest setting in the codec used, with a good sound quality. However in an emergency situation, a lower sound quality may be accepted when the message is of most importance.

This simulation looks on the scenario where we want to send message to a large group. This is achieved by decreasing the codec bit-rate from 12.2 kbit/s down to 4.75 kbit/s and is the lowest setting. And measure how much the capacity is increased, resulting from the lower bit-rate in the codec. However, this makes it hard to compare directly to other simulations, as the codec bit-rate is changed and is seen as a reference for the performance with a low bit-rate codec. The simulations will however still use the same definitions for capacity limit.

5.4.3 Bundling of speech frames

In the last simulation scenario, the total bandwidth was lowered by using a lower codec bit-rate, subsequently decreasing the sound quality. Here, the codec bit-rate will stay at 12.2 kbit/s by attacking the headers instead with bundling of speech frames into a single packet. Reducing the total header size by half and, leaving more room for voice data.

The cost of using this method is increased latency. Because only every other packet is sent, increasing the time between each transmission, from 20 ms to 40 ms.

5.4.4 Combining decreased bit-rate with bundling

In sections 5.4.2 and 5.4.3 the bit-rate is reduced by decreasing the headers or speech data. However, this will create two problems, first if decreasing the speech data, the proportional size on the headers will be larger, and in the other case because we are using ROHC that reduces the size of the headers to only 3 bytes instead of 40 bytes relative to the size of the speech data, it is only reducing each packet by 12 bits.

The last optimization is therefore, a combination of reduced codec bit-rate and bundling of speech frames for reduced header size. The specific bit allocations in all the different scenarios can be seen in table 5.3.

	12.2 kbit/s bit-rate		4.75 kbit/s bit-rate	
AMR frame per IP packet	20 ms	40 ms	20 ms	40 ms
Speech data (AMR)	244 bit	$2 * 244$ bit	95 bit	$2 * 95$ bit
ToC + CMR [24]	10 bit	$10 + 6$ bit	10 bit	$10 + 6$ bit
Octet aligned padding	4 bit	0 bit	7 bit	2 bit
Total AMR data	264 bit	504 bit	112 bit	208 bit
IPv4/UDP/RTP (ROHC)	24 bit	24 bit	24 bit	24 bit
Link level headers	80 bit	80 bit	80 bit	80 bit
Total data	368 bit	608 bit	216 bit	312 bit
Total bit-rate	18.4 kbit/s	15.2 kbit/s	10.8 kbit/s	7.8 kbit/s
Procent voice data	66%	80%	44%	61%

Table 5.3: Table showing the bit-allocation for bundled and unbundled speech configurations with two codecs.

Chapter 6

Results

Similar to the last chapter, we divide this chapter into two sets of simulations, showing the results. Each scenario is presented and evaluated for the capacity, intra group latency, jitter, and resource allocation. To compare the result, you must have something to compare with. The capacity is most interesting to compare, and included where necessary. At the end of this chapter, we compare and summarize the results.

6.1 Reference comparison between one-to-one and one-to-many traffic

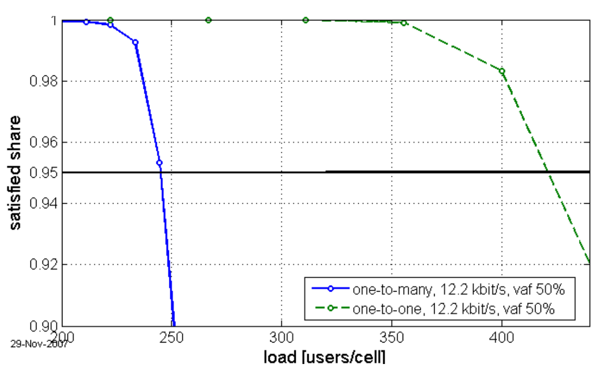

Figure 6.1: User satisfaction with one-to-many compared with one-to-one scenario with 50% voice activity

As described in section 5.3.1 our reference simulation compares the difference in capacity between one-to-one and one-to-many communication. At the capacity limit we evaluate the intra group latency, jitter, and resource allocation.

Figure 6.1 shows a performance comparison between one-to-many and one-to-one communication with a voice activity factor of 50%, each of the following measurements in this scenario will also be with 50% voice activity. The figure shows a decrease in capacity with almost 50% using one-to-many communication. The decrease in capacity is probably because of the lack of diversity

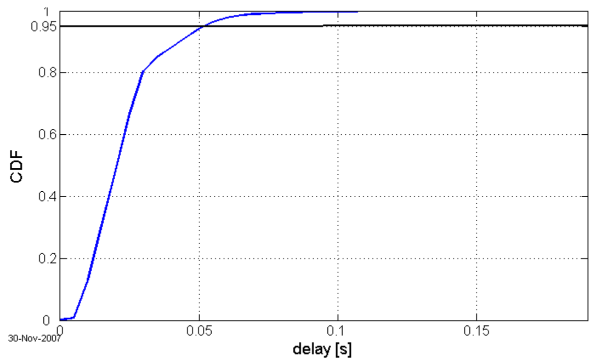

Figure 6.2: Intra group latency at capacity limit with 50% voice activity

between the users, and we investigate this further in the next section when we look at voice activity.

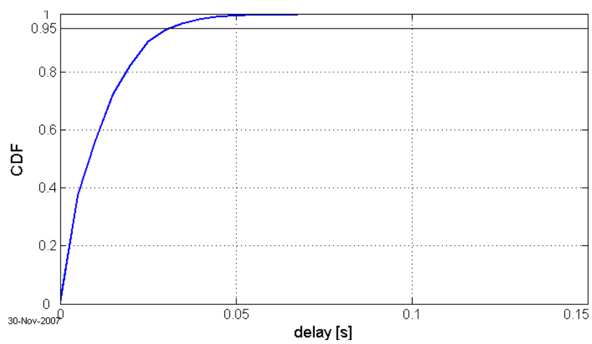

Figure 6.3: Jitter at capacity limit with 50% voice activity

The figure 6.2 shows the Intra Group Latency as defined in 5.1.1. We see a latency around 45 ms, which should not cause any problems in the network. The other latency measurement of importance is jitter, shown in figure 6.3, and is within reasonable numbers. Both measurements are snapshots of the load at the capacity limit.

Looking at figure 6.4 it shows a lack of traffic 50% of the time. This we expect with a voice activity set to 50% and comparing the one-to-one scenario in figure 6.5 we can see a decrease in non-scheduled resource blocks.

6.2 Voice activity effects on capacity

Changes in voice activity are often something the operators calculate to know how much traffic there will be in the network, and how it will affect the capacity limit. The figure 6.6 shows the comparison between the one-to-one and one-to-many traffic models, using 90% voice activity. By comparing the previous figure 6.1 with 50% voice activity we see a large change in the one-to-one result.

As we see in figure 6.7 with the voice activity factors set to 10%, 50%, and 90%. As seen both 50% and 90% are approximately similar with 10% only

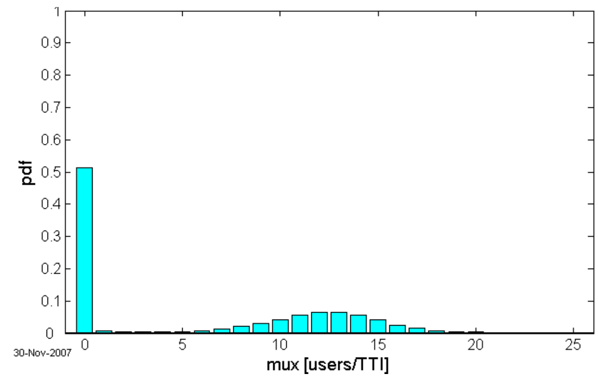

Figure 6.4: One-to-many traffic resource allocation at capacity limit with 50% voice activity

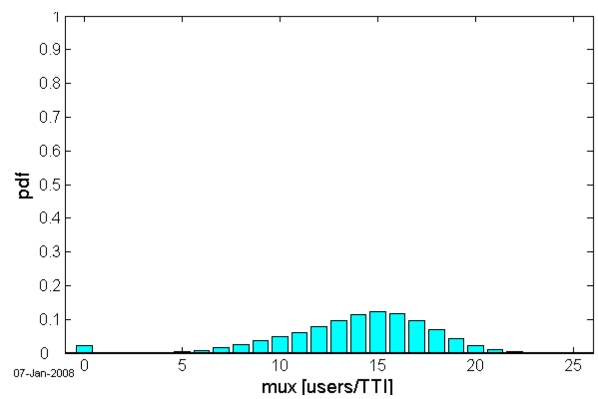

Figure 6.5: One-to-one traffic resource allocation at capacity limit with 50% voice activity

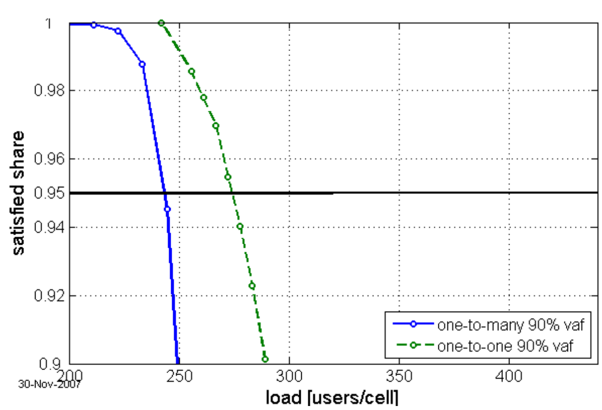

Figure 6.6: User satisfaction with one-to-many conversation compared with a one-to-one VoIP scenario with 90% voice activity factor

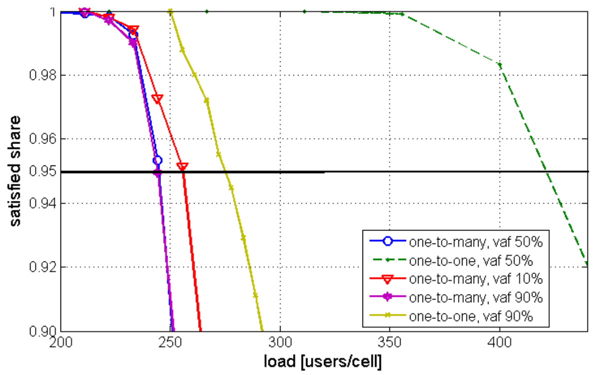

Figure 6.7: One-to-many user satisfaction with voice activity factors of 10%, 50% and 90% with comparison to one-to-one traffic with 50% and 90%

slightly higher. However, the higher capacity with 10% voice activity can be a result from the short talk-spurts used.

6.3 Multiple groups effect on capacity

The common scenario for public safety users is the day-to-day traffic, where the groups are only one or a few users. The worst case scenario, discussed earlier, with a large amount of public safety users distributed within a confined area is less common. There are many numbers of groups that are in use in a network, so a look at their performance is interesting.

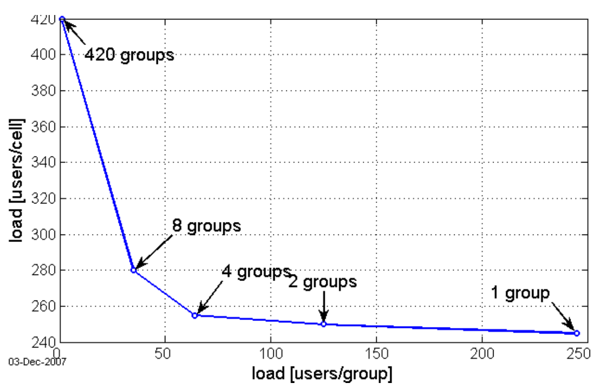

Figure 6.8: Maximum user satisfaction compared to group size

Shown in figure 6.8, is the capacity plot for different number of groups. The plot describes the relationship between the numbers of groups and the total capacity in a cell. Looking closely for one to four groups, we see only a moderate increase in capacity. For the rest of the graph, showing capacity figures for eight groups and increasing until every group contains only one user, the increase in capacity is significantly larger than for smaller number of groups. The increase in capacity for larger number of groups is a result from an increase in diversity between the groups, making the traffic less bursty.

6.4 Improved scheduling

Some packets might fail to be sent, the reasons for this can be many things like a bad connection to the user's area, or a large amount of traffic with higher priority. These packets get old, and in a real-time scenario we consider them lost if older than a few hundred milliseconds. Setting a higher priority would help these packets to be sent and so improving the capacity. The figure 6.9 shows a capacity increase by 12% per cell when setting a higher priority to packets that have been in the sending queue for a while.

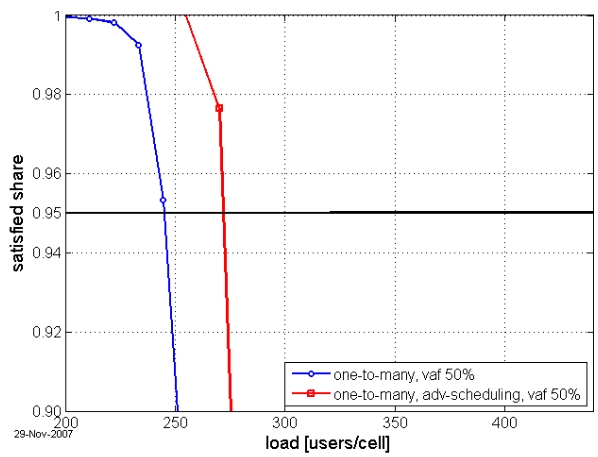

Figure 6.9: User satisfaction with advanced scheduling

As the users increase, looking at figure 6.10 the jitter created in the radio channel is almost constant and should not cause a problem, even if some percentage of users sees a larger jitter, because of more packets gets sent often later in its life and have a greater success of getting transmitted then in reference scenario where they would have the same priority as everyone else.

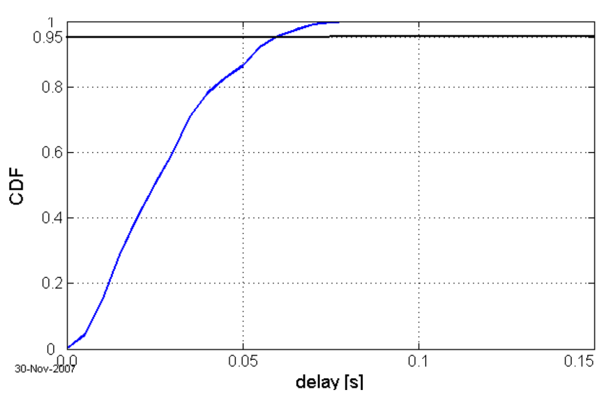

Figure 6.10: Jitter at capacity limit with 50% voice activity and advanced scheduling

Also, the same is with intra group latency seen in figure 6.11. And would be because, at the beginning of transmission the channel is still empty. Meaning that there are no old packages, and the situation is the same as in the reference scenario.

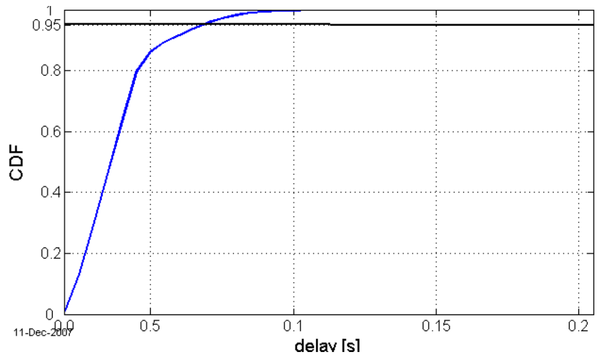

Figure 6.11: Intra group latency at capacity limit with 50% voice activity and advanced scheduling

6.5 Decreasing the codec bit-rate

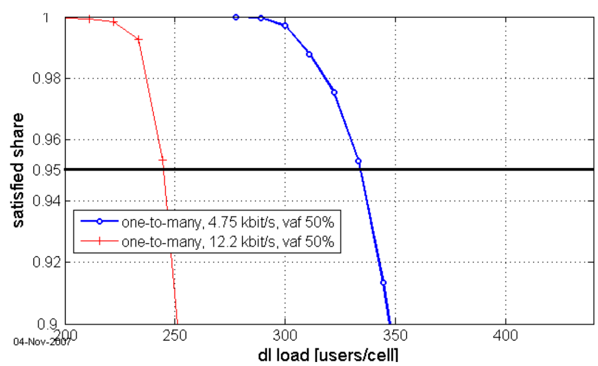

Figure 6.12: User satisfaction with AMR speech codec with bitrate at 4.75 kbit/s

If the capacity is insufficient having a large amount of users in a confined area there may have to be compromises. One of the most obvious compromises is to decrease the speech quality using a speech codec with a low bit-rate. Seeing the results from the simulation in figure 6.12 we can see a 36% capacity increase. Considering that the data part of each speech frame is decreased from 12.2 kbit/s to 4.75 kbit/s one could expect more. However, looking at the whole bit structure in table 5.3, we see only a reduction in speech data. So for a total reduction in data, both the headers and speech data would have to be reduced for a better effect.

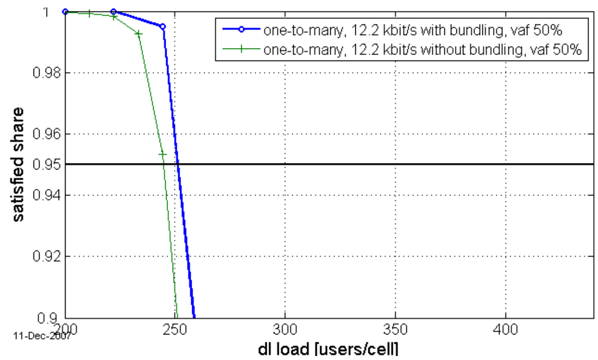

Figure 6.13: User satisfaction bundling two AMR 12.2 kbit/s speech frames in each IP packet

6.6 Bundling of speech frames

As seen in the previous section the headers and overhead data are a large part of the total traffic. One method to reduce overhead data is bundling, as discussed in section 5.4.3. By including two speech frames for each IP packet the overhead data are decreased, only needing the header information for every other speech frame instead of all of them. However, looking at figure 6.13 the increase in capacity is rather small. Going back to table 5.3 the bit-rate is reduced by 3.2 kbit/s, because the high bit-rate is only a small percentage in actual capacity in the cell.

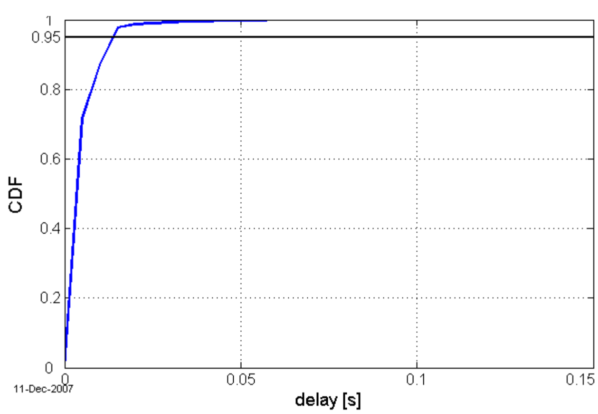

Figure 6.14: Jitter at capacity limit and speech frame bundling using AMR 12.2 kbit/s speech codec

Bundling uses a method to include two, or more speech frames in each IP packet. The issues appear when every other packet have a larger delay. An examination of jitter in figure 6.14 shows similar result as the unbundled scenario, this because we measure only the delay between each IP packet and not each speech frame. A measure on each speech frame would have resulted in a similar result except a spike at zero and twenty milliseconds, as half of them are delayed by twenty milliseconds.

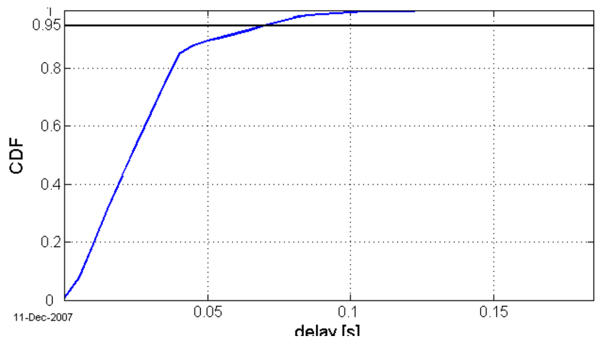

Figure 6.15: Intra group latency capacity limit and speech frame bundling using AMR 12.2 kbit/s speech codec

Looking at the intra group latency in figure 6.15, as expected the delay does increase with a few milliseconds compared with the unbundled scenario. The increase is probably a result from the extra delay in half of the speech frames.

6.7 Combining decreased bit-rate with bundling

In the previous section, we saw bundling used to lower the header size. However, a small header size has only a small effect if the speech data included is large. In the case with decreased speech codec rate the header size is still large. So the combination of should minimize the header more, and seen in figure 6.16 the maximum load per cell does increase with a significant amount because of the decreased header data.

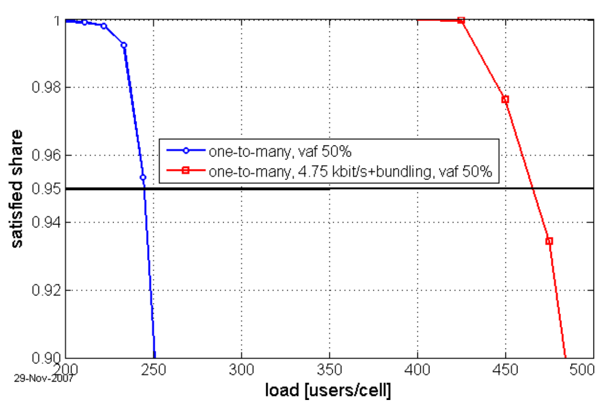

Figure 6.16: User satisfaction bundling 2 AMR 4.75 kbit/s speech frames in each IP packet

6.8 Comparison between different performance improvement techniques

Making a comparison of the optimization techniques discussed in previous sections we can make some conclusions. In table 6.1 we list and compare all results with the reference scenario. Judging from the list, the most interesting method to rapidly increase the capacity is the combination of codec with bit-rate 4.75 kbit/s and bundling, resulting in almost a double in capacity. Some of the drawbacks with bundling and low bit-rate are lower speech quality and longer delays, but within reasonable boundaries. The second method of interest is improving the scheduler, giving almost no effect on the speech quality. The drawback with more complex schedulers is on the system side, increasing the complexity on the base station, with only moderate improvement in capacity.

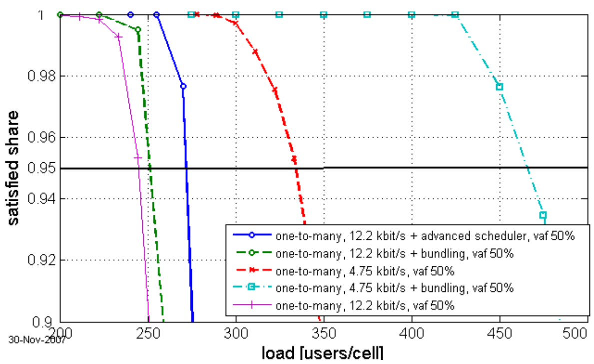

Figure 6.17: Comparison with all performance improving techniques

Simulation scenario	Intra group latency (ms)	Jitter (ms)	Increased capacity (%)
Reference	45	40	-
Improved scheduler	55	65	12
Low bit-rate (AMR 4.75)	45	40	36
Bundling (AMR 12.2)	84	40	2
Bundling (AMR 4.75)	85	40	90

Table 6.1: Comparative performance improvement between each simulation

Chapter 7

Discussion

The object of this work is to simulate traffic behaviors in an one-to-many group communication scenario. From the first couple of simulations in chapter 6 we can draw some conclusions of its behavior. The reference simulations give an idea of the worst case scenario where everyone is in the same group. In the one-to-many scenario, the traffic load on the network will be larger when transmitting, and during the quiet periods there is no traffic at all. This gives 100% load at talk-spurts, and reduces the capacity to 40-50% compared to the one-to-one scenario in a cellular network.

Other simulations suggest a larger diversity in the cell with more groups and an increase in capacity. However, with these simulations we distribute the users throughout the whole network. So the results reflect more the average of number of users in each cell. However, this does give a good idea how the network load increases as the diversity increases.

When the traffic load increases, one expects that the capacity will decrease. However in the simulations on voice activity effects on capacity in the network, the results do not show this. As seen earlier in figure 6.7 the capacity is almost unaffected. Only in the low traffic situation with small talk-spurts we get a small increase in capacity. The diversity at the start and the end of the talk-spurt is likely the cause for the small increase. In this small time window, the load is less then 100% that one gets for the rest of the talk-spurt. Giving a larger percentage to this traffic helps the users transmit their data and get a slightly higher user satisfaction. Also of interest is that the behavior of an one-to-one traffic model increases the capacity, since the traffic is spread out more in time than in the one-to-many traffic model. The one-to-one model is the more conventional and the result is as expected. However, these are the best and worst case scenarios, using a more normal situation with smaller groups we get something between these two.

In an operational network, the operator can tweak the scheduler to increase the capacity. The work done in the simulations with an improved scheduler does increase the capacity to some degree. However, we want to keep the complexity low in the network. A complex scheduler in a computational sense would put larger stress on the network equipment and more expensive to operate. However, a better scheduling algorithm may help optimize the traffic in the time domain. When optimizing traffic in time domain you either help packets that are old and about to expire, or you exploit the time diversity in the channel. Both methods will however increase the delay between listening users and can cause reduce sound quality when users close to each other hear the same sound with a delay between them from multiple speakers.

Another way of increasing capacity is by bundling speech frames. Bundling is often used to increase the capacity in networks where we don't use ROHC to reduce the relative size of the overhead. When the percentage of data in each packet for services likes VoIP when the raw data is low, the header data become very high. However, in LTE networks we use ROHC as common practice. For a relatively large packet like the AMR 12.2 kbit/s speech codec used as standard, gives with ROHC a higher percentage of data, resulting in only a marginally increase in the maximum load in the network cell.

While trying to reduce the header overhead failed the attempt to decrease the AMR speech codec bit-rate to 4.75 kbit/s which is the lowest possible setting succeeded. By this, we increase the headers percentage of each packet. However, the total size per packet is decreased. We notice that the correlation between the traffic per packet and the total traffic in the cell. But, a bit-rate as low as 4.75 kbit/s does not give as good speech quality as 12.2 kbit/s and the public safety users would maybe prefer a better sound quality.

The last attempt made to increase the traffic load in a cell is by identifying that the header overhead in using a low bit-rate speech codec is large, some 50% of the data is nothing but headers. So having a low bit-rate, with bundling we saw in simulations a significant increase in traffic. A combination of bundling and low bit-rate is the best option to handling large groups with one-to-many communication with unicast traffic. Other techniques would be to use multicast to transmit the data. This however might give other problems not investigated in this report. And for smaller group sizes for most day-to-day communication we can see that users in the network receiving the group communication are larger than the already high number of users saw in the simulation. The object of this work is to simulate traffic behaviors in an one-to-many group communication scenario. From the first couple of simulations in chapter 6 one can draw some conclusions of its behavior. The reference simulations give an idea of the worst case scenario where everyone is in the same group. In the one-to-many scenario, the traffic load on the network will be larger when there is traffic, and during the quiet periods there is no traffic at all. This gives 100% load at talk-spurts, and reduces the capacity to 40-50% compared to the one-to-one ungrouped situation that is normally in a cellular network.

7.1 Future work

These simulations have many assumptions. The simulations that we have done are based on a small number of assumptions and specific scenarios that we want to investigate. There are however, many other scenarios and settings that need attention which are beyond the scope of this thesis. Some complementary work could investigate other scenarios with for example, other types of traffic combinations. As the combination of more than one type of communication traffic might be of interest in a scenario where both voice and video is of interest for public safety users.

Another addition to this work would be to make a more extensive review of different schedulers and how they relate to each other. As one aspect is to make as small increased load on the network as possible, is there a way to not hurt the performance for other users with lower priority in the network. And still match the requirements identified as important for public safety users.

One aspect is when transmitting information to a group with so called one-to-many communication and the traffic is in some sense synchronized. The normal model for the traffic patterns at the packet level is often analyzed with a queue theory. However with this new traffic pattern where all packets receive simultaneously. This is however not investigated fully and might need more work later to fully understand this traffic pattern and to develop new scheduling techniques better adapted to this traffic behavior.

Furthermore, the models used here can be developed further, with a more complex scenario. Where both non-prioritized users and prioritized public safety users is on the same network using different traffic models for voice, video, and data, and look at it at a more normal situation then our very simplified version.

This also results in the idea for a work that explores more types of traffic that could be useful. What services needed and unknown today might be useful tomorrow if the network is no longer an issue, and no technical limitation? That makes it also of interest that when using somewhat normal equipment. What improvements need to be done to the software platforms available today to support these new services?

Chapter 8

Conclusions

This master thesis work simulates traffic behaviors in a one-to-many group communication scenario in a 3GPP LTE cellular network.

The simulations show that using one-to-many communication may decrease the capacity by 40-50% in a network compared to one-to-one communication using 50% voice activity. This because of the different traffic models used. Also, the voice activity only slightly affects the capacity with one-to-many communication as the number of users will be the same for each TTI as the talk-spurt duration increases.

As expected, when simulating more groups then a single group, the capacity will increase because the diversity between groups increases from the worst case with only one large group to the best case that is the one-to-one scenario.

The use of complex schedulers is shown to have a moderate effect on the capacity. However, depending on the wanted results; other improvements may help more. The lower codec bit-rate, in particular if combined with bundling speech frames, gives a large effect on the capacity.

Bibliography

[1] 3GPP. Summary of Requirements identified during 3GPP RAN long term evolution workshop. REV-WS 044, 3rd Generation Partnership Project (3GPP), 2004.

[2] 3GPP. Proposed Study Item on Evolved UTRA and UTRAN. RP 040461, 3rd Generation Partnership Project (3GPP), Mar 2006.

[3] 3GPP. Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN). TR 25.913, 3rd Generation Partnership Project (3GPP), Mar 2006. v7.3.0.

[4] 3GPP. Feasibility study for evolved Universal Terrestrial Radio Access (UTRA) and Universal Terrestrial Radio Access Network (UTRAN). TR 25.912, 3rd Generation Partnership Project (3GPP), June 2007. v7.2.0.

[5] 3GPP. IP Multimedia Subsystem (IMS); Multimedia Telephony; Media handling and interaction. TS 26.114, 3rd Generation Partnership Project (3GPP), Sep 2007. v7.3.1.

[6] 3GPP. Technical Specification Group Services and System Aspects; Performance characterization of the Adaptive Multi-Rate (AMR) speech codec. TR 26.975, 3rd Generation Partnership Project (3GPP), Jun 2007. v7.0.0.

[7] 3GPP. Vocabulary for 3GPP Specifications. TR 21.905, 3rd Generation Partnership Project (3GPP), Sep 2007. v8.2.0.

[8] TETRA MoU Association. Push To Talk over Cellular (PoC) and Professional Mobile Radio (PMR). *TETRA*, 2004.

[9] C. Bormann. RFC3095: RObust Header Compression (ROHC): Framework and four profiles: RTP, UDP, ESP, and uncompressed. *RFC*, 2001.

[10] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming. The Long-Term evolution of 3G. *Ericsson Review*, (No.2), 2005.

[11] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming. *3G Evolution HSPA and LTE for Mobile Broadband*. Academic Press, 2007.

[12] Lijing Ding and Rafik A. Goubran. Assessment of Effects of Packet Loss on Speech Quality in VoIP. *Department of Systems and Computer Engineering, Carleton University*, 2003.

[13] ETSI. Digital Video Broadcasting (DVB) Framing structure, channel coding and modulation for digital terrestrial television. EN 300 7449, ETSI, Jun 2004. v1.5.0.

[14] WiMAX Forum. WiMAX End-to-End Network Systems Architecture Stage 2-3. Technical report, WiMAX Forum, July 2007. v1.1.0.

[15] S. Frost and Benjamin Melamed. Traffic Modeling For Telecommunications Networks. *IEEE Communications Magazine*, 1994.

[16] S. Frost and Benjamin Melamed. Can we talk, Public safety and the Interoperability challange. *National Institute of Justice*, 2000.

[17] ITU-T. The E-model, a computational model for use in transmission planning. G 107, Telecommunication standardization sector of ITU, Mar 2005.

[18] Krisberedskapsmyndigheten. Drift- och tjänstespecifikation för Rakelsystemet. KBM 300, Krisberedskapsmyndigheten, Jun 2007. v5.0.

[19] Daniel Electronics LTD. P25 Training Guide. http://www.p25.com/resources/P25TrainingGuide.pdf.

[20] Macom. P25 Technical overview. http://www.macom-wireless.com/federal/P25

[21] OMA. Push to talk over Cullular (PoC) - Architecture. Technical report, Open Mobile Allience (OMA), Nov 2006. v1.0.1.

[22] National Task Force on Interoperability. Why Can't We Talk, Working Together To Bridge the Communications Gap To Save Lives. *NTFI*, 2003.

[23] Peter Sievering and Björn Robbel. A Statistical Traffic Analysis of Group Speech Communications in the German TETRA Trial Aachen. In *5th European Personal Mobile Communications Conference*, pages 391 –396. IEEE, 2003.

[24] J. Sjoberg, M. Westerlund, A. Lakaniemi, and Q. Xie. RFC3267: Real-Time Transport Protocol (RTP) Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs. *RFC*, 2002.

[25] TETRA. TETRA: General Design. EN 300 392-1, ETSI, Jun 2005. v1.3.1.