Master's Thesis

Analysis of QoS and QoE aspects for copper-based mobile backhaul

Carl Bjerggaard
Johan Hallberg

Department of Electrical and Information Technology,
Faculty of Engineering, LTH, Lund University, June 2014.

Analysis of QoS and QoE aspects for copper-based mobile backhaul

A master thesis performed in collaboration with Lund University and
Research Area Broadband Technologies at Ericsson Research
Lund, Sweden, 2013

Carl Bjerggaard and Johan Hallberg
Faculty of Engineering, LTH
Lund University

LTH examiner: Stefan Höst

LTH supervisor: Jens A Andersson

Ericsson supervisor: Daniel Cederholm

Sammanfattning

Uppskattning av kvaliteten hos strömmad media i en kopparbaserad mobil backhaulförbindelse

Framtidens smarta nät kommer att ställa stora krav på kapacitet, tillgänglighet och kvalitet då trafiken till stor del kommer att utgöras av strömmade medietjänster utanför nätägarnas kontroll. Genom att jämföra fel under strömmande av en film med kvaliteten på den mottagna filmen har vi lyckats skapa en matematisk modell som uppskattar kvaliteten hos strömmad media, redan innan filmen börjat skickas.

Mobiltelefoner har länge varit ett vanligt inslag i våra liv. I början nöjde vi oss med att ringa och skicka SMS och senare även MMS men runt år 2007 inträffade något som radikalt ändrade vårt användande av samt våra förväntningar på mobiltelefoni, nämligen lanseringen av smarta telefoner. Helt plötsligt förväntade vi oss förutom att kunna ringa och SMS:a även att kunna skicka e-mail, surfa, strömma film och musik, spela online spel m.m. Vidare blev det under denna tid mer regel än undantag för bärbara datorer och surfplattor att ha tillgång till en bra internetuppkoppling oavsett plats. För tio år sedan hade få hört talas om IPTV och Video on Demand men idag använder många Netflix, HBO och diverse play-kanaler, så som SVT Play, TV4 Play och liknande, dagligen. Trots att användare av IPTV ofta delar uppkoppling med flera andra förväntar de sig ändå samma kvalitet av IPTV som i det markbundna nätet.

Denna utveckling har lett till att fler enheter än någonsin är uppkopplade mot det mobila nätet och för att tillgodose kundernas nya hårdare krav på kapacitet, tillgänglighet och kvalitet har EU startat projektet COMBO i samarbete med ett flertal europeiska mobiloperatörer, telekommunikationsföretag och universitet för att avlasta dagens mobila nät med hjälp av det befintliga fasta nätet. Genom att placera många små celler nära användarna, dvs. små mobilmaster inte helt olika en vanlig trådlös uppkopplingspunkt, kan trafik till och från mobila enheter som t.ex. mobiltelefoner och surfplattor avlastas

från de stora mobilmasterna till de små cellerna och vidare till backhaulen, vilken utgör den del av det fasta nätet som är närmst användaren. Detta är ett av många sätt att lösa framtidens krav på det mobila och fasta nätet och innan ett så omfattande projekt kan genomföras krävs det att man noga undersöker de olika alternativen samt konsekvenserna av dessa.

Då DSL är en av de vanligaste teknikerna för fast internetanslutning i Europa idag har vi undersökt hur den ökade belastningen påverkar möjligheterna att strömma media via denna teknik samt hur man kan uppskatta hur användaren upplever kvaliteten hos denna i sin mobila enhet. Då DSL är en teknik känslig för yttre störningar, så som radiosignaler och brus etc., är det viktigt att undersöka hur dessa påverkar prestandan i nätet och därmed den upplevda kvaliteten hos slutanvändaren.

Genom att analysera vilka fel som inträffar under strömning av en film och jämföra detta med kvaliteten på den mottagna filmen hos slutanvändaren har vi lyckats skapa en matematisk modell som uppskattar kvaliteten hos strömmad media. Faktum är att modellen kan beräkna vilken kvalitet en film teoretisk skulle få på nätet just nu vilket innebär att ingen film behöver sändas för att modellen ska fungera. I ett framtida s.k. smart nät skulle modellen kunna användas för att avgöra vilken förbindelse i det fasta nätet som vid en viss tidpunkt är mest lämplig för att avlasta det mobila nätet, baserat på förbindelsens potential för att strömma media just då. Detta skulle bidra till en konstant hög kvalitet för slutanvändaren och även göra det möjligt att ingå avtal om vilken kvalitet på strömmad media ett nät ska tillhandahålla i olika delar av nätet.

Abstract

As discussed in the FP7 project COMBO, future mobile backhaul technologies will partly rely on fixed access networks, with in turn may be based on Digital Subscriber Line (DSL) technology. With the expected Fixed Mobile Convergence (FMC) new challenges regarding Quality of Service (QoS) and Quality of Experience (QoE) will be faced.

Mobile Internet Protocol Television (IPTV), together with Over The Top (OTT), are amongst the most important Internet based services today and will continue to be so in the future. To provide the mobile end-users with the same Quality of Experience as with Terrestrial Broadcasted Television a higher capacity in the network, and especially in the Radio Access Network (RAN), is needed. To attain this, small cells can be connected to the pre-existing copper, fibre or microwave access network near the subscribers, thus extending the mobile backhaul. This way of reusing the pre-existing access networks is only one amongst other possible solutions to the problem above but has the benefit of a reduced time to market as well as an improved cost-effectiveness. Since a majority of the existing wired access networks consists of Digital Subscriber Line based technology, some future mobile backhauls will have to rely on this technology. Due to the DSL transmission method's sensitivity to external interference it is of great importance to understand how this interference affects the network performance metrics and QoS parameters, and thereby the QoE, of a network.

Throughout this report, relations and correlations between the most suitable network performance metrics and parameter estimation methods are analysed, tested and verified by several experiments on a real DSL scenario in a lab environment. By measuring the network performance metrics on the physical-level together with the objective QoE parameter VQM at user-level, a method for predicting QoE based on network performance metrics is developed. Finally, the developed method is evaluated by introducing traffic on OSI-layer 3 and above and verifying that the method is still valid.

Acknowledgements

At first, we would like to show our deep respect and gratefulness to our supervisor Jens A. Andersson, Lecturer at Faculty of Engineering LTH, for being a great support, not only during the thesis work but throughout our whole education. He gave us this fantastic opportunity to work with a fun, challenging and important area of research in cooperation with Ericsson and has taught us almost everything we know within the area of Network Communication. Furthermore, we would like to thank our examiner Stefan Höst, Associate Professor at Faculty of Engineering LTH, for many good and thorough explanations about the basics of our research as well as for his support during our work. In addition, we would like to thank our supervisor at Ericsson, Daniel Cederholm, for answering a ton of questions during the time of our work as well as supporting us in our research, by inviting us to Ericsson multiple times for guidance. Finally, we would like to thank our girlfriends for patiently supporting us during our education and thesis work.

Contents

List of Figures

List of Tables

Abbreviations

3GPP	Third Generation Partnership Project
BER	Bit error rate
CPE	Customer-premises equipment
CRC	Cyclic Redundancy Check
CV	Code violation
DLCM	Dynamic Line Code Management
DMOS	Differentiated Mean Opinion Score
DMS	Degradation Mean Opinion Score
DMT	Discrete Multi-Tone modulation
DQM	DSL Quality Management
DSL	Digital Subscriber Line
DSLAM	DSL access multiplexer
DSM	Dynamic Spectrum Management
ES	Errored Second
FDM	Frequency Division Multiplexing
FEC	Forward Error Correction
FEXT	Far End Crosstalk
FMC	Fixed and Mobile network Convergence
FTTCurb	Fiber To The Curb
INP	Impulse Noise Protection
IPTV	Internet Protocol Television
ISP	Internet Service Provider
ITU-T	International Telecommunication Union
LATN	Loop attenuation
LOF	Loss of Frame
LOS	Loss of Signal
LOSS	Loss of Signal Second
LPR	Loss of Power
MOS	Mean Opinion Score
MSE	Mean Squared Error
NEXT	Near End Crosstalk

NR-B	No-Reference Bit stream
NR-P	No-Reference Pixel
OTT	Over the Top
PEIN	Prolonged Electrical Impulse Noise
PMD	Physical Media Dependent
PSD	Power Spectral Density
PSNR	Peak Signal to Noise Ratio
PSTN	Public Switched Telephone Network
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
REIN	Repetitive Electrical Impulse Noise
RFI	Radio Frequency Interference
RS-code	Reed-Solomon code
RTP	Real-time Transport Protocol
SATN	Signal attenuation
SDTV	Standard Definition Television
SEF	Severely Errored Frame
SES	Severely Errored Second
SHINE	Single High level Impulse Event
SLA	Service Level Agreement
SNR	Signal-to-Noise Ratio
SNRM	Signal-to-Noise Ratio Margin
SSM	Static Spectrum Management
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
UAS	Unavailable seconds
UDP	User Datagram Protocol
VN	Virtual Noise
VQM_VFD	Video Quality Metrics Variable Frame Delay

Chapter 1

Introduction

1.1 Background

Mobile Internet Protocol Television, together with Over The Top, are amongst the most important Internet based services today and will continue to be so in the future. This puts new requirements on both fixed access and mobile networks to provide the end-users with the same Quality of Experience as with Terrestrial Broadcasted Television. To achieve the same quality in mobile networks a higher capacity in the network, and especially in the Radio Access Network (RAN), is needed. To attain this, small cells, e.g. microcells, picocells and femtocells, can be placed close to the subscribers and connected to the pre-existing copper, fibre or microwave access networks. Thus, the mobile backhaul, the intermediate link between the core network and the wireless access network, will be extended to the small cells near the subscribers. This way of reusing the pre-existing access networks is only one amongst other possible solutions to the problem above but has the benefit of a reduced time to market as well as an improved cost-effectiveness. Since a majority of the existing wired access networks consists of Digital Subscriber Line based technology, some future mobile backhauls will have to rely on this technology. Moreover, the fixed and mobile access networks of today have been developed independently and are strictly separated both functionally and physically. Hence, they differ in regard to network interfaces, access technologies, network functionalities, etc., and thus must be able to share the same network infrastructure in the future. This change will have an impact on the QoS for mobile IPTV and OTT services, e.g. signal-to-noise ratio, latency, bit-rate, packet loss and particularly the QoE [**?**][**?**].

1.2 Problem Description

Since xDSL is a transmission method sensitive to external interference such as crosstalk, impulse noise and radio frequency interference (RFI), it is of great importance to understand how this interference affects the network performance metrics and QoS parameters of an xDSL based network. To obtain a high capacity in copper-based mobile backhauls and to guarantee a certain level of network performance, QoS and QoE, the interrelations between different metrics and parameters on different lower layers of the so-called OSI-model need to be investigated. By understanding this interaction, information about QoS and QoE parameters on higher layers can be deduced and analysed to give an understanding of the perceived quality at the end-user.

As described in Section 2.8 many studies have been performed on the relation between QoS and QoE parameters on OSI-layer 3 and above, while cross-layer investigations on lower layers, considering network performance metrics, are conspicuous by their absence. Since OTT content cannot be prioritised on lower layers in the same way as for example IPTV content, knowledge of how lower layer metrics and parameters affects the QoE of the OTT content is incredibly valuable.

Further, many methods for estimating QoS and QoE parameters are performed on the subscriber end of the loop and hence not always feasible from the network operator's perspective. Therefore, it is necessary to analyse the applicability of the different methods on the network operator end of a copper-based mobile backhaul [**?**][**?**].

1.3 Limitations

Since the scope of this thesis was quite broad and included both compilation and discussion of QoS parameter estimation methods and their applicability on mobile backhauls, as well as verification of at least one such method, some limitations on the project had to be stated.

With the provided lab setup different xDSL technologies such as ADSL, ADSL2plus and VDSL2 could be configured and with VDSL2 as the current state of the art technology in Europe, this technology was chosen [**?**]. The lab setup also offered the possibility to configure a multitude of different protection mechanisms for the xDSL connection and since it was not possible to choose all configurations we chose to use a Minimum Impulse Noise Protection of 2 DMT symbols together with a Maximum Interleave Delay of 4ms for the tests. However, the second parameter was automatically determined during line initialisation based on the first parameter together with the current maximum achievable data rate [?].

Another consequence of the lab setup was that the mobile access part of the network was missing. Thus, we where not able to test the impact on the perceived QoE caused by the wireless part of the backhaul, leaving us with the option to make assumptions about this.

Furthermore, we decided to only use half standard definition television (SDTV) quality for streaming during the tests due to media with higher qualities taking far too much time to analyse after streaming. We also chose to only stream OTT content since the majority of the today's mobile network traffic consists of non-operator traffic such as OTT [?]. Thus, we considered this to be the most important technique to investigate regarding QoS.

Finally, we chose to measure QoS parameters on the physical layer of the OSI-model while estimating the QoE on user-level, thus omitting measurements on the layers in between. As described in Section 2.8, a lot of research has been done on the relationship between OSI-3 and the layers above and therefore we considered the relation from OSI-1 to user-level as more interesting.

1.4 Structure of Report

In Chapter 1 the background and motivation for this thesis is presented together with the limitations for the scope of the thesis. In Chapter 2 the theory needed to comprehend the thesis is compiled. Next, Chapter 3 is divided into two sections; Method and Test Results. In method, the steps and preparations needed in order to perform the tests are shown, followed by the test suite. In the second half, test result, the results without any conclusions are presented. Then, a discussion about the models is conducted, with a focus on the meaning of the findings. Finally, the thesis is ended by the authors' conclusions regarding the results, their value and their relations to COMBO together with some recommendations for future enhancement of the research.

Chapter 2

Theory and Empiry

2.1 Quality of Service

QoS is defined in ITU-T E.800 as:

> *Totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service.*

This means that QoS is a set of measurable parameters which describes the performance of the underlying system, e.g. bit rate, bit error rate etc. QoS is highly dependent on the system which performance is monitored and it can be measured on one link, in a net or all the way between two users, i.e., end-to-end QoS, as illustrated in Fig. 2.1. Further, the user experience is an important QoS parameter which ensures the quality of the delivered content. This parameter will be discussed further in Section 2.2.

The QoS can be divided into four different parts:

- The QoS offered by the service provider.
- The QoS achieved by the service provider.
- The QoS perceived by the customer.

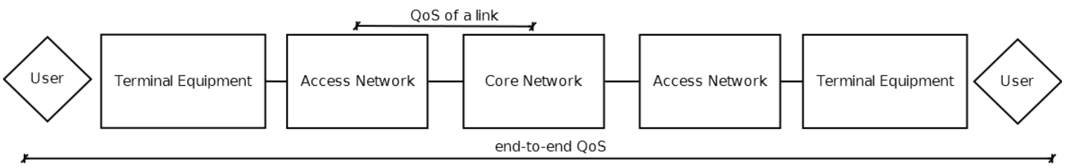

Figure 2.1: *Scope of QoS* Showing the scope of where QoS parameters can be measured.

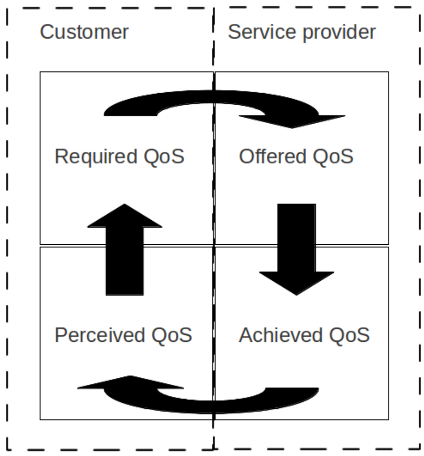

Figure 2.2: *The four parts of QoS* Illustrates the four different parts of QoS as well as how they affect each other.

- The QoS required by the customer.

This somewhat reflects the complexity of QoS; a service provider offers say a bit rate of 100Mb/s but only achieves 10Mb/s, however the customer may still perceive this as a good service due to the low price since the customer requires a cheap internet connection. A connection between the four parts is also implied, i.e., the QoS perceived by the customer will change the QoS required by the customer. In turn, this will change the QoS offered and achieved which will affect the QoS perceived by the customer. The four parts and their relations are illustrated in Fig. 2.2 [**?**].

2.2 Quality of Experience

As mentioned in Section 2.1, QoE is a subset of QoS meaning it is a way of evaluating the quality of a system. What distinguishes QoE from other QoS parameters is its subjective nature, i.e., to accurately measure QoE tests with real users are required. In [**?**] ITU-T suggest several different methods of quantifying a subjective observation of the quality of a service. One suggested method is Degradation Mean Opinion Score (DMS), which is based on the adopted Mean Opinion Score (MOS) method. In this method the user rates the quality of the system on the following five points degradation scale:

- 5 - Degradation is unnoticable.
- 4 - Degradation is noticable but not annoying.
- 3 - Degradation is slightly annoying.
- 2 - Degradation is annoying.
- 1 - Degradation is very annoying.

Since subjective tests takes a lot of time and are in need of continuous evaluation of test persons they are not feasible to perform in a real-time system. However, there are methods to estimate the QoE based on QoS parameters with a fairly high accuracy. When estimating QoE of media streams the methods can be divided into three different categories depending on the degree of access to the original media:

- Full-reference methods.
- Reduced-reference methods .
- No-reference methods.

The full-reference methods has full access to both the sent and received media and thus can evaluate the degree of degradation experienced by the user by comparing the received and the original media. The reduced-reference method on the other hand has some knowledge of the sent media but full access to the received media and combines this with an analyse of the real-time data stream to estimate the QoE. Finally, the no-reference methods has no access to the sent media and has to rely on analysing the real-time data stream to estimate the QoE [**?**].

Further, no-reference methods can be divided into two types:

- No-reference pixel (NR-P) type, refers to no-reference methods which uses the decoded video stream to evaluate the QoE.
- No-reference bit stream (NR-B) type, refers to no-reference methods which uses the bit stream to evaluate the QoE.

There are also proposed no-reference methods which combines both NR-P and NR-B to evaluate the QoE, such as [**?**]. Some of these methods will be discussed further later in this thesis.

2.3 Over The Top Content and IPTV

Over the top content is defined as a network service without the direct involvement of the internet service provider (ISP). According to [?] Youtube is the most popular OTT media service today with over three billion hours of video being streamed each month, this roughly corresponds to 2.2 million full-time jobs. Further, Youtube has grown from 63 million visitors each month in 2006 to 800 million visitors in 2012 and is expected to continue to grow [?][?]. In 2011, 51% of all internet traffic was streamed video and it is expected to grow to 55% in 2016 [?]. From 2010 to 2011 video streaming had increased by 93% to 39% of the total mobile data sent and the usage of mobile video is estimated to have an annual growth of 28% over the next five years [?].

This development puts new demands on the capacity of the mobile networks which in turn pursues the need for convergence between fixed and mobile networks, i.e. FMC. Subscribers have gotten used to the quality from terrestrial TV transmissions and will expect nothing less from OTT services. Failure to deliver this quality may result in dissatisfied customers complaining or even cancelling their subscriptions. Hence, it is important to be able to monitor the quality of the delivered content to avoid complaints.

2.4 xDSL

xDSL is an umbrella term for a family of technologies that achieves high transmission rate over the existing PSTN cables. The xDSL family contains several techniques but since the lab is set up to use VDSL2 only that technique will be described here. VDSL2 is based on the recommendation for VDSL and is compatible with multimode operability. VDSL2 utilises frequencies from 0 up to 30MHz, depending on DSL version and settings, of the copper cable to send and receive data, or rather 25kHz - 30MHz since the regular phone and a guard band uses 0 - 25kHz. Because an ordinary subscriber is more dependent on downstream traffic, e.g. streaming a movie or downloading a file, most DSL techniques are designed to be asymmetric. This means that the frequency band is not divided equally between downstream and upstream but rather giving the downstream a much wider frequency band. Further, VDSL2 uses Discrete Multi-tone modulation (DMT) which divides the frequency band into several channels, or tones, each 4.3125kHz wide. Depending on the signal-to-noise ration (SNR) of the frequency spectrum the tones can transmit between one and fifteen bits. This gives VDSL2 a top capacity of 200 Mbps at short distances. The capacity of a VDSL2 line, depending on version and settings, may deteriorate quickly, from 100 Mbps at 0.5 km to 50 Mbps at 1 km [?][?]. Since the physical link of DSL is a copper cable susceptible to a variation of different noises, see Section 2.6, it is imperative to use different Dynamic Quality Management levels, see Section 2.5, to minimise the distortion. An important aspect in DSL used in this thesis is the fact that the DSLAM is always transmitting on the line and if no data is available for transmission it will send idle cells. This is because the re-initialise process takes time and to avoid it the DSLAM keeps transmitting cells.

2.5 DSL Quality Management

DSL Quality Management (DQM) is a concept invented by Broadband Forum and is described in TR-197 [?]. DQM comprises a variation of techniques and strategies to assure an adequate quality and stability of the DSL line. By altering different configuration parameters the quality of the DSL line can be changed. The following parameters may influence line quality and stability:

- Data Rate
- Transmit Power Spectral Density (PSD)
- Spectral Masking
- Signal-to-Noise Ratio Margin (SNRM)
- Impulse Noise Protection (INP)
- Virtual Noise (VN)
- Re-initialisation profile

DQM can be divided into two main categories: Dynamic Spectrum Management (DSM) and Dynamic Line Code Management (DLCM).

2.5.1 Dynamic Spectrum Management

The main purpose of DSM is to diminish the crosstalk between different lines in a cable. This is mostly accomplished by adjusting the transfer power and frequency band used. DSM is further divided into four levels: static spectrum management (SSM) and DSM L1 - L3. In SSM no actual real-time management is performed but rather the system is preconfigured to limit the transmit spectrum to a certain predefined value. DSM L1 deals with impulse noises and DSM L2 addresses crosstalk from short lines in to long lines. DSM L3, also known as vectoring, virtually removes crosstalk between lines which greatly increases the performance of DSL [**?**][**?**].

2.5.2 Dynamic Line Code Management

The idea of DLCM is to improve the performance of a DSL system by continuously altering the line coding parameters from without the current state of the line. DLCM techniques primarily uses a combination of INP and delay parameters to prevent the influence of various noise types. By identifying crosstalk and other impairments mentioned in Section 2.6 on the line the reliability and data rate of the DSL system can be increased.

2.5.3 Embedded DSL Quality Techniques

A DSL modem contains several embedded features which can be used to improve the performance of the system [**?**]. However, the most important ones in this case is the impulse noise protection (INP) which refers to a systems ability to protect itself against impulse noise, see Section 2.6.2. In DSL a good INP can be achieved either by Forward Error Correction (FEC) and interleaving or by retransmission [**?**].

DSL utilises Reed Solomon code as FEC technique which together with interleaving is a very powerful tool to mitigate impulse noise. Reed Solomon code (RS-code) is a technique developed to detect and/or correct errors in a sequence of symbols. By adding t extra symbols to the data up to t erroneous symbols can be detected and up to $t/2$ symbols can be corrected. Interleaving is a technique used to improve the performance of the RS-code by scrambling the symbols in the outbound data before sending it and then reordering it upon arrival. This reduces the likelihood that several consecutive symbols are destroyed by noise. Assume that the following data are to be sent over a noisy DSL line using RS code and no interleaving.

Data:	*aaaabbbbccccddddeeee*
Received distorted data:	*aaaab____cccddddeeee*

The received data is to distorted to be restored using RS-code and thus must be discarded and/or retransmitted. However, consider now the case where RS-code is used in conjunction with interleaving.

Data:	aaaabbbbccccddddeeee
Interleaved:	abcdeabcdeabcdeabcde
Received distorted data:	abcde____eabcdeabcde
Deinterleaved data:	a_aab_bbbc_ccd_dddeeee

In the same scenario the received data can be fully restored using RS-code simply by interleaving the outbound data. However there is a trade-off to consider when using INP with RS-code and interleaving. Each RS-code symbol added increases the overhead on the channel and thus decreasing the effective data rate. Further, in order to interleave the system must gather some data to interleave, thus adding an interleave delay. The interleave delay is a constant delay, and deeper interleaving results in a longer delay, which may result in an virtually jitter free environment. The interleave delay may however have a negative impact on certain real-time applications.

2.6 DSL Quality Impairments

2.6.1 Crosstalk Noise

Crosstalk is one of the noise types that affects the performance of DSL the most and is typically caused by the signal in one line leaking into an adjacent line, i.e. electromagnetic coupling, and hence decreasing the SNR of that line. When a transmitted signal leaks into an arriving signal of an adjacent cable pair at the same end of the transmission line, it is called near end crosstalk (NEXT). This type of interference can severely impair the transmission due to the transmitted signal being much stronger than the attenuated incoming signal, leading to an already attenuated signal being heavily impaired. On the other hand, if a transmitted signal leaks into another transmitted signal of an adjacent cable pair at the same end of the line, the second signal will consist of an interfered signal being gradually attenuated by the loop. This noise is called far end crosstalk (FEXT) [**?**].

Crosstalk originating from a transceiver running the same type of DSL as the disturbed transceiver is called self crosstalk while crosstalk between different DSL types, for example SHDL and VDSL, is referred to as foreign crosstalk, or more commonly, alien crosstalk. By using frequency division multiplexing (FDM) in an DSL system, the upstream and downstream frequency bands are separated from each other, thus effectively eliminating any self NEXT in the received signals. However, foreign crosstalk cannot be eliminated in this way due to different DSL types transmitting on different frequencies and thus overlapping each other [**?**][**?**].

The crosstalk noise can further be divided into stationary, short-term stationary and time-varying crosstalk. Stationary crosstalk can occur on a line with constant line conditions and a constant number of disturbing sources, leading to a constant crosstalk. This constant noise will lower the SNR and thus the performance and range of the DSL connection if not handled properly, but does not decrease the stability of the connection. Short-term stationary crosstalk can occur on passive DSL systems, e.g. time division multiple access (TDMA) systems, where the transmitted signal has the same effect as a stationary signal during a short time interval, and hence can disturb a stationary DSL system leading to high error rates. Therefore, stationary and short-time stationary DSL systems are normally separated from each other physically or technically [**?**].

2.6.2 Impulse Noise

Impulse noise is normally caused by electromagnetic radiation on the subscriber end of the loop, caused by electrical devices such as power lines. This type of noise typically occurs with short bursts with a high amplitude covering a wide band of frequencies. Since impulse noise bursts occurs rather infrequently the SNR is less affected than the bit-error rate. Without a proper mechanism protecting against this an impulse noise peak can corrupt one or several consecutive DMT symbols causing the quality of the end service to be significantly lowered [**?**].

Impulse noise is divided into three categories [**?**]:

- Repetitive Electrical Impulse Noise (REIN)
- Prolonged Electrical Impulse Noise (PEIN)
- Single High level Impulse Noise Event (SHINE)

REIN occurs with a fixed inter-arrival time and is often originated from the power network giving it a frequency of 100 Hz in Europe. The noise burst are typically short, around 100 μs, and will only corrupt one or two DMT symbols. PEIN is a long noise, usually between 1 ms and 10 ms, with a random inter-arrival rate. Due to the long noise bursts in REIN it is difficult to avoid symbol loss using only INP and in this case retransmission might be preferable. Finally, SHINE is burst of noises with a duration above 10 ms making it difficult to mitigate.

2.7 Performance metrics and parameters

In this section the most important performance metrics used in this thesis are explained. Further, the performance metrics are divided into sections depending on which OSI-layer they can be measured on.

2.7.1 User level

2.7.1.1 Mean opinion score

Mean opinion score is commonly used to quantify users' subjective opinions about different services. By evaluating the user experience on a scale between 1 and 5, an estimation of the average user experience can be found. It is common to try to map some of the OSI-layer 7 parameters from the section below with MOS to achieve an estimation of the QoE, as discussed in Section 2.8.

2.7.1.2 Differentiated Mean Opinion Score

Differentiated Mean Opinion Score (DMOS) is based on MOS but uses another scale, usually 0 to 100.

2.7.2 OSI-layer 7

2.7.2.1 Video quality metrics

A challenge when using a full reference QoE estimation method to calculate video quality is time varying delay of the video frames. Video Quality Metrics Variable Frame Delay (VQM_VFD) is a model developed to handle the varying frame delay which allows it to track subjective quality over a wide range of different videos. Further, VQM_VFD achieves 0.9 correlation with the subjective tests' DMOS [**?**].

2.7.2.2 Peak Signal to Noise Ratio

Peak Signal to Noise Ratio (PSNR) is a measurement of the ratio between the maximum possible signal, the peak signal, and the noise affecting the signal. It is often used in reconstruction of lossy compression and attempts have been done to map it to QoE [**?**].

2.7.3 OSI-layer 3

2.7.3.1 Packet-loss probability

Packet-loss probability refers to the probability that a packet sent from the sender will not reach the receiver. According to RFC 6374 [**?**] this measurement comprises both the transmit loss and the receive loss between two arbitrary nodes connected via a bidirectional channel.

2.7.3.2 Delay

Delay refers to the time it takes for a bit of data to travel from one node or end point to another and is measured in milliseconds. According to RFC 6374 [**?**] there are three different delays to measure between two arbitrary nodes A and B:

- The forward one-way delay (A to B);
- The reverse one-way delay (B to A);
- The two-way delay (A to B to A).

2.7.3.3 Delay variation

According to RFC 5481 [**?**] delay variation refers to either:

- Packet-delay variation - the difference in delay between an arbitrary packet in a specified interval and the packet with the least delay in said interval.
- Inter-Packet delay variation - the difference in delay between two consecutive packets.

2.7.3.4 Net bit-rate

Net bit-rate refers to the rate at which bits are being sent over a network, excluding the physical layer overhead, and is measured in kbps.

2.7.4 OSI-layer 1

2.7.4.1 Attenuation

Attenuation refers to the deterioration of the signal and is divided into two categories:

- Loop attenuation (LATN),
- Signal attenuation (SATN).

2.7.4.2 Bit-rate

See Section 2.7.3.4.

2.7.4.3 Bit error rate

Bit error rate (BER) is the ratio between the amount of bit errors and the total amount of sent bits, i.e. $BER = number\ of\ errors/number\ of\ bits\ sent$ [**?**].

2.7.4.4 Code violation

Code violation (CV) is a count of far-end block errors, i.e., CRC-faults, occurring on the bearer channel [**?**].

2.7.4.5 Errored second

Errored second (ES) is a count of one seconds intervals where at least one of the following errors occurred [**?**]:

- CRC anomalies,
- Loss of Signal (LOS),
- Severely Errored Frame (SEF),
- Loss of Power (LPR).

2.7.4.6 Forward Error Correction

Forward error correction (FEC) is a method, or rather several methods, to find and correct a predetermined amount of symbol errors. The FEC parameter is a count of the number of corrected code words occurring on a channel during a period of time [**?**].

2.7.4.7 Impulse noise protection

As a performance metric, Impulse noise protection is measured as the number of sequentially symbols that the protection methods can protect against, see Section 2.5.3.

2.7.4.8 Interleave delay

Interleave delay refers to the delay which occurs because of interleaving and this delay is tunable and measured in milliseconds [**?**].

2.7.4.9 Loss of Frame

Loss of Frame (LOF) is declared after $2.5 \pm 0.5s$ of contiguous SEF defect and cleared after $10 \pm 0.5s$ of no SEF defect [**?**].

2.7.4.10 Loss of signal second

A loss of signal second (LOSS) is a count of one second intervals containing at least one LOS. A LOS occurs when there is no signal on the line for $2.5 \pm 0.5s$ [**?**].

2.7.4.11 Severely errored seconds

Severely errored seconds (SES) is a count of one second intervals where at least one of the following is true [**?**]:

- 18 or more CRC anomalies,
- 1 or more LOS defects,
- 1 or more SEF defects,
- 1 or more LPR defects.

2.7.4.12 Signal-to-noise ratio margin

Signal-to-noise ratio margin refers to how much the SNR can fluctuate before the connection is lost. The SNRM parameter ranges from -64dB to 63dB with 0.1dB steps due to amount of bits used [**?**].

2.7.4.13 Signal-to-noise ratio

Signal-to-noise ratio refers to the ratio between the received signal power and the noise power for each subcarrier. It shall be measured during diagnostic and initialisation mode and can be requested during showtime, i.e., it can be measured in real time. During showtime the SNR must be measured over at least 256 symbols [**?**].

2.7.4.14 Unavailable seconds

Unavailable seconds (UAS) is the total amount of seconds during which the interface is unavailable [?].

2.8 Related work

After having performed an extensive literature review with focus on the relation between Network Performance Metrics and QoS parameters on xDSL physical-level, and QoE on user-level, very little research related directly to this area was found. However, multiple technical articles regarding QoS to QoE mapping on OSI-layer 3 and above together with two articles regarding mapping from OSI-1 and OSI-2, respectively, to QoE where found.

In [?] experiments using a real xDSL test bed, similar to our own test bed, are performed. By using different line configurations, and varying loop lengths and noise patterns, the authors succeeds in estimating both application and network metrics based on physical metrics in an ADSL scenario. The authors conclude that the relation between network layer metrics, such as packet loss, and application layer metrics is non-linear, which indicates that information on the network layer contributes with important information to the estimation of QoE. However, in our thesis a relatively good estimation of the QoE is found without including the network layer metrics, which indicates that the relationship from physical layer directly to QoE can still be found. Furthermore, the authors of [?] uses PSNR for estimating QoE in contrast to the model VQM_VFD used in this report, which is able to handle the calibration errors that the PSNR model is so sensible to [?]. Finally, [?] lacks experiments with traffic induced on middle layers as well as experiments on VDSL2, as performed by us.

In [?], QoS at each level of the Internet protocol stack is defined, which can be directly related to the OSI-model, as depicted in Fig. 2.3. This definition is then used in [?] for mapping QoS on application-level to QoS on user-level, i.e QoE, by performing a multiple regression analysis with application-level parameters as predictor variables and the user-level parameter as the criterion value.

Furthermore, the authors in [?] and [?] use this relationship for mapping QoS on node-level to QoE in an IEEE 802.11e EDCA wireless LAN by also performing a multiple regression analysis together with a principal component analysis. In those articles, the authors use the so called Psychological Scale as a representation of the perceived quality, QoE, which

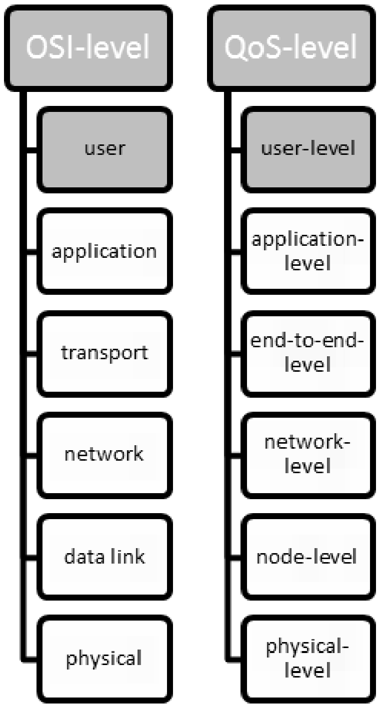

Figure 2.3: *OSI-model vs. QoS-levels* Showing the relations between the OSI-model and the QoS-levels, from physical level up to user-level.

is an improvement of the Mean Option Score, MOS, scale. In the latter article, estimations of the QoE for six audio-video content types is presented as functions of MAC-level QoS parameters. The article concludes that node-level QoS parameters can be used for estimating the QoE on user-level, but cannot reflect the processing of media on higher layers, for example buffering and media synchronisation, which of course affects the QoE. The article also concludes that more research about QoE estimation for other encoding rates and transmission rates is needed.

Chapter 3

Work and Results

3.1 Method

The aim of this master thesis was to study and analyse different known network performance metrics as well as estimation methods for QoS with a focus on QoE. To acquire sufficient knowledge of the recent research within this area a general literature review followed by a more focused review was carried out.

Afterwards, different network performance metrics together with QoS and QoE estimation methods where compiled and scrutinised in regard to their applicability on the operator end of a mobile backhaul, relying on an xDSL based access network. In addition, relations and correlations between the most suitable network performance metrics and parameter estimation methods where analysed to avoid false positives and to form a hypothesis feasible for testing with experiments, as shown in Section 3.1.2.

Furthermore, experiments where performed to find the best correlation from OSI-layer 1 to OSI-layer 7 without estimating or measuring the QoS of the traffic between those layers.

The network performance metrics and QoS parameters where measured at the physical layer of the OSI-model while the QoE parameters where estimated at the user-level. By analysing the correlations between QoS parameters and QoE parameters with analytical tools in Matlab, two models for predicting QoE by analysing network performance metrics where found, tested, evaluated and verified.

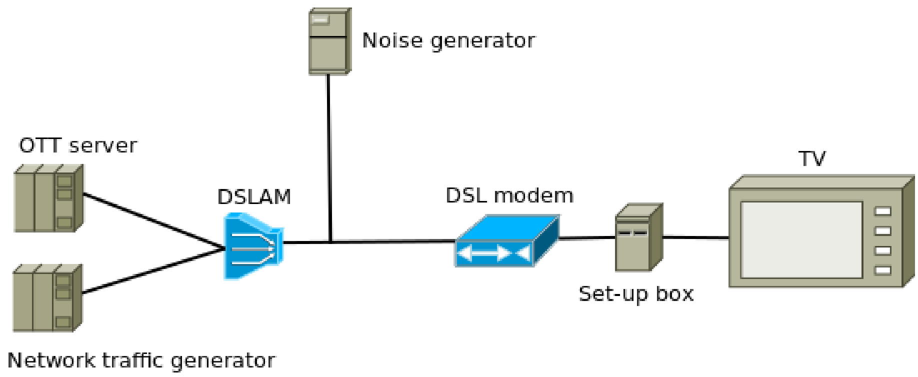

Figure 3.1: *Lab setup* Shows the components used in the lab environment.

3.1.1 Lab environment

To test how interference on a DSL based system affect the QoE of streamed video content a lab were used. Fig. 3.1 shows the lab setup used during the tests. The OTT server was used to create a UDP stream of the movie and send it to the set-top box. Instead of a traditional set-top box a computer running Linux Mint and VLC were used. The network traffic generator was used to induce other UDP data on the DSL line in an attempt to simulate a real life scenario with several subscribers concurrently using the line. The DSLAM and the DSL modem was used to transport the data over the 100 m to 1400 m long cabel using VDSL2 and a variation of settings discussed in Section 3.1.2. Attached on the DSL line is a noise generator capable of generating a variety of different signals. By inducing these signals onto the DSL line different kind of distortions can be created, the signal configuration used during the tests is described in Section 3.1.4. The set-top box was used to receive the data stream sent by the OTT server and store it to a file to be analysed later on.

Further, the lab was equipped with four different softwares:

- MatLab was used to retrieve various different parameters from the DSL line, see Section 3.1.6.
- Video Lan Client (VLC) was used to stream the movie from the OTT server to a file on the set-top box.
- avconv was used to convert to and from the YUV 4:2:2 format used in BVQM.

- Batch Video Quality Metric (BVQM) was used to analyse the quality of the streamed movie compared to the original one.

3.1.2 Test description

The QoE can be estimated from without the delay, throughput and loss for both the video and the audio stream [**?**]. Since it is not practical to distinguish between the audio and video data when observing a DSL based system the movie, including audio and video, will be seen as one data stream. How delay, throughput and loss can be measured will be discussed further in Section 3.1.6. From without this knowledge a hypothesis was formulated:

> *The quality of a streamed video in a DSL system depends on some low level parameters reflecting the delay, throughput and loss of data on the stream.*

This implies that by measuring some given parameters it should, to some extent, be possible to estimate the QoE of the video from without the video stream. Further, it should be possible to alter the QoE by disturbing the system, i.e. changing the given parameters, by inducing interference on the DSL line.
Based on this hypothesis an attempt to find correlation between DSL parameters and the QoE will be performed. To study correlation between the QoE of a video stream and DSL parameters, situated on OSI-1, several problems had to be solved.

- How can the QoE be measured objectively?
- Which OSI-1 parameters, accessible from the DSLAM, affects the QoE?
- How can the QoE be varied in a more or less controlled manner?
- Does internal DQM methods effects the QoE?
- How can the parameters measured in the DSLAM scale to measurements performed when several subscribers utilises the line?
- How does the bitrate affect the measurements?

3.1.3 Measurement of QoE

By utilising the VQM tool listed in Section 2.7.2.1 the QoE of the streamed movie can be measured objectively. Since this tool calculates the average VQM between two 15 seconds YUV 4:2:2 encoded video clips one original clip was created, *yuv_original.avi*. To produce the second clip the following steps where performed, see Fig. 3.2:

- First the clip *yuv_original.avi* was compressed using lossless h.264 to *h264_original.avi*.
- *h264_original.avi* was then sent as RTP over UDP stream across the DSL-line using the program VLC-media player.
- The stream was received by a computer on the other side of the DSL-line using the VLC-media player and stored to a new file *h264_received.avi*.
- *h264_received.avi* was then decompressed to YUV 4:2:2 again and stored in *yuv_received.avi*.
- Finally the file *yuv_original.avi* and *yuv_received.avi* was analysed using the program VQM to find the VQM score of the received video clip.

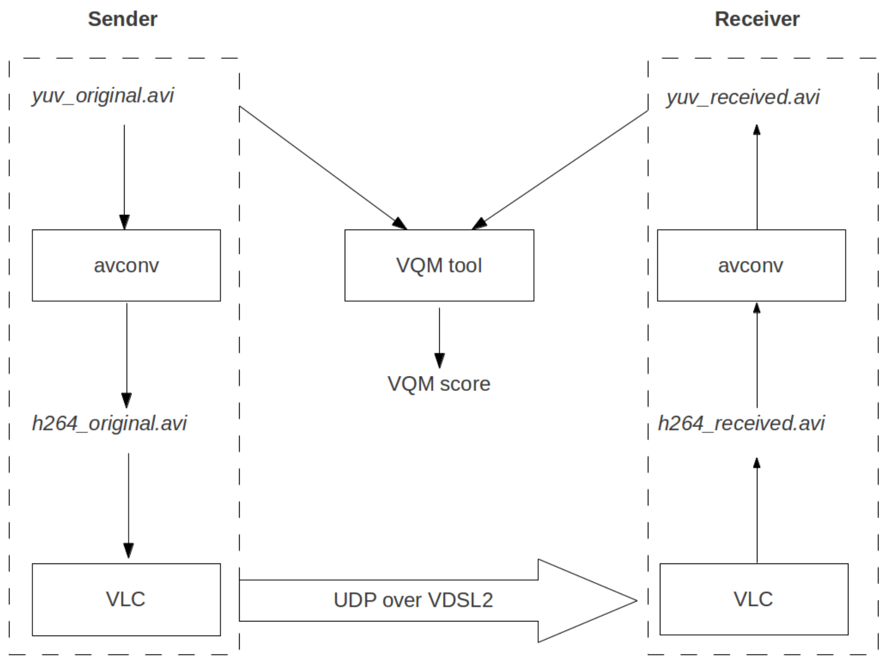

Figure 3.2: *Video streaming scenario* Shows how to obtain two video clips from which VQM analysis was performed.

3.1.4 Modify QoE

In order to alter the QoE of the video stream the quality of the loop has to be changed. This can either be done by changing configurations in the DSLAM, such as locking the bit-rate or lower the output power, or applying external noise on the loop. To more reflect reality the latter, in combination with different loop length, were chosen. The DSLAM is very good at adapting against stationary noise, e.g. using bit-swapping, and to prevent this a time varying noise had to be used. To achieve this the loop will continuously be exposed to an interference which iterates throughout the 30 MHz frequency band used by VDSL2 using 100 kHz steps each 50th ms. The interference were a sawtooth wave generated by a vector signal generator. To ensure that the noise is sufficient to alter the QoE and that it will be worse when the noise has higher voltage a series of configuration test where performed. A video clip was streamed over the VDSL2 line, with length 100 m to minimize the effect of attenuation, with thirteen different voltage of the noise. For each voltage level the clip was streamed four times to achieve a mean and a variance. In Fig. 3.3 it is shown that the VQM value of a video clip is increased when the voltage of the noise experienced by the channel is increased.

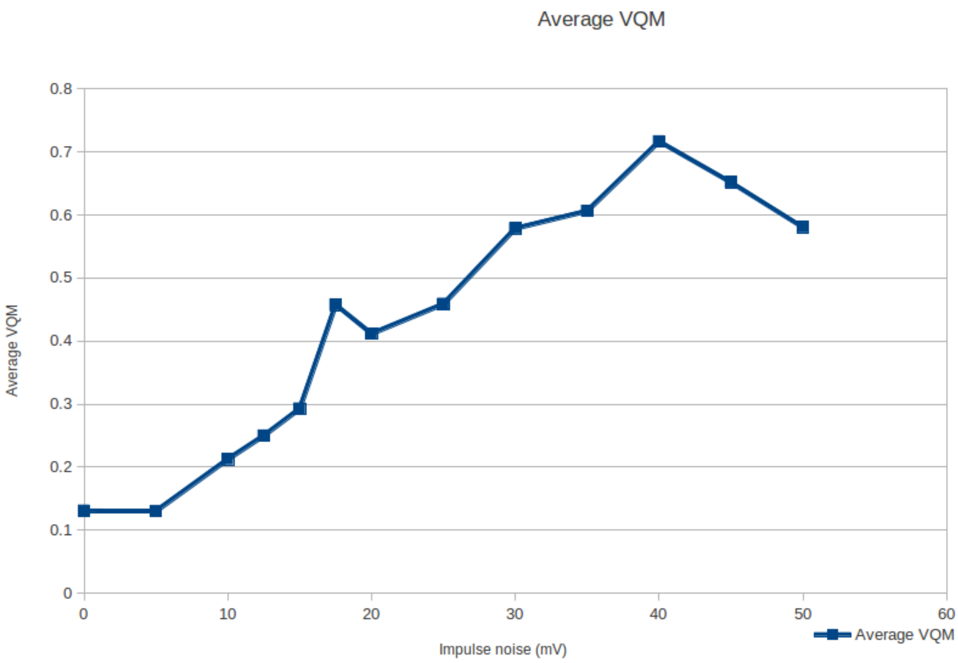

Figure 3.3: *Mean VQM plotted as a function of mV of the disturbance.* Illustrates how the mean VQM value of a video clip changes when the voltage of a disturbance increases.

Further, Fig. 3.4 shows the variance of the VQM experienced at different voltage levels of the disturbance. Even though no clear trend can be seen the variance is most present where the derivative of Fig. 3.3 has its highest value, i.e. between 10 mV and 40 mV. This shows that without interference the VQM value is always low and with high interference, at this cable length above 40 mV, the VQM value is almost always very high. Around and above 40 mV the DSLAM would reset frequently and thus the data collected at those levels are unreliable at best. Hence no conclusions will be made about the drop in VQM score for values above 40 mV.

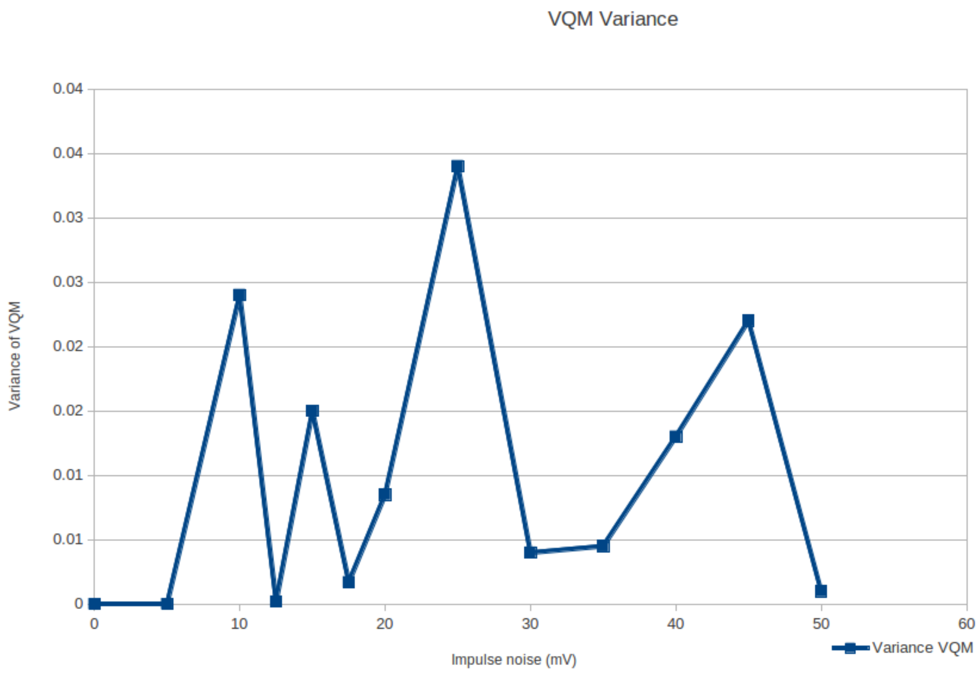

Figure 3.4: *The variance of VQM plotted as a function of mV of the disturbance.* Illustrates how the VQM variance of a video clip changes when the voltage of a disturbance increases.

3.1.5 DQM methods

As mentioned in Section 2.5 DSL utilizes several methods to ensure a good line quality. Since the QoE will be altered by injecting external noise to the line the methods preventing the effects of noise are the most important ones. Firstly, the DSLAM will bit-swap, i.e., send less bits on the tones which are experiencing noise. This is avoided by, which is mentioned in Section 3.1.4, using a noise which varies over time. Secondly, the DSLAM can be configured to use different levels of INP, as mentioned in Section 2.5.3, which may greatly increase the DSL performance in a noisy environment. To study how usage of INP effects the QoE all test cases, see Section 3.2, will be performed with and without INP. In the case where INP will be used the usage of both reed-solomon code and interleaving are implied.

3.1.6 Parameters affecting QoE

To estimate QoE from without parameters accessible in a DSLAM it is important to find how said parameters correlate with the QoE. Further, it is important to find which parameters that potentially correlates with the

QoE. As shown in Section 2.8, when estimating QoE of a video stream from without different network performance metrics, delay (including jitter) and packet loss are the parameters with the greatest impact. Since the purpose is to estimate the QoE in real-time, only parameters that changes values in real-time and fairly often, i.e., atleast each 15 seconds, are relevant. The following parameters where chosen, see Section 2.7 for a description of the parameters:

- Errored Second (ES)
- Forward Error Correction (FEC)
- Severely Errored Second (SES)
- Code Violation (CV)
- Signal-to-Noise Ratio Margin (SNRM)
- Loss of Frame Second (LOFS)
- Forward Error Correction Second (FECS)

The common denominator amongst these parameters is that all of them, in some sense, measures fault occurrences on the DSL line. Which is directly or indirectly connected to the packet loss mentioned above. Since DSL system configured to use INP and interleaving are virtually free of jitter, see Section 2.5.3, jitter and delay will be excluded. To cope with the fact that BVQM only can find the average VQM value of clips over a time interval of 15 seconds the parameters will be measured once a second for the duration of the video clip. Then every parameter will be summarised as the average or sum, depending on the parameter, of the sampled values.

3.1.7 Scale parameters

A DSL line is always sending data at max capacity and if no payload is sent the line will be filled with idle cells, see Section 2.4. This idle cells will invoke and increase the error counters mentioned in Section 3.1.6 which means that the parameters will have the same value weather any data is being sent or not. This does indeed indicate that the parameters will be valid when other traffic is on the line, however this will be further investigated in the traffic load tests described below. The traffic load tests are designed to verify weather the relations between the parameters chosen and the calculated QoE are the same when there is other traffic present on the line. By using the program ngen the line will be exposed to low traffic, traffic at half the capacity and traffic above the capacity of the line.

3.1.8 How does bitrate affect measurements?

The fact that error parameters are measured on idle cells proposes a new challenge; how can we know if an error occur on an idle cell or not? For example if we send a movie at 10 Mbps over a line with max capacity 60 Mbps only 16% of the traffic will be the movie, i.e. 84% of the errors occurring on the line will not affect the quality at all. And the tricky part about this problem is that it varies from different scenarios, i.e. if a movie of higher quality was streamed or if the lines capacity was lowered this ratio would be changed.

3.1.9 Test protocol

To cover all the cases discussed above a test suite was created. In the sections below the combination of noise strength and loop length used in each test is presented. The load tests consist of a combination of noise strength and line traffic and the loop length is constant at 700 m.

3.1.9.1 Main test without INP

In Table 3.1 the different lab environment combinations used during each test case are shown.

Table 3.1: *Main tests without INP shows the test set up for the main tests without the usage of INP.*

	Loop Length (m)		
Noise strength (mV)	100	700	1400
0	Test A1	Test A7	Test A13
0.05	-	-	Test A19
0.1	-	-	Test A20
0.25	-	-	Test A21
0.5	-	Test A10	Test A14
1	-	Test A11	Test A22
2.5	-	Test A15	-
5	-	Test A16	-
7.5	-	Test A17	-
10	Test A2	Test A8	-
15	Test A12	Test A18	-
20	Test A3	Test A9	-
30	Test A4	-	-
40	Test A5	-	-
50	Test A6	-	-

3.1.9.2 Main test with INP

In Table 3.2 the different lab environment combinations used during each test case are shown.

Table 3.2: *Main tests with INP* Shows the test set up for the main tests with the usage of INP.

	Loop Length (m)		
Noise strength (mV)	100	700	1400
0	Test B1	Test B7	Test B16
0.05	-	-	Test B17
0.1	-	-	Test B18
0.25	-	-	Test B19
0.5	-	-	Test B20
1	-	Test B8	Test B21
2.5	-	Test B9	Test B23
5	-	Test B10	Test B24
7.5	-	Test B11	Test B25
10	Test B2	Test B12	Test B26
15	Test B3	Test B13	Test B27
20	Test B4	Test B14	Test B28
30	Test B5	Test B15	-
40	Test B6	Test B29	-
50	Test B22	Test B30	-
60	Test B34	Test B31	-
70	Test B35	Test B32	-
80	Test B36	Test B33	-
90	Test B37	-	-
100	Test B38	-	-
110	Test B39	-	-
120	Test B40	-	-
130	Test B41	-	-
140	Test B42	-	-
150	Test B43	-	-
160	Test B44	-	-
170	Test B45	-	-
180	Test B46	-	-
190	Test B47	-	-
200	Test B48	-	-
210	Test B49	-	-
220	Test B50	-	-
230	Test B51	-	-
240	Test B52	-	-
250	Test B53	-	-

3.1.9.3 Verification test without INP

In Table 3.3 the different lab environment combinations used during each test case are shown.

Table 3.3: *Verification tests without INP* Shows the test set up for the verification tests without the usage of INP.

	Loop Length (m)		
Noise strength (mV)	100	700	1400
0.1	-	-	Test C10
0.25	-	-	Test C11
0.5	-	-	Test C12
1	-	-	Test C13
2.5	-	Test C5	-
5	-	Test C6	-
7.5	-	Test C7	-
10	-	Test C8	-
15	Test C1	Test C9	-
20	Test C2	-	-
30	Test C3	-	-
40	Test C4	-	-

3.1.9.4 Verification test with INP

In Table 3.4 the different lab environment combinations used during each test case are shown.

Table 3.4: *Verification tests with INP* Shows the test set up for the verification tests with the usage of INP.

	Loop Length (m)		
Noise strength (mV)	100	700	1400
5	-	-	Test D11
7.5	-	-	Test D12
10	-	-	Test D13
15	-	-	Test D14
20	-	-	Test D15
40	-	Test D6	-
50	-	Test D7	-
60	-	Test D8	-
70	-	Test D9	-
80	-	Test D10	-
160	Test D1	-	-
170	Test D2	-	-
180	Test D3	-	-
190	Test D4	-	-
200	Test D5	-	-

3.1.9.5 Load test without INP

In Table 3.5 the different lab environment combinations used during each test case are shown. Note that since the cable length was constant at 700 m during these tests whilst the traffic load over the DSL connection was varied, each test case is a combination of noise strength (mV) and traffic load (Mbps).

Table 3.5: *Load tests without INP* Shows the test set up for the load tests without the usage of INP.

	Bit rate (Mbps)		
Noise strength (mV)	30	60	10
0	Test E1	Test E9	Test E12
2.5	Test E2	Test E10	Test E13
5	Test E3	Test E11	Test E14
7.5	Test E4	-	Test E15
10	Test E5	-	Test E16
15	Test E6	-	Test E17
20	Test E7	-	Test E18
30	Test E8	-	Test E19

3.1.9.6 Load test with INP

In Table 3.6 the different lab environment combinations used during each test case are shown. Note that since the cable length was constant at 700 m during these tests whilst the traffic load over the DSL connection was varied, each test case is a combination of noise strength (mV) and traffic load (Mbps).

Table 3.6: *Load tests with INP* Shows the test set up for the load tests with the usage of INP.

	Bit rate (Mbps)		
Noise strength (mV)	30	60	10
0	Test F1	Test F9	Test F12
7.5	Test F2	Test F10	Test F13
10	Test F3	Test F11	Test F14
15	Test F4	-	Test F15
20	Test F5	-	Test F16
30	Test F6	-	Test F17
40	Test F7	-	Test F18
50	Test F8	-	Test F19

3.1.10 Test execution

When the test suite had been compiled, see Section 3.1.9, it was time to perform the tests. The test execution can be divided into three phases: the set-up phase, the streaming-phase and the analysis-phase. In the set-up phase the lab environment was set up according to the test case to be tested, i.e., the interference level, the cable length and in some cases network traffic were set. In the streaming-phase the actual content was streamed over the DSL line and the received data was stored in a file. This was repeated for all test cases in the test suite. Afterwards, when all data had been collected, the analysis-phase begun. During the analysis-phase all collected data was simply analysed using the VQM-tool. The result of the tests are presented in Section 3.2.

3.2 Test Results

In this section the results of the tests mentioned in Section 3.1.2 are presented. The meaning of the results will be further discussed in Chapter 4. As mentioned above all test groups are divided into two test batches, with and without INP, and will be presented individually. In the figures presented the data collected from each respective test batch, described in Section 3.1.9, are shown. The mean squared error (MSE) between the data estimated by the model and the real values will also be presented to further provide insight in the models.

3.2.1 Main tests

To find the correlation between VQM and the selected parameters, see Section 3.1.6, the data from the tests described in Table 3.1 and Table 3.2 were collected. For each test case the corresponding VQM values were calculated. The collected and calculated data were separated in to two groups, the data found without using INP and the data found using INP. In the sections below the data will be presented in regard to those groups. By performing a regression analysis in Matlab on the data in each group separately several potential models where found. The quality of the models where determined from two criteria:

- goodness of fit in form of MSE and
- the adjusted R-squared value.

The two best models will be further analysed below. The models chosen will henceforth be referred to as model A, the model not using INP, and model B, the model using INP.

Model A

$$VQM = 6.8107*10^{-2} + 4.8039*10^{-1}*ES - 7.2516*10^{-1}*SES + 8.1623*10^{-4}*CV + 1.6442*10^{-1}*(ES*SES) - 9.7086*10^{-2}*ES^2 - 4.238*10^{-2}*SES^2 - 1.5258*10^{-7}*CV^2$$

Model B

$$VQM = 6.4329*10^{-2} + 9.851*10^{-2}*ES - 7.7241*10^{-9}*FEC - 2.022*10^{-2}*SES - 2.8215*10^{-6}*CV + 6.0093*10^{-6}*SNRM - 6.7713*10^{-5}*FECS$$

3.2.1.1 Without INP

Fig. 3.5 illustrates how well model A follows the real VQM values.

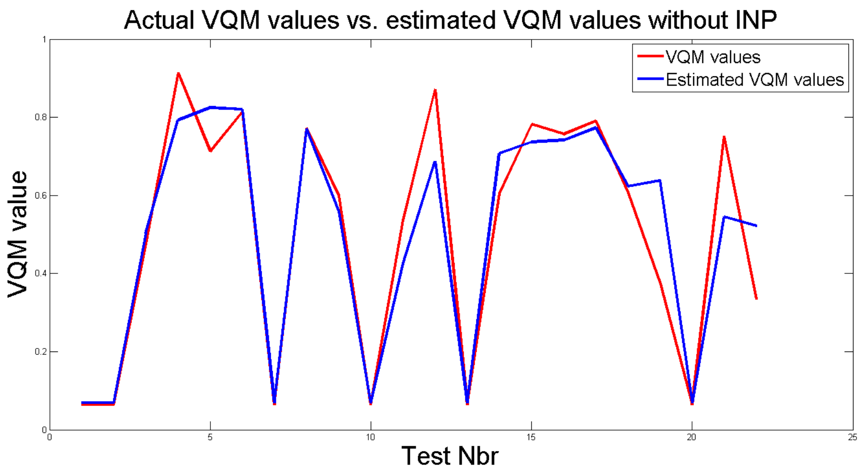

Figure 3.5: *Main tests without INP* Shows how well model A estimates VQM without the presence of INP.

3.2.1.2 With INP

Fig. 3.6 illustrates to what degree model B follows the real VQM values.

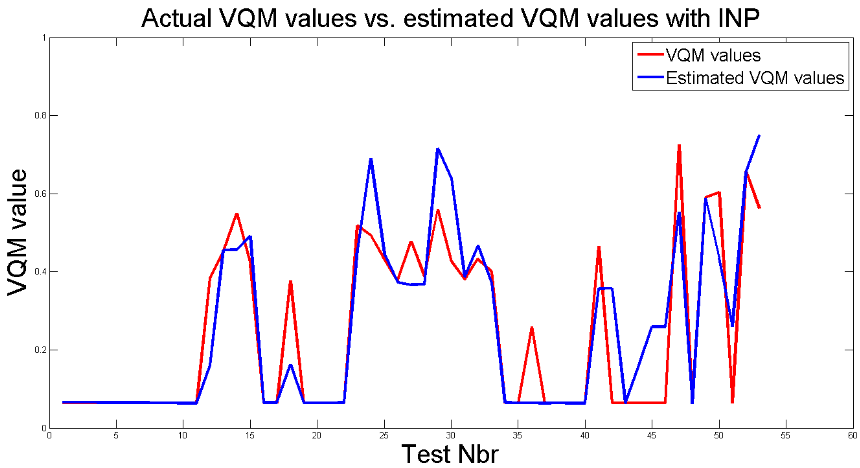

Figure 3.6: *Main tests with INP* Shows how well model B estimates VQM with the presence of INP.

3.2.2 Verification tests

During the main tests two models were created and evaluated. However, the models were only evaluated against the data they where created from. To avoid false correlation it is important to investigate how well the models behave when presented with new data.

3.2.2.1 Verification of model A

To verify model A, found in Section 3.2.1.1, new data were collected. By using matlab it was found that the mean squared error between the estimated VQM of the new data and the real VQM of the new data was 0.12. In Fig. 3.7 the estimates of new data created from without model A versus the real VQM values of that data is shown.

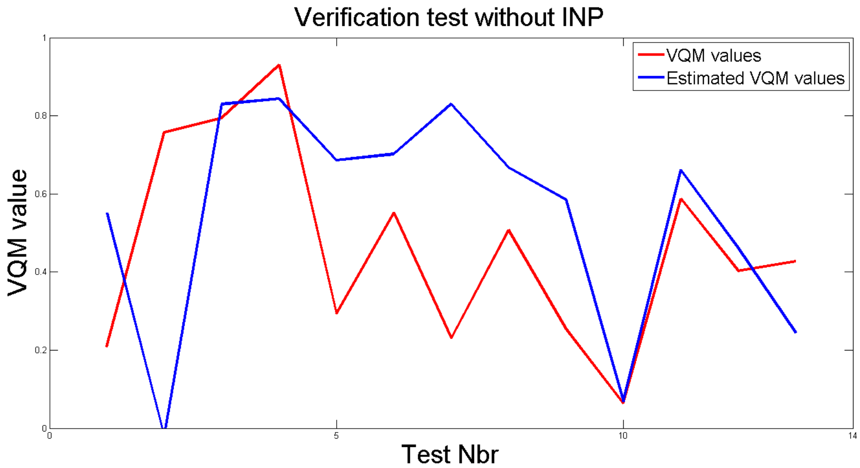

Figure 3.7: *Verification tests without INP* Shows how well model A estimates VQM of new data without the presence of INP.

3.2.2.2 Verification of model B

To verify model B, found in Section 3.2.1.2, new data were collected. By using matlab it was found that the mean squared error between the estimated VQM of the new data and the real VQM of the new data was 0.005. In Fig. 3.8 the estimates of new data created from without model B versus the real VQM values of that data is shown.

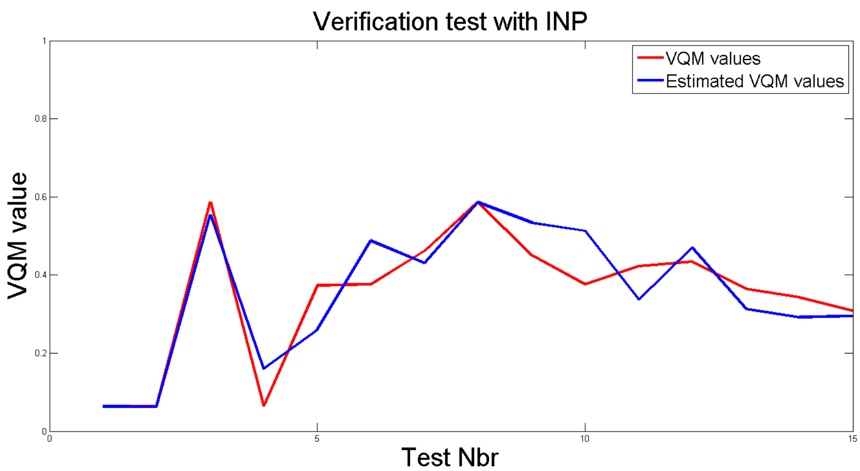

Figure 3.8: *Verification tests with INP* Shows how well model B estimates VQM of new data with the presence of INP.

3.2.3 Load tests

An important aspect of both model A and model B is its ability to perform on a busy channel, i.e., a channel utilised by several concurrent subscribers. Note that the line is flooded during test 9, 10 and 11 during both load tests scenarios.

3.2.3.1 Load test on model A

In Fig. 3.9 the performance of model A on a busy line is shown. Provided that the line is not flooded the MSE was calculated to 0.07.

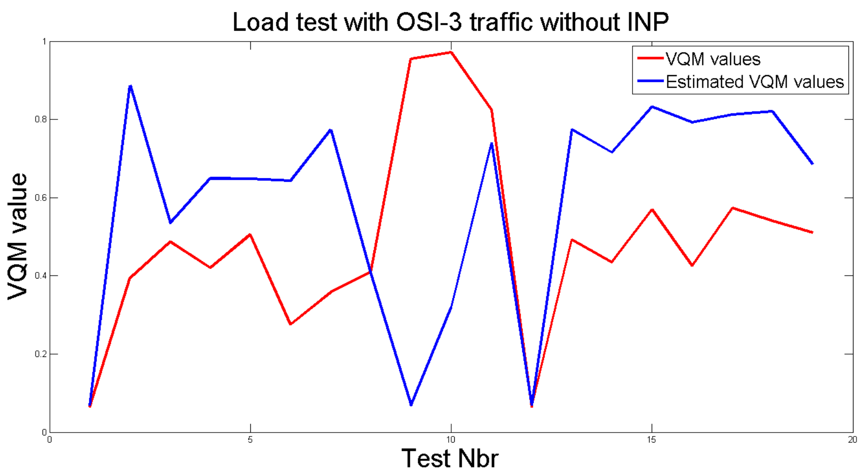

Figure 3.9: *Load tests without INP* Shows how well model A estimates VQM of new data without the presence of INP on busy channel.

3.2.3.2 Load test on model B

In Fig. 3.10 the performance of model B on a busy line is shown. Provided that the line is not flooded the MSE was calculated to 0.005.

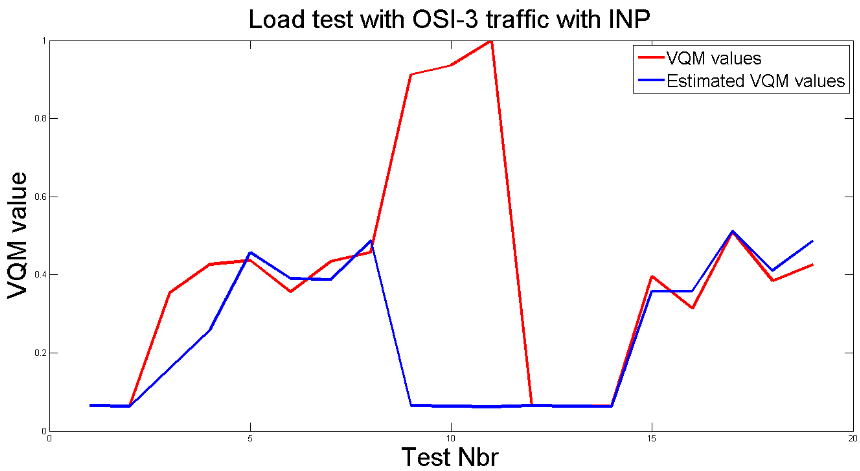

Figure 3.10: *Load tests with INP* Shows how well model B estimates VQM of new data with the presence of INP on a busy channel.

Chapter 4

Model Discussion

As mentioned in Section 3.2.1 two different models where created from without the test data, model A and model B. Considering the Mean Squared Error (MSE) of model A and model B during the verification tests it is clear that model B is more accurate; $MSE = 0.17$ for model A versus $MSE = 0.005$ for model B. This is likely due to either of two different factors:

- the usage of INP or
- the fact that model B is based on a wider selection of test data.

Considering Fig. 3.7 and Fig. 3.8 it is clear that model A has its weakest point at estimating VQM values between 0.2 and 0.5, whilst model B is even for all VQM values. By using INP, as in model B, many of the small random errors introduced in this interval can be adjusted, which could explain why model B is superior to model A. Model B is based on 53 tests whilst model A is based on 22 tests and on top of that model B uses two more parameters (FEC and FECS) than model A. This could also explain why model B is superior to model A.

As shown in Fig. 3.9 and Fig. 3.10 both model A and model B performs just as well, with just a slight and negligible improvement in MSE for model A, on a busy line. However, in the figures it is clear that when the line is flooded, during test 9, 10 and 11, neither model can estimate the VQM score. As mentioned in Section 3.1.7, DSL utilises idle cells and the chosen parameters are based on all cells sent, including the idle cells. As a consequence of this only errors occurring within the scope of the DSL channel, i.e. errors on the actual cells, will be detected. In the case of test 9, 10 and 11 the poor video quality is caused by congestion in the DSLAM, which results in data being discarded outside of the DSL line. Since no errors actually occurs on the DSL line but rather in the DSLAM, both models estimates the VQM score to a low value when it actually is very high.

Due to cells being sent continuously, either as data cells or idle cells, both model A and B can actually determine the potential VQM score of an idle DSL line. This concludes that the models are not applicable to a given video stream per say but instead presents diagnostics of the DSL line.

> *Given that a video in perfect condition arrives at the DSLAM and that the line is not flooded then those models can predict the VQM score of the video upon arrival at the CPE modem.*

This can be used in the future to determine which channel a roaming subscriber should use when streaming video, as discussed in Section 5.2.

The models are based on tests where video was streamed with RTP over UDP and thus only best-effort channels have been tested. A consequence of this is that other, non best-effort channels such as TCP for example, will probably not work. Further studies relating to the usage of connection oriented protocols should be conducted.

The entire test suite was performed under certain perfect lab circumstances and the distortion was not a realistic noise and thus the model might not work outside the laboratory with real impulse noises. The constants used in model A and B, or possibly the entire formulas, might have to be adapted to each new scenario and DSL setup. Further testing on this matter is needed.

As mentioned in Section 3.1.8 the quota of payload, i.e. movie bit rate to total line bit rate is an important factor to consider. All tests have been performed using a single and rather lightweight video clip. This results in that this quota has been kept rather low and how this model will work with other clips higher and/or lower quality is uncertain. This is also a subject for further studies.

The models where created from without measurements of ES, FEC, SES, CV, SNRM and FECS. An analysis of the internal correlation amongst the parameters are shown in Table 4.1. From the correlation matrix it is quite clear that there is heavy internal correlation between parameters and these correlations can to a degree be explained by studying how each of the parameters are connected to one and another. For example, the CV and ES has a high positive correlation because a CV invokes an ES, as described in Section 2.7.4.5. However, heavy internal correlation can mean that some underlying parameter which correlates with the other parameters may exist. To improve the accuracy of the model, this should be investigated further.

Table 4.1: *Correlation matrix* Shows the internal correlation amongst the chosen parameters.

	ES	FEC	SES	CV	SNRM	FECS
ES	1	0.5995	0.7017	0.6599	-0.2767	0.462
FEC	0.5995	1	0.156	0.1599	0.3023	0.7525
SES	0.7017	0.156	1	0.9246	-0.4744	0.02532
CV	0.6599	0.1599	0.9246	1	-0.444	0.04987
SNRM	-0.2767	0.3023	-0.4744	-0.444	1	0.3874
FECS	0.462	0.7525	0.02532	0.04987	0.3874	1

Despite the drawbacks of the models mentioned here it is still concluded that there indeed is some truth in the hypothesis formulated in Section 3.1.2. Given further studies and refinements of the models those may prove to be powerful tools to ensure a satisfactory QoE of video sent over copper based DSL systems.

Chapter 5

Conclusions

5.1 Value of the results

The aim of this master thesis was to study, compile and analyse different known network performance metrics as well as estimation methods for QoS with a focus on QoE. By investigating their applicability on copper-based mobile backhauls, information about where in the mobile network to apply those methods, could be deduced.

By conducting a massive literature review within the scope of this topic, we found that little or none research had been done about the relation between QoS metrics at the physical layer and perceived QoE and therefore decided to focus on that. From this, a hypothesis was formed, as stated in Section 3.1.2. On the basis of this hypothesis, a multitude of experiments where performed to find correlation from OSI-layer 1 to OSI-layer 7. The experiments resulted in two estimation models for estimating QoE at user-level based on measured QoS metrics at physical-level. The models are to no extent claimed to be a universal solution to the problem of ensuring a good quality for the end-user, but should instead be seen as indications on the non-linear, but existing, relation between the physical-layer and the perceived QoE at user-level. These finding can hopefully contribute to the future development of QoS and QoE methods for FMC networks.

5.2 Contribution to COMBO

As discussed in the FP7 project COMBO [?][?], future mobile backhaul technologies will partly rely on fixed (non-3GPP) access networks, with some consisting of DSL technology. With fixed broadband networks in Europe currently dominated by different DSL technologies and Fiber To

The Curb (FTTCurb) with VDSL2 as the State of the Art, the expected FMC together with a predicted exponential growth of data traffic will put new challenges on the DSL technology regarding QoS and QoE.

By investigating and comprehending the interrelations between different parameters on lower levels of the OSI-model, information about the QoS and QoE on higher layers can be deduced. With our models the quality of the loop, in the context of streaming video, in a copper-based backhaul can be diagnosed.

5.2.1 COMBO scenario 1

A typical use case for our models would be the scenario where local stations, e.g femtocells, offloads macro mobile stations and routes best-effort data through the local fixed access backhauls. Here, an intelligent FMC network would be able to decide on which of the available backhauls to use based on, among other criteria, their current potential for delivering OTT content, i.e. based on their current estimated VQM.

By continuously monitoring OSI-1 parameters for the different available copper-based mobile backhauls in the CO/NG-POP or in the small cells themselves, the FMC network would hold information about the different backhauls' potential QoS/QoE even if they where unused at the moment. This would contribute to the optimal and seamless QoE for the end-user that the COMBO project strives to achieve and also make it possible to negotiate SLAs promising a specific level of QoE for OTT content over specific parts of the FMC network [**?**][**?**].

5.2.2 COMBO scenario 2

Another scenario that could make this type of performance monitoring even more relevant is the current trend to move video and content distribution servers from large data-centres placed in the core area of the network into the metro area of the network. By bringing the source of content closer to the subscribers and thus closer to their local backhauls, the source of errors on higher OSI-layers induced by the core and aggregation networks would be minimised, thus improving the accuracy of our model [**?**][**?**].

5.3 Future Work

Since research within the area of estimating QoE from low level QoS metrics is so rare, quite a lot of work needs to be done on this topic in the future.

More research on how changes in lower layer parameters affect network layer parameters, which in turn affect application layer parameters, is definitely needed.

With more time, resources and experience, our test result data could probably be segmented in the same way as is done in [**?**] to achieve an even better accuracy. Moreover, experiments should be performed for other line configurations and setups than those performed here, and multiple protection mechanisms together with different SNR and power thresholds should be tested to disqualify possible false relations and positives. Furthermore, different video qualities should be tested, both lower and higher than the half SDTV tested by us, to adapt the models to a greater set of varying media quality. Other streaming mechanisms than RTP over UDP should also be tested to investigate the impact of connection oriented transport protocols, as is also discussed in Section 3.2. Finally, future tests should take the amount of idle cells in the xDSL transmission into account to calculate the ratio between the idle data and useful data. In this way a higher level of accuracy could probably be achieved.

References

[1] 7th Framework Programme, “Annex i - description of work - combo,” 2012.

[2] 7th Framework Programme, “Large-scale integrating project (ip) proposal - combo,” 2012.

[3] “Tr-197: Dqs: Dsl quality management techniques and nomenclature,” 2012.

[4] E. AB, “Ecn web interface user guide - eda 1200,” 2009.

[5] ITU-T, “E.800: Definitions of terms related to quality of service,” 2008.

[6] ITU-T, “P.800: Methods for subjective determination of transmission quality,” 1996.

[7] F. Kuipers, R. Kooij, D. D. Vleeschauwer, and K. Brunnström, “Techniques for measuring quality of experience,” 2010.

[8] T. Yamada, Y. Miyamoto, and M. Serizawa, “No-reference video quality estimation based on error-concealment effectiveness,” 2007.

[9] “The top 5 video streaming websites,” 2012. [Online; accessed 2013-12-17] URL: http://www.techsupportalert.com/top-5-video-streaming-websites.htm.

[10] “Youtube press statistics,” 2012. [Online; accessed 2013-12-17] URL: http://www.youtube.com/t/press_statistics/.

[11] “700 million youtube subscribers by 2016?,” 2010. [Online; accessed 2013-12-17] URL: http://www.trefis.com/stock/goog/articles/13489/700-million-monthly-youtube-visitors-by-2016/2010-03-19.

[12] "The zetabyte era," 2012. [Online; accessed 2013-12-17] URL: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/VNI_Hyperconnectivity_WP.html.

[13] "The growth of video streaming is on the rise," 2012. [Online; accessed 2013-12-17] URL: http://www.videoindustrynews.com/2012/03/14/the-growth-of-video-streaming-is-on-the-rise/.

[14] C. Franklin, "How dsl works," 2011. [Online; accessed 2013-12-17] URL: http://computer.howstuffworks.com/dsl1.htm.

[15] P.-E. Eriksson and B. Odenhammar, "Vdsl2: Next important broadband technology," 2006. [Online; accessed 2013-12-17] URL: http://www.ericsson.com/ericsson/corpinfo/publications/review/2006_01/files/vdsl2.pdf.

[16] "Digital subscriber line (dsl)," 2008. URL: http://www.scholarpedia.org/article/Digital_subscriber_line_(DSL).

[17] "G.998.4: Improved impulse noise protection for dsl transceivers," 2010.

[18] S. Wolf and M. H. Pinson, "Video quality model for variable frame delay (vqm_vfd)," 2011.

[19] G. Gonçalves, R. Dantas, A. Palhares, J. Kelner, J. Fidalgo, D. Sadok, H. A. Berg, and D. Cederholm, "Estimating video quality over adsl2+ under impulsive line disturbance," 2009.

[20] "Rfc 6374: Packet loss and delay measurement for mpls networks," 2011.

[21] "Rfc 5481: Packet delay variation applicability statement," 2009.

[22] "Ber bit error rate tutorial and definition," 2013. [Online; accessed 2013-12-17] URL: http://www.radio-electronics.com/info/rf-technology-design/ber/bit-error-rate-tutorial-definition.php.

[23] "G.997.1: Physical layer management for digital subscriber line transceivers," 2012.

[24] "Tr-176: Adsl2plus configuration guidelines for iptv," 2008.

[25] "G.992.3: Asymmetric digital subscriber line transceivers 2 (adsl2)," 2010.

[26] S. Tasaka and Y. Ishibashi, "Mutually compensatory property of multimedia qos," 2002.

[27] S. Tasaka and Y. Ito, "Psychometric analysis of the mutually compensatory property of multimedia qos," 2003.

[28] Y. Ito and S. Tasaka, "Feasibility of qos control based on qos mapping in audio–video transmission over ieee 802.11 wireless lans," 2006.

[29] T. Suzuki, T. Kutsuna, and S. Tasaka, "Qoe estimation from mac-level qos in audio-video transmission with ieee 802.lle edca," 2008.

[30] T. Suzuki, T. Kutsuna, and S. Tasaka, "Qoe estimation from mac-level 008 in audio-video transmission with ieee soz.lle edca," 2008.

LUND UNIVERSITY

http://www.eit.lth.se