

**User-Centric Video Streaming Quality of Service and Experience
Evaluation over Next Generation Mobile Networks.**

F.L Lugayizi

17058198

**Thesis submitted for the degree of Doctor of Philosophy in Computer
Science at the Mafikeng Campus of the North-West University**

Supervisor: Prof. M.B Esiefarienrhe
Co-Supervisor: Dr. N Gasela

May 2016

DEDICATION

I dedicate this endeavor firstly to the almighty GOD – the one who gave me the gifts of life, good health, strength, wisdom and courage; above all, the gifts of the Holy Spirit to work through the various challenges this work brought with it. I shall forever remain grateful for the Lord’s intervention whenever my morale was down, I will forever trust in you Lord.

Secondly, to my beloved parents; Mr. and Mrs. JMC Kiwanuka for the parental, financial and moral support they accorded me up to now and more especially during this work, may the almighty God continue to reward you abundantly with the gift of life to make you enjoy the fruits of your investments.

In a very special and heartfelt way, with so much love I dedicate this work to my lovely Son Jairus-John Francis Bayiga Lugayizi and the mother; my wife Damalie K. Lugayizi, thank you for the belief, patience, encouragement and moral support you sacrificed to give to me during this work despite the challenging nine months you went through, I shall forever be grateful.

Lastly to my wonderful siblings and their families; Denis (“the Mutebis”), Lawrence (“the Kigonyas”), Winnie (“the Winnies”), Felix (“the Ssekatawas”), Philip and Joseph-thank you for showing the family love and emotional support for all these years, I dedicate this achievement to you all.

ACKNOWLEDGMENTS

I thank God once again for having brought me through my education stages until this stage of completing my work successfully, thus I go ahead to salute all the people the Lord has used to help me through this time.

Firstly, I express a lot of gratitude to the people who have been part of making this whole journey a success; people God chose to do that; my Supervisors Prof. Michael Esiefarienrhe and Dr. Naison Gasela for having guided and mentored me so well throughout the stages of this work; may God reward you abundantly and also give us more opportunities to work together in future.

Special thanks go out to the Acting Vice Rector of Research and planning; Prof. Eno Ebenso, the Acting Dean-Faculty of Agriculture, Science and Technology; Prof H. Drummond and the Acting Director School of Mathematical and Physical sciences; Prof. Ashmore Mawire for having portrayed professional management skills as far as running the Faculty was concerned with regard to the various challenges the faculty experienced during this work; this professionalism channeled well through to our work hence the successful completion. God Bless you all.

In another special way, I would like to thank my big brother Felix who has been a great role model throughout my University life experience, without your support, I do not think I would be where I am right now.

Lastly, I thank the North West University Post graduate Bursary and emerging researcher's fund for the financial support that was given to me to be able to complete this work, thank you very much.

DECLARATION

I _____ solemnly declare that most of this work is original from me and instances where another author's or writer's work was used, it was fully acknowledged both within text and in the references or bibliography at the end of the work. I further confirm that this work hasn't been submitted in any form for another degree or diploma for that matter to any institution of learning, and even after, there is no intention for re-submitting this work to any other institution for the purposes of obtaining another degree or diploma.

F.L Lugayizi
17058198
May 2016

Prof. M.B Esiefarienrhe
Supervisor
May 2016

Dr. N. Gasela
Co-Supervisor:
May 2016

ABSTRACT

Quality of Experience (QoE) is a prototype of human feelings, wishes, impressions and motives. Quality of Service (QoS) on the other hand is a technology-based metric that is employed to gauge how a network performs at any given time. Latterly, QoE in Next Generation Mobile Networks (NGMNs) - mostly referred to as 4G Long Term Evolution (LTE) networks have continued to attract interest from network researchers, developers and designers. This is mainly because QoE on its own comes with a degree of relationship to QoS or network conditions, yet to end users QoE matters much more than QoS. QoS is basically about network metrics such as jitter, pack delay and loss. Much as these metrics are critical too, they are only a small part of satisfying users, this is because QoE is more related to the perception of a service by a mobile user.

The main challenge that remains is; mobile users always expect their mobile devices to have full connectivity and high network performance, if any of this is absent, providers of this service do not have sufficient time to wait for these quality complaints from the users to come through; instead such users will simply change service providers in the hope of finding a better service on another network. This confirms that excellent network metrics can easily mean nothing especially if users decide to find other service providers. This occurrence poses two challenges; first, mobile users will continue to look for better quality video content streaming without breakouts elsewhere, secondly the networks they may keep moving to are bound to experience a knock-on effect as far as network resource utilization is concerned, this is mainly because of a sudden increase in the number of users accessing the same service at the time. In this same period, video streaming also continues to have a significant influence on network traffic especially in NGMNs. The major force behind this influence has been satisfaction of user experience as far as service delivery is concerned.

Therefore, it is important that service providers hold a means of continuity in the measurement of the QoE since QoE may not only be judged on network speed improvement such as deploying advanced technologies. The notion of bridging this gap between QoS and QoE in NGMNs has indeed led to significant and successful works being published by network researchers. However, much of these systems and mechanisms suggested in these works mainly focus on presenting specific technologies. In this thesis, we introduce an independent mechanism that is compatible enough with other technologies. The mechanism has the ability

to monitor network QoS and at the same time evaluate trends in QoE as far as user perception of the service is concerned. We used both subjective lab environment simulations and a simulated environment using the NETEM simulator. To generate the video streams we used the VLC media player, a router and a switch hosted the Network Address Translator (NAT) on the subnet and a tool known as Distributed Internet Traffic Generator (DITG) was used to emulate TCP bulk traffic. To emulate packet loss, Netem was our tool of choice, which was integrated into the Linux iproute2 toolset that we also used for the queuing disciplines required to implement and realize the required behaviour. After this setup, we then ran several simulations in order to study the performance of our mechanism.

According to the results, the mechanism performed as expected and each of the simulations resulted in improved quality for the relevant users. It should be noted that, customer differentiation presented the most expected results, which was an important aspect to Internet Service Providers (ISPs). The simulations also showed that the perceived quality of the users can be improved without sacrificing the quality of streams belonging to ordinary users. The general conclusion from the implementation of these experiments indicated that it is indeed possible to monitor QoE of the users so as to retain them, this was entirely made possible because the mechanism that was implemented can equally keep track and evaluate network QoS in terms of video quality performance, resource utilization and optimization. Thus addressing the major challenges that QoS and QoE pose to NGMNs as the number of mobile video streaming users increase for different reasons.

Keywords: *Quality of Service, Quality of Experience, Administrative System, Performance Evaluation, Mean Opinion Score.*

TABLE OF CONTENTS

DEDICATION.....i

ACKNOWLEDGEMENT..... ii

DECLARATION..... iii

ABSTRACT iv

TABLE OF CONTENTS v

TABLE OF FIGURES.....vii

LIST OF TABLES..... xiv

LIST OF ACRONYMS and ABBREVIATIONS..... xv

CHAPTER 1 1

1.0 INTRODUCTION 1

1.1 Background and Motivation 1

1.2 Problem Statement 3

1.3 Research Goal 5

1.4 Research questions 5

1.5 Research Objectives.....6

1.6 Research Limitations 6

1.7 Research Contribution.....7

1.7.1 Publications..... 10

1.8 Thesis outline 10

CHAPTER 2..... 12

2.0 RELATED LITERATURE 12

2.1 Chapter Overview 12

2.2 NGMNs Theoretical Background..... 12

2.3 Challenges in Next Generation Mobile Networks 16

2.3.1 QoS..... 16

2.3.2 Mobility Management and Seamless Connectivity..... 17

2.3.3 Energy Efficiency..... 18

2.3.4	<i>Call Admission Control (CAC)</i>	18
2.3.5	<i>Resource Allocation</i>	18
2.3.6	<i>Intelligent Billing and Cost</i>	19
2.3.7	<i>Security and Fraud Management:</i>	19
2.4	<i>Video Streaming in NGMNs</i>	20
2.4.1	<i>Video Models</i>	22
2.4.2	<i>User Abandonment</i>	29
2.4.3	<i>Video Compression</i>	31
2.4.4	<i>Video Format</i>	31
2.4.5	<i>Video Downloading vs. Video Streaming</i>	33
2.4.6	<i>Types of Video Streaming</i>	35
2.4.7	<i>Video Streaming Protocols</i>	36
2.4.8	<i>Assessment of Videos</i>	36
2.5	<i>Communication Ecosystem</i>	40
2.6	<i>Quality of Service(QoS) and Quality of Experience(QoE) in NGMNs</i>	43
2.6.1	<i>Quality of Service</i>	46
2.6.2	<i>Quality of Experience (QoE)</i>	59
2.7	<i>QoE and QoS evaluation in testbed-oriented Living Lab setting</i>	65
2.8	<i>Beyond State-of-the-Art</i>	67
2.8.1	<i>5G HetNet architecture and enabling technologies.</i>	71
2.9	<i>How mobile networks improve QoS</i>	75
2.9.1	<i>QoE and the Techno Centric Approach</i>	85
2.9.2	<i>QoE and the User Centric Approach</i>	88
2.10	<i>Techno centric approach in advanced mobile networks</i>	90

2.11	Quality of Service Class Pointer	91
2.12	Allocation and Retention Priority (ARP).....	91
2.13	Guaranteed Bit Rate (GBR).....	91
2.14	Maximum Bit Rate (MBR)	92
2.15	Parameters for the Network Policy Controller.....	92
2.16	Uplink and Downlink Base Station Policies	94
2.17	Identification and Designing of QoS Networks.....	94
2.17.1	<i>Policing</i>	94
2.17.2	<i>Traffic modelling and management.</i>	95
2.17.3	<i>Administration of QoS policy and management</i>	95
2.17.4	<i>Data planning and scheduling</i>	95
2.18	Interrelating network QoS to User QoE.....	96
2.19	Quality of Service and Quality of Experience Layers.....	97
2.20	Ways of Measuring QoE	100
2.21	Objective Quality Assessment Modelling	103
2.21.1	<i>Bit-stream model</i>	106
2.21.2	<i>Parametric planning model</i>	106
2.21.3	<i>Parametric packet-layer model</i>	106
2.21.4	<i>Hybrid model</i>	106
2.21.5	<i>Media layer model</i>	107
2.22	Mapping subjective scores to objective indices	107
2.22.1	<i>Linear mapping function</i>	109
2.22.2	<i>Non-linear mapping functions</i>	109
2.23	Different approaches in interrelating QoS to QoE.....	110
2.23.1	<i>VQM-BASED Mapping Model</i>	110
2.23.2	<i>Statistical Analysis QoE Modelling</i>	112

2.23.3	<i>Machine Learning based QoE models.....</i>	113
2.23.4	<i>Crowdsourcing for subjective simulations as a QoE model.....</i>	115
2.23.5	<i>Resource Arbitration based QoE modelling</i>	116
2.23.6	<i>Consideration of environment and equipment factors in QoE modelling.</i>	117
2.23.7	<i>Quantitative and Qualitative Assessment based QoE modelling</i>	118
2.24	Summarized findings of studies that confirm links between QoS and QoE.....	118
2.25	Conclusion	122
2.26	Chapter Review.....	124
CHAPTER 3		126
3.0 RESEARCH METHODOLOGY		126
3.1	Chapter Overview	126
3.2	Introduction.....	126
3.2.1	<i>Quantitative Research Methods</i>	126
3.2.2	<i>Qualitative Research Method</i>	127
3.2.3	<i>Mixed Research Methods.....</i>	127
3.3	Methodology used in the thesis.....	127
3.3.1	<i>NGMNs State-of-the Art review</i>	128
3.3.2	<i>QoS/QoE Administration Mechanism.....</i>	128
3.4	QoE Intervention Manager.....	133
3.4.1	<i>Dynamic Network Resource Provisioning</i>	133
3.4.2	<i>Shaping</i>	134
3.4.3	<i>Access Control</i>	134
3.4.4	<i>Handover.....</i>	134
3.4.5	<i>Marking and Alerting</i>	135
3.5	Chapter Review.....	135

CHAPTER 4.....	136
4.0 IMPLEMENTATION	136
4.1 Chapter Overview	136
4.2 Prototype implementation	136
4.2.1 Policy.....	140
4.2.2 Performance Indicator	142
4.2.3 Performance Model	143
4.2.4 Decision.....	144
4.3 Test bed Evaluation setup.....	151
4.4 Chapter review	154
CHAPTER 5.....	155
5.0 RESULTS ANALYSIS AND DISCUSSION.....	155
5.1 Chapter Overview	155
5.2 Simulation 1: AP performance during congestion without the mechanism.....	155
5.2.1 Discussion	157
5.3 Simulation 2: User-Centric differentiation other than application-centric differentiation.....	159
5.3.1 Discussion	159
5.4 Simulation 3: Performance evaluation of the administration system.	166
5.4.1 Discussion	166
5.5 Simulation 4: Effects of traffic and client differentiation.	170
5.5.1 Discussion	174
5.6 Chapter Review.....	175
CHAPTER 6.....	176
6.0 CONCLUSION AND FUTURE WORK.....	176
6.1 Chapter Overview	176
6.2 Research Summary	176
6.3 Research contributions	178

6.4 Future Work..... 180

6.5 Closing Remark 181

7.0 REFERENCES 182

APPENDIX A - EMULATING WIRELESS/MOBILE NETWORK DELAYS 194

APPENDIX B – ABSTRACTS OF PUBLICATIONS 197

TABLE OF FIGURES

Figure 1-1 Conceptual view of research contributions	8
Figure 2-1: How 3GPP standards have evolved[11].....	13
Figure 2-2 LTE Wireless architecture[14]	15
Figure 2-3 A video model based on chunks[30]	25
Figure 2-4. The status of the video file buffer in FFD and the Throttling based video model[30]	28
Figure 2-5 Viewer abandonment average rate by viewing time[31]	30
Figure 2-6 Communication Ecosystem Terminology[55]	41
Figure 2-7 Relationship between various QoS and QoE approaches	47
Figure 2-8 ITU_T E.800 framework[76]	50
Figure 2-9 Mapping of user centric QoS requirements and model for user centric QoS categories based on ITU-T Rec G 1010[80]	52
Figure 2-10 End-to-end Delay Calculation	55
Figure 2-11 Quality of Experience factors to consider	60
Figure 2-12 New technological advancement towards 5G[109].....	68
Figure 2-13 HETNETS[110].....	70
Figure 2-14 challenges that 5G is going to offer to all of the users[115].	74
Figure 2-15 User satisfaction is easily affected by mobile video QoE[116]	77
Figure 2-16 Improvement in mobile video QoE across 5 different mobile networks[117]	79
Figure 2-17 Improved Quality of Experience with Smart Capacity Optimization[117]	81
Figure 2-18 Higher resolution videos create the most traffic volume on networks[117]	84
Figure 2-19 Elements that influence QoE.....	87
Figure 2-20 QoE and QoS Mapping Curve[127].....	99
Figure 2-21 User-centric ways of measuring the quality of network services.....	102
Figure 2-22. The methods of mapping QoE to QoS	108
Figure 3-1 Customer experience administration mechanism (CEAM).....	130
Figure 4-1: A conceptual view of the access point software.	139
Figure 4-2: The queuing policy.....	146
Figure 4-3 Access control decision process	148

Figure 4-4 Per-flow QoE decision process.....	150
Figure 4-5: Evaluation test bed setup.....	152
Figure 5-1 RTP video MOS score in run no.1and 2	156
Figure 5-2: Bulk transfer and the throughput of the RTP vide in the two runs	158
Figure 5-3: RTP videos' MOS score in run no.1	161
Figure 5-4 RTP videos' MOS score in run no.2	162
Figure 5-5: RTP videos' throughput in run no.1	164
Figure 5-6: RTP videos' throughput run no.2	165
Figure 5-7 MOS scores for the videos when admission control is inactive.....	168
Figure 5-8 MOS scores for the videos when admission control is active.....	169
Figure 5-9 MOS scores for the videos when the administration system is inactive ...	172
Figure 5-10: MOS scores for the videos when the administration system is active ..	173

LIST OF TABLES

Table 2-1 Five level scale for overall video quality rating 39

Table 2-2 Comparison of QoE Measurement approaches 64

Table 2-3 Difference between User Centric and Techno Centric Approach 89

LIST OF ACRONYMS AND ABBREVIATIONS

1G – 1st Generation

2G – 2nd Generation

3G - 3rd Generation

3GPP – 3rd Generation Partnership

4G – 4th Generation

4G-LTE – 4th Generation – Long Term Evolution

5G – 5th Generation

ANFIS – Adaptive Neural Fuzzy Interference System

ANN – Artificial Neural Network

ANSI – American National Standard Institute

AP – Access Point

AppPol – Application Policy

ARP – Address Resolution Protocol

ARP – Allocation and Retention Priority

ARPU – Average Revenue per User

ATM – Asynchronous Transfer Mode

AVC – Advanced Video Coding

CAC – Call Admission Control

CBR – Constant Bit Rate

CBWFQ – Class-Based Weighted Fair Queuing

CCA – Catalog, Categorize and Analyze

CEAS – Customer Experience Administration system

CM – Control Manager

CPU – Central Processing Unit

CRAN – Cloud Radio Access Networks

D2D – Disk to Disk

DA – Discriminate Analysis

DCR – Degradation Category Rating

DiffServe – Differentiated Services

D-ITG – Distributed Internet Traffic Generator

DSCP – Differentiated Services Code Point

DT – Decision Trees

DVD – Digital Versatile Disc

EIGRP – Enhanced Interior Gateway Routing Protocol

EstQoE – Estimated Quality of Experience

ETSI – European Telecommunications Standards Institute

FFD – Full File Download

FIFO – First In First Out

FPS – Frames Per Second

FR – Full Reference

FTP – File Transfer Protocol

Gbps – Giga bits per second

GBR – Guaranteed bit Rate

GPRS – General Packet Radio Service

GSM – Generation Mobile System

HCI – Human Computer Interaction

HD – High Definition

HetNeT – Heterogeneous Network

HSPA – High Speed Packet Access

HTB – Hierarchical Token Bucket

HTTP – Hyper Text Transfer Protocol

ICT – Information Communication Technology

IEEE – Institute of Electrical and Electronics Engineers

IETF – Internet Engineering Task Force

IoT – Internet of Things

IP – Internet Protocol

IPTV – Internet Protocol Television

IPv6 – Internet Protocol version 6

ISO – International Organization for Standardization

ISP – Internet Service Provider

ITU-T International Telecommunications Union

KPI – Key Performance Indicators

LAN – Local Area Network

M2M – Machine to Machine

Mbps – Megabits per second

MBR – Maximum bit Rate

MHz – Megahertz

MIMO – Multiple-Input and Multiple Output

MMQ – Modeling Media Quality

MMS – Microsoft Media Service

MOS – Mean Opinion Score

MPEG – Moving Picture Expert Group

MPLS – Multiprotocol Label Switching Networks

Ms – Milli second

MSE – Mean Square Error

NAT – Network Address Translation

Netem – Network Emulator

NGMNs – Next Generation Mobile Networks

NNET – Neural Networks

OFDMA – Orthogonal Frequency Division Multiple Access

OSI – Open Systems Interconnection

OSPF – Open Shortest Path First

PCEF – Policy and Charging Enforcement function

PDA – Personal Digital Assistant

PDN – Packet Data Network

PDN-GW – Packet Data Network Gateway

Pe_Mod – Performance Model

Per_Ind – Performance Indicator

PEVQ – Perceptual Evaluation of Video Quality

PSNR – Peak Signal-to-Noise Ratio

PVC – Permanent Virtual Circuit

QBS – QoE Based Scheduling

QCI – QoS Class Indicator

QCIF – Quarter Common Inter-mediate Format

QoE – Quality of Experience

QoEInd – Quality of Experience Indicator

QoEMod – Quality of Experience Model

QoS – Quality of Service

QoSInd – Quality of Service Indicator

QoSMod – Quality of Service Model

RAN – Radio Access Networks

RATs – Radio Access Technologies

RDT – Real Data Transport

RF – Radio Frequency

RO – Research Objective

RQ – Research Question

RR – Reduced Reference

RRM – Radio Resource Management

RST – Rough Set Theory

RSVP – Resource Reservation Protocol

RT – Real Time

RTCP - Real Time Transport Control Protocol

RTMP - Real Time Message Protocol

RTP – Real-Time Transport Protocol

RTSP – Real-Time Streaming Protocol

SC-FDMA – Single Carrier Frequency Division Multiple Access

ServLevAg – Service Level Agreement

SLA – Service Level Agreement

SLAP – Service Level Agreement Policies

SSQ – Surveying Subjective QoE

SubsPol – Subscriber Policy

SVM – Support Vector Machines

TCP – Transmission Control Protocol

TCP/IP – Transmission Control Protocol/Internet Protocol

TFT – Traffic Flow Templates

TUQ – Testing User perceived QoS

UDP – User Datagram Protocol

UE – User Equipment

UMTS – Universal Mobile Telecommunications Systems

UX – User Experience

VANETs – Vehicular Adhoc Networks

VBR – Variable Bit Rate

VoD – Video On Demand

VoIP – Voice over Internet Protocol

VQ – Video Quality

VQM – Video Quality Metric

WCDMA – Wide Band Code Division Multiple Access

WFQ – Weighted Fair Queen

Wi-Fi – Wireless Fidelity

WiMAX – Worldwide Interoperability for Microwave Access

WRED – Weighed Random Early Detection

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

In the present technological dispersion, it is true that the concepts of Convergence and Next Generation Networks (NGNs) are now being widely accepted and embraced by telecommunications stakeholders. However, the evolution of NGNs has further meant mobile evolution of broadband, wireless and mobile communication technology which has paved the way for a new concept known as Next Generation Mobile Networks (NGMNs) or commonly known as 4th Generation(4G) wireless technology. 4G is the Fourth Generation of technology for mobile and other telecommunication equipment. This technology provides multimedia services anywhere/anytime globally and also provides other services like GPRS. This is the next generation of wireless technology, which has enhanced the data rate and multimedia service[1]. The technologies are fully Internet Protocol(IP) based integrated systems and networks that own unique IP addresses that are achieved after convergence of wired and wireless networks. They provide a data rate between 100Mbps to 1Gbps according to the environment. They also provide good quality of service (QoS) and secure transmissions to the user. 4G is based on IPv6 and uses the same bandwidth (5-20 MHz) as 3G, but gives a high data rate of 20Mbps or more. It only uses packet switching. The advantages of 4G technology are that it has a higher data rate than 3G[1]. It is an IPv6 based network and supports multimedia anywhere at low transmission costs than 3G. It is more secure and reliable than 3G. This network uses only packet based switching. The disadvantages of 4G technology are that equipment required to implement this technology is still very expensive. Globally, there are still some areas which do not have access to 4G. The other disadvantage is that 4G compatible

phones are more expensive than 3G phones[1]. Due to the fact that NGMNs base their services on mobile internet and multimedia streaming [2-4], mobile internet further means increased capability of mobile devices and applications to process information, store information, display images and stream video content, among other functionalities. User generated video is one of these applications, such as Hulu, You-tube, Netflix, among others. For instance, according to [5] , an individual would require 5 million years to be in position to watch all the video content that will cross IP networks every month in 2019. This fact uncovers the growing concern on the video quality(QoS) concept as far as the end user's experience(QoE) is concerned. This is in terms of the degree to which, when and how the latter becomes quality conscious[4]. The authors in [6] found out that a user would normally share their bad experience of service with other 13 prospective or potential users. This brings serious challenges to network developers and application designers with regard to retaining these users since they will simply change service providers in the hope of finding a better service on another network. In turn, if there's a migration of users onto other networks, there's bound to be a knock-on effect on the network infrastructure and resources of the network these users may have identified. For example, Wi-Fi hotspots.

ITU-T in [7] defined QoE as: "the overall acceptability of an application or service, as perceived subjectively by the end-user." But we need to note that ITU-T also acknowledges the following statements in relation to QoE.

- "QoE implies the overall end-to-end effects on a given system, such as terminals, networks, clients, services infrastructure, etc."
- "QoE can be the general acceptability of a service that can be influenced by expectations and context of use from a user's point of view."

The main observation here though is, ITU-T fails to indicate the detailed meaning of “context” and how “user expectations” are to be measured.

Furthermore, QoS can be seen as the process a network provider implements to deliver a satisfactory service with an assured level of service. It is mainly measured using jitter, packet loss, and throughput and delay parameters. Thus, if we are to compare the definitions of QoE and QoS, the latter is mainly network centric as far as achieving service quality is concerned while the former solely focuses on the user in achieving the same service quality.

Therefore, this work seeks to bridge the gap between QoS and QoE by addressing the main challenges in NGMNs that directly impact on QoE of users with regard to video streaming applications in these networks. In the process, both subjective and objective experiments are conducted to assess the quality of video streaming in NGMNs in the eyes of the User. The work ultimately intends to contribute to the service quality improvement of video streaming and network resource optimization in future Next Generation Mobile Networks and beyond. To achieve this, we presented a mechanism that monitors, controls, manages and evaluates QoE without affecting network QoS.

1.2 Problem Statement

The ever increasing portion of total internet traffic in audio and video multimedia services demands the availability of fast, efficient and cheap mobile connectivity. The need to meet the higher prospects of users and their necessities for economical mobile connectivity, reliable and efficient communication and access of any service across various operator domains all pose severe challenges to mobile operators. The capability of mobile devices for streaming the high definition content, poses another challenge for mobile service providers[8]; this is entirely because consumers have now become familiarized with additional service demand of resources and better quality services.

The major challenge for QoS and QoE frameworks therefore is to provide traffic control mechanisms based on the estimation, subscription and differentiation amongst application based traffic. Other challenges that have not efficiently been addressed for these networks to enjoy the envisaged success include: 1) End-to-end QoS; 2) Energy efficiency; 3) Resource allocation; 4) Mobility management; 5) Intelligent billing and 6) Call admission control. In our work though, we focus mainly on QoS in terms of its relationship and implications to QoE as a way of making sure the user has the best experience when connected through NGMNs.

The rapid growth of content data distribution over the internet has greatly reduced the global traffic over the internet. This has allowed network service providers to create workable business through the use of innovative streaming methods. The internet network provides the solution for data delivery and end-to-end system methods to address and achieve a high level of network QoE and services. The increasing infrastructure demands on the base station, Wi-Fi hot spots and delivering of video over HTTP makes it difficult to distinguish it from other web traffic. Failing to keep acceptable QoS standards may affect user expectations of the availability of fast, efficient but cheap mobile connectivity. In order to handle the different types of applications in the network as a way of improving the service quality in video and voice transmission, there has to be proper and efficient resource and traffic management. For this to be possible; there must be a mechanism to manage the resources and also handle the traffic in the most efficient way possible within the network[9].

In this thesis, we mainly concentrated our efforts on bridging the gap between QoS and QoE by implementing an independent mechanism compatible enough with other technologies, and also have the ability to monitor network QoS. At the same time, we should be able to evaluate trends in QoE as far as user perception of the service is concerned. While designing the mechanism, the following areas will also be put into consideration:

- Identification of all media flows;
- QoE differentiation and prioritization based on application types, classes of service subscriptions;
- QoE expectations.

Thus, the independent QoE monitoring and evaluation mechanism presented in this work has a high degree of quality of experience consideration by placing the user at the center. Therefore, it aids in providing the satisfactory connection between the user and network conditions. It also provides the information to the consumers with proper satisfaction during the network management process in a real time environment.

1.3 Research Goal

The main goal of this research work is to formulate, design and implement an independent user centered QoE and QoS monitoring system with an evaluation mechanism for video streaming in NGMNs..

1.4 Research questions

The following Research Questions (**RQ**) will be used to realize the research objectives as outlined in section 1.5.

RQ1: To what extent have network developers, network operators, researchers and manufacturers addressed the challenges of video streaming over the latest advancement in communication technology such as Next Generation Mobile Networks?

RQ2: Should end-user based studies on QoE always be conducted in “living lab” environments?

RQ3: How can QoS objective metrics be linked and related to QoE subjective metrics as a way of achieving network performance and user satisfaction?



RQ4: Is it possible to design and implement an independent mechanism for network developers to be able to monitor and evaluate a network using application quality degradation statistics from user feedback?

1.5 Research Objectives

In order to achieve the main goal of the research presented in Section 1.3, we shall employ the following research objectives (**RO**).

RO1: Analyze and review the extent to which the challenges of video streaming QoS and QoE in NGMNs have been addressed.

RO2: Carry out an investigation to establish whether QoE assessment and measurement approaches should be conducted within 'living lab' environments.

RO3: Evaluate and present a correlation between objective QoS and Subjective QoE mechanisms

RO4: Formulate, design and implement an independent QoE mechanism to administer, evaluate, monitor and provide network developers with application quality degradation statistics.

1.6 Research Limitations

The quality limitations of the mechanism in Section 1.5-(RO4) that may impact on the overall performance of the system include the following:

- a) The QoE mechanism has a limitation related to a specific video data traffic path. The system has specified anonymous senders, so data tracking is difficult. It also involves the limiting of several recognizable incursions of access controls within a video stream.
- b) The policy of video and data streaming has inflow of the data streaming for specific interruptions to the whole traffic sharing of bandwidth. This data flow has the limitation

about the data bit rate and signal to noise ratio. The precluding of the continuous data flow has particular data flow features that are persistent and similar to network dynamics that may limit the total signal flow.

- c) The issue about the implementation of the mechanism is data and quality compensation.

If the network system is not capable of providing sufficient resources to meet the customer demands then the users should be compensated.

1.7 Research Contribution

QoE involves bringing together user experience, expectations, emotions and perceptions onto an application and the performance of a network (expressed by QoS parameters). Quantifiable links between QoS and QoE are of paramount importance for the implementation of QoE control management systems with quantifiable QoS parameters. Therefore, the mechanism presented here in our work can be used by a service provider to manage and monitor the QoE of a user and be in position to effectively respond to complaints of video degradation during a streaming session. Figure1.1 shows a conceptual view of the mechanism implemented in our work. This mechanism can be located on any layer of the TCP/IP model, for example at the network layer on routers and access points or on the application layer where you can find a user workstation.

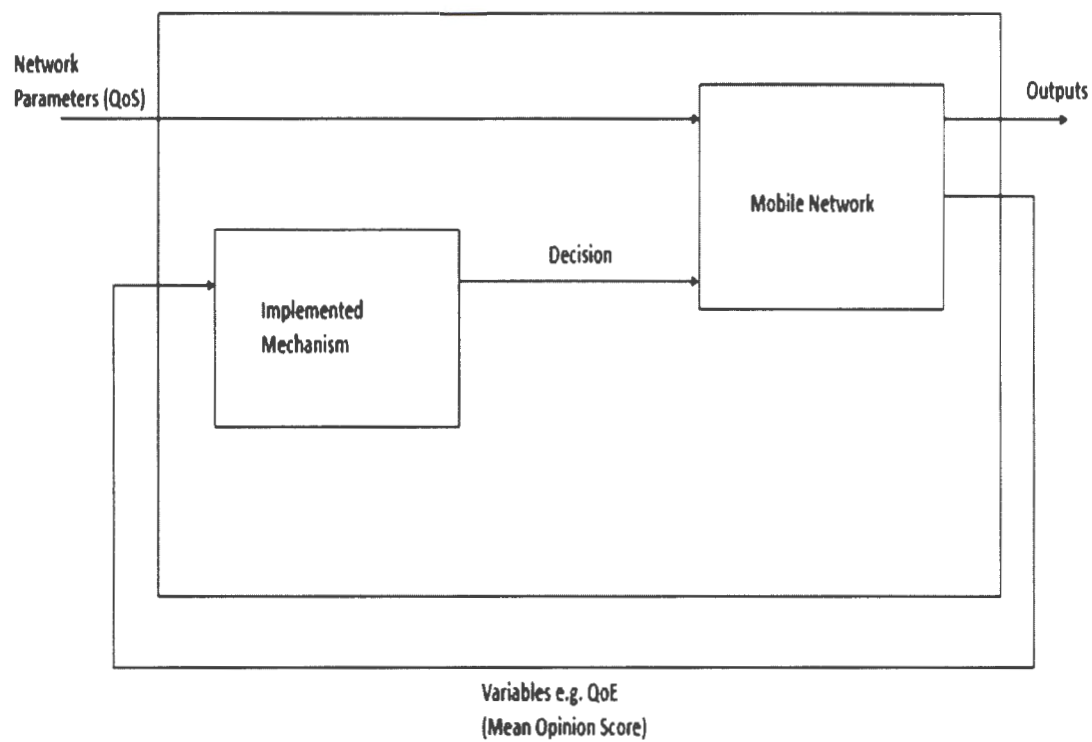


Figure 1-1 Conceptual view of research contributions

The main contributions of this research can be grouped in the following ways:

- **Monitoring network QoS and QoE.** The mechanism has the ability of gathering the necessary network data and use this data to monitor the network or a user application and be able to take measures where it is deemed necessary. In Figure 1.1, the variables act as the main input to the mechanism. The same variable can also be used to highlight network conditions in terms of performance. For instance, the sizes of queues at the router and load of traffic. On the other hand, the performance of the application in the network can be monitored (for example, QoE, data rates and delays).
- **Network traffic management.** The mechanism addresses the main challenges in network traffic management, prioritization, differentiation and resource optimization based mainly on QoE evaluation.
- **Network convergence.** The mechanism also supports the convergence in networks with a variety of infrastructure models of access, transport, and control service solutions to cater for services in a multimedia environment.

Over all, the mechanism makes it possible for network operators to pay more attention to their users as far as QoE is concerned.

- **User expectations.** The mechanism that's presented adheres to the various user requirements to realize reliable, cost effective and efficient way of communication albeit geographic constraints.
- **Network performance.** To achieve the desired network performance, the mechanism has the ability to:
 - i. Reduce compression of data to achieve efficient delivery of data packets minus a drop in quality;
 - ii. Minimize network data packet loss;
 - iii. Improve on privacy and minimize latencies in data transfer;

- iv. Manage and administer data interoperability and security;
- v. Optimize the connection and handover setup;
- vi. Reduce the handover and connection setup;
- vii. Minimize video streaming resolution pitfalls as a way of increasing the user data rate to realize the desired network experience.

1.7.1 Publications.

- *Francis L Lugayizi, Naison Gasela, Esiefarienrhe Bukohwo Michael* “Performance Evaluation of OSPF and EIGRP Routing Protocols for Video Streaming over Next Generation Networks”, Published in Advances in Computer Science Journal, Volume 4, Issue 5, September 2015.
- Evaluation of Video Transmission Quality of Service over Next Generation Networks. By: *Francis L Lugayizi and Obeten Ekabua*, proceedings of the 11th International Conference on Wireless Networks (ICWN’12: July 16-19, 2012 Las Vegas, USA).

1.8 Thesis outline

The rest of this thesis is organized as follows. **Chapters 2** presents the various challenges that NGMNs experience, it is also in this chapter that RQ1, RQ2, RO3 and RO1, RO2, RO3 respectively were all addressed. **Chapter 3** introduces the various concepts behind the data modelling and design of the mechanism mentioned in RO4. The chapter begins by introducing the research methodologies that exist to execute scientific research. It then introduces the methodology that was used in this thesis. The ultimate research contribution of this work is presented in detail in **Chapter 4**. This chapter introduces the various steps that were taken to implement the prototype mechanism presented in chapter 3 along with the various simulations that were carried out. The chapter also introduces the evaluation set up that was used to assess

the prototype mechanism implemented. RQ4 is answered in this chapter and hence achieving RO4. **Chapter 5** presents the results from Chapter 4 and then gives a full discussion and analysis of these results. The concluding thesis remarks, re-echoing of the research contributions and future work is then presented in **Chapter 6**. Cited references and appendix are then attached.

CHAPTER 2

RELATED LITERATURE

2.1 Chapter Overview

This chapter firstly presents an extensive review of the theoretical background and the basic concepts surrounding the evolution of NGMNs, QoS and QoE in these networks which basically form the backbone of this thesis. Then an analysis of the various challenges video streaming experiences follows. After that, the chapter then covers state-of-the-art works of other authors done in the area of NGMNs, video streaming QoS and QoE evaluation and measurement. The chapter also presents a detailed analysis on how QoE is implemented and improved. in order to achieve this, a comparison between QoE in the technical approach and QoE in the user centric approach is done. The chapter concludes by presenting the answers to RQ3 in order to achieve RO3 of this thesis.

2.2 NGMNs Theoretical Background

Next Generation Mobile Networks are a brain child of the 3rd Generation Partnership (3GPP). This partnership was launched in 1998 as a way of introducing technical specifications and reports for 3rd Generation (3G) mobile systems that evolved from GSM core networks[10]. Figure2.1 shows the various 3GPP standards and how they have evolved over the years.

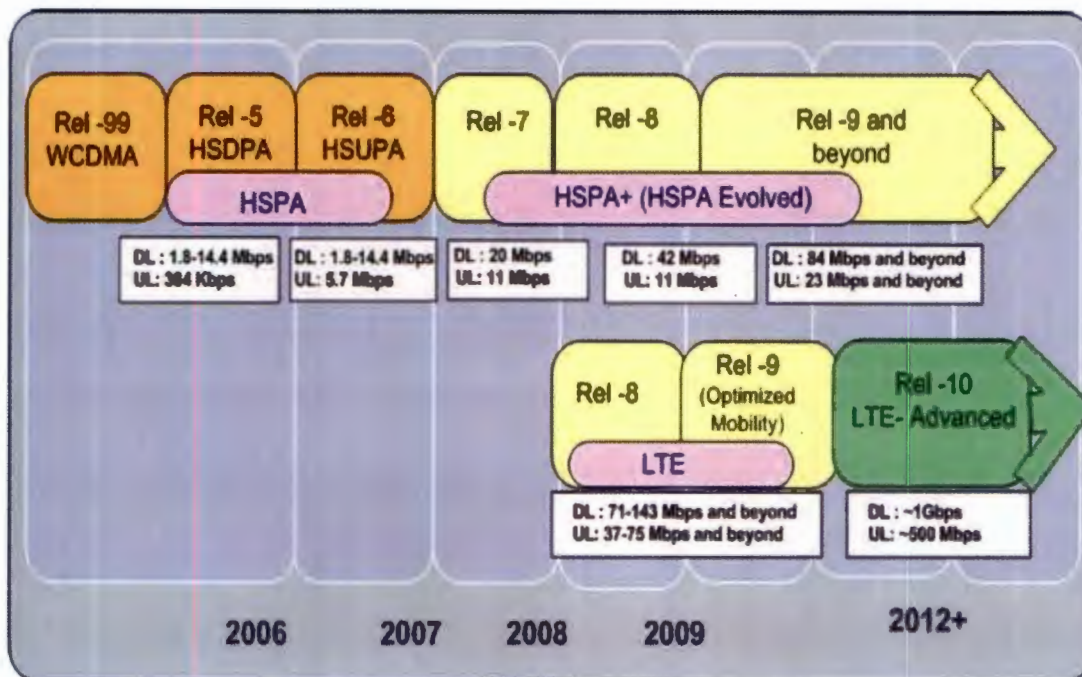


Figure 2-1: How 3GPP standards have evolved[11].

As seen in Figure 2-1, LTE was 3GPP's release 8 and remains one of its new releases; this explains why LTE is always referred to whenever Next Generation Mobile Networks are being discussed. In LTE, cell-edge users enjoy satisfactory data rates. For example LTE can support 300 Mbps of downlink and uplink of 75 Mbps plus a delay of less than 5ms[12]. With the latest LTE-Advanced (3GPP Release 10) peak data rates can reach 1 Gbps. The deployment of LTE networks is still ongoing and according to [10] it is estimated that there are 175 LTE networks with slightly over 90 million subscribers. If LTE is to be compared to 3G networks such as High Speed Packet Access (HSPA), LTE supports high spectral efficiencies and flexible system bandwidth. In comparison with Wideband Code Division Multiple Access (WCDMA), orthogonal frequency division multiple access (OFDMA) is the backbone of down-link access in LTE since it provides orthogonal access to many users in a given cell[13]. Up-link access on the other hand bases its operations on single carrier frequency division multiple access (SC-FDMA) when users are assigned a single virtual subscriber as a way of compensating for the variations in power of the transmitted signal. These variations normally originate from multi-carrier transmissions[10]. This generally turns out to be an advantage to battery-powered devices during this uplink.

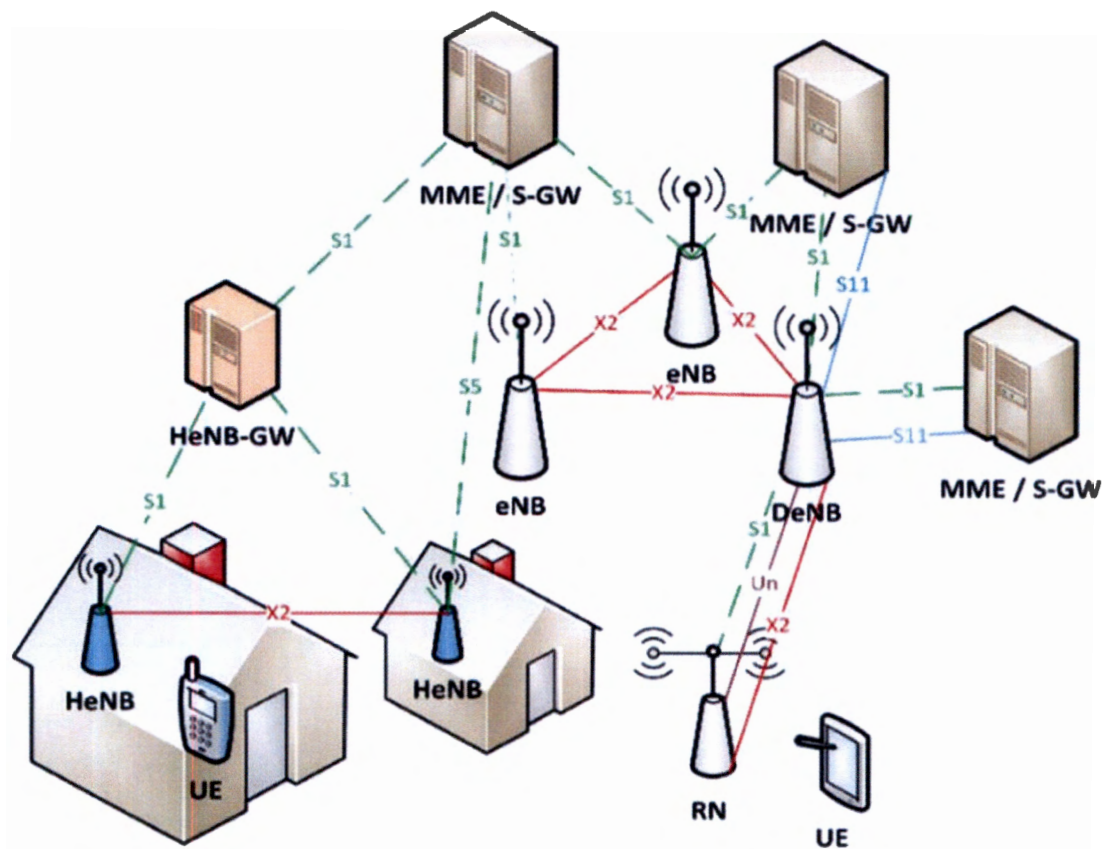


Figure 2-2 LTE Wireless architecture[14]

As illustrated in Figure2-2, the backbone components of access in the LTE network architecture are mainly the base station (eNodeB) and user equipment (UE). In a mobile cell, eNodeB mainly takes care of radio resource management (RRM)[15]. Further on, RRM is divided into two functional categories; dynamic resource allocation and scheduling executed on a millisecond (ms) criteria and then the other category is static (admission control and QoS management) mainly executed during the setup phase of the data transmission.

LTE's main goals in design remain lower delays, user-centric bit rates, increased spectrum efficiency, operational simplicity and lower cost[16]. MIMO and OFDM are the technologies that have greatly improved data rates in LTE. If this is to be compared to 3G HSPA, LTE uses OFDM mainly because this technology makes it possible to accommodate a wide variety of users and combining the latter with MIMO brings added advantages[17].

2.3 Challenges in Next Generation Mobile Networks

This section discusses the challenges in NGMNs. Research in these areas is ongoing and they have not efficiently been addressed for these networks to enjoy the envisaged success. The major challenges to this regard include: 1) End-to-end QoS; 2) Energy efficiency; 3) Resource allocation; 4) Mobility management; 5) Intelligent billing and 6) Call admission control. In our work though, we focus mainly on QoS in terms of its relationship and implications to QoE as a way of making sure the user has the best experience when connected through NGMNs.

2.3.1 QoS

Next Generation Mobile Networks (Wi-Fi, WiMAX and 4G-LTE) are multi-career networks with many different back-bone and access technologies. Because of this aspect, network providers experience a challenge in delivering the expected end-to-end QoS to users. This can negatively affect a user's experience of a particular service. It is for this reason that there is still a demand for more solutions that will be implemented to guarantee user application end-

to-end QoS in the mobile networks domain. Another aspect that complicates this QoS challenge is that NGMNs have the ability to classify the quality of applications basing on the user or the application itself. For example, applications such as video conferencing and video streaming always need more resources compared to other applications. This observation among others has led researchers to come up with various models to try and implement efficient end-to-end QoS [18]. As networks have evolved, it has been observed that in homogenous networks calls normally move from one cell to another cell with the same radio access technology. Coming to the heterogeneous nature of NGMNs, the same call can easily move from a single network to a totally different network with different access technology. This might seem an advantageous functionality but it comes with challenges as far as implementing end-to-end QoS since the user has no knowledge of this shift from one network to another. This can easily translate into negative QoE and in the long run a network provider can lose customers. Thus NGMNs still need mechanisms to close this gap between implementing QoS and QoE while focusing on the user and their applications.

2.3.2 Mobility Management and Seamless Connectivity.

Akyildiz *et al* in [15] made an observation that base stations always fail to support users switching base stations in the event that other base stations have weak signals to support current user sessions. Mobility management is the process of managing this migration from poorly performing base stations to others in the hope of better services. The ability of the base stations to transfer sessions to other base stations is referred to as handover. In NGMNs, this handover process turns out to be a challenging and complex task mainly because of the difference in the radio access technologies and cores of a network. Different researches have emphasized and recommended that in mobility, the handovers from different cells should be properly handled to maintain the desired QoS[15, 16].

2.3.3 Energy Efficiency.

The research community in computer networks continues to have interest in energy efficient network modeling. The authors in [17] state that in mobile networks, almost all networking equipment operate on limited battery power. Yet the heterogeneous nature of NGMNs requires much more energy compared to the traditional networks. The first challenge of this is that video streaming applications and other multimedia applications require more energy compared to their counterparts such as voice and data applications. Secondly, NGMNs do involve a lot of switching from one network to another and this process consumes a lot of energy. These challenges are being addressed at different levels in wireless networks, such as at the application, network and device levels. In order to strike a balance between QoS and energy efficiency, more work needs to be done because these two resources remain critical in NGMNs.

2.3.4 Call Admission Control (CAC)

The sole purpose of CAC is the reduction of the dropping probability at which connection of handover and new calls are take place in NGMNs. Dropping probability is the quantity of connection requests that are dropped by CAC in relationship to the total network requests. In a network scenario where congestion exists, calls are either delayed, rejected or accepted. A sub-system in CAC is mainly tasked with this decision. Incoming traffic in NGMNs is either real time in nature such as voice and video or non-real time for example, texts and images [19]. CAC therefore plays a big role in managing the main factors that may influence the performance of real time applications, these mainly include; packet loss, jitter and delay.

2.3.5 Resource Allocation

In NGMNs, upon a call being accepted, it has to be allocated the appropriate network resources to perform as expected. The allocation of the required resources such as bandwidth mainly

depends on the application in question (video, voice or data). Research is ongoing in designing resource allocation schemes as presented in [20-22].

2.3.6 Intelligent Billing and Cost.

NGMNs dictate that all applications in these networks can access different services and various resources far away from their own service providers. This fact necessitates network providers to have in place billing and customer relationship systems to make it possible to accommodate newer technologies as NGMNs continue to evolve technologically. Users should be in position to get billed for services they use regardless of their network service provider. However, service providers continue to work on expanding the existing systems in order to deliver flexibility in customized billing [23]. This too remains a challenging task on its own. Cost is another concern that comes along with billing. Prior to migrating to other networks, users may have been taking advantage of low service costs but as soon they move to other networks, the same might not happen. This is yet another challenge that various researchers are putting their heads together to strike equilibrium of service costing without affecting the QoS and the users at large[24].

2.3.7 Security and Fraud Management:

The architecture of NGMNs is vulnerable to organized crime groups and hackers. This is mainly because there is easy access to information, services and resources as well as identity theft. The biggest point of entry of these attacks is at the IP layer. Data on this layer moves in form of packets, this makes it possible for attackers to mark these packets with “borrowed” IP addresses that make it possible for restricted users to pretend to be the legitimate users. As the authors in [25] indicated, the hackers misuse the network services and end up benefiting more than the authorized users of a service who may be unaware of such an act and only find out when paying their bills. Yongsuk in [26] carried out a survey that indicated that the threats

discussed in this section cause service delivery interruptions, thus he ended up recommending that network and service providers have a responsibility of protecting their networks against different kinds of threats while providing end-to-end user security when they access their services. It is on this basis that there is still more work and innovation needed to guarantee the security of NGMN applications and the users of these applications.

2.4 Video Streaming in NGMNs

Lately, multimedia service applications have become an absolutely necessary characteristic as far as networking is concerned. This is entirely backed up by the amount of video and audio content that pops up hour by hour on the World Wide Web. Because of this development, a variety of systems continue to be designed to handle these media applications. For example, Video on Demand (VoD), Internet Television (IPTV), Video conferencing, Internet Protocol Telephony, to mention but a few.

Multimedia networking is defined by a set of hardware and software infrastructure that support multimedia streams on network so that users can communicate effectively. Using IP based network as a communication infrastructure is not a trivial task. Many problems arise when delivering multimedia data over IP.

Firstly, multimedia applications need a lot of bandwidth unlike the traditional text-based applications. A regular uncompressed video movie with a 1 minute length period, with a temporal and spatial resolution of 25 images per second, and 320 x 240 pixels respectively requires 330Mb and about 46 Mbps(Megabits per second) bandwidth for real-time transmission. As a consequence, audio-visual compression is vital for transporting such media on networking environments. This compression process leads to new problems that we expose later.

Secondly, a greater number of multimedia applications demand movement of data in a real-time fashion. Thus, video and audio data of these applications has to be played back in a steady or unbroken fashion along with the rate of sampling. Any failure of this data to arrive on time as expected, will lead to the play back process of the video being halted. However, buffering some parts of this data on the receiver can reduce this problem, however, the latency remains a challenge. For example, with IP telephony, one can tolerate an end-to-end transmission delay of about 250 milliseconds. If the latency exceeds the limit, there will be poor quality in the sound and an echoes will be encountered. In addition to high bandwidth and low latency, network congestion has more serious effects on real-time traffic. If the network is congested, the transfer takes longer to complete and real-time data becomes obsolete and unused if it doesn't arrive at the expected time.

Thirdly, the compression process will transform the traffic characteristic of audio video data stream from CBR (Constant Bit Rate) to VBR (Variable Bit Rate). For most multimedia applications, the receiver has a limited buffer. Smoothing the data can reduce the burstiness problem of multimedia traffic. If bursty traffic is sent, the application's buffer may easily overflow or underflow. The application will easily lose data if the data arrives too fast since the buffer easily overflows. This can result into poor quality video reception. On the other hand, if the same data arrives too slowly, the application can easily "be starved" because of the underflow of its buffer.

Lastly, since the Internet has millions of users using the network at a given time, a fairness problem in bandwidth share arises.

Nowadays, sharing of video is at the top of the various ways of entertainment and knowledge searches. To share a video effectively, video data has to be firstly stored and then transported over a channel of communication. However, the transportation of raw video over

communication channels is a challenge since raw video carries a large amount of data. Even if this raw video was to be stored, it would require one to have a significant amount of storage space on their device to store the raw video data. Camera footages are one source of redundant video data. The reduction of this redundancy in the raw video is required without compromising the quality of the video itself. This is where video compression comes into the picture, to reduce the redundant data by considering the quality of the video[27].

2.4.1 Video Models

Video models are the various ways that video is transmitted from its source to where the video is to be consumed. Below are the most common models that are used.

a) Real time (RT) Streaming video model

Real time (RT) streaming is a video model which is mostly used in mobile TV or real-time streaming services which use protocols such as real time streaming protocol (RTSP), real time messaging protocol (RTMP), and Microsoft media service (MMS) [28]. Upon reception of an initial request from the client to the server, the server then sends out one video frame at a time with equal time interval. If the video is encoded with 25 frames per second (FPS), then the server sends one video frame after every 40ms. Here, the server initiates the conversation with the client. Then, there are two separate communication channels set up in between them for data and control messages. The control messages include the commands like 'play', 'pause', 'stop' and 'seek'[28].

b) Hyper Text Transfer Protocol(HTTP) Streaming video model

Video contents are delivered over hypertext transfer protocol (HTTP) in this video model. This is also called as progressive download. HTTP streaming is the most popular video



algorithm used in cellular networks with over 98% of total video services [29]. This is due to the popularity of YouTube, which supports progressive download to watch video online.

A video file is not considered as real time streaming anymore; rather it is downloaded from a web server. The video is temporarily buffered at the local computer or mobile device. The media player can start playing the video as soon as some seconds of video data is available at the local buffer, which is configured by the video server. A user can view the video again and again within the same user session without downloading it again[28].

In case of HTTP streaming, the video file is segmented at the server. The media player at the client first downloads the manifest file, which contains the information about the URLs of all the available media segments. The media segments are much smaller than the actual media file and contain around 10s of media data. The duration of the media data per media segment is described in the manifest file. The media player then downloads all the media segments one after another using HTTP. Those received media segments are concatenated at the play-out buffer of the client. HTTP streaming is more robust towards the radio channel variations. It is possible to watch 500 kbps encoded video in a 256 kbps connection, which is otherwise not possible in case of the RT streaming. It is obvious in this case that it will take a longer time to download the file and viewer will also experience frequent interruption during the video session. However, buffered data in case of HTTP streaming helps to overcome the above problem, if the period of low bandwidth is temporary. Progressive download possibly uses either chunk based video model or full file download based video model or throttling based video model or some variations of these three models.

c) **Chunk based video model**

In the chunk based video model, the server sends a chunk of data, consisting of many video frames, in response to a request from the client instead of sending these frame by frame at a regular interval.

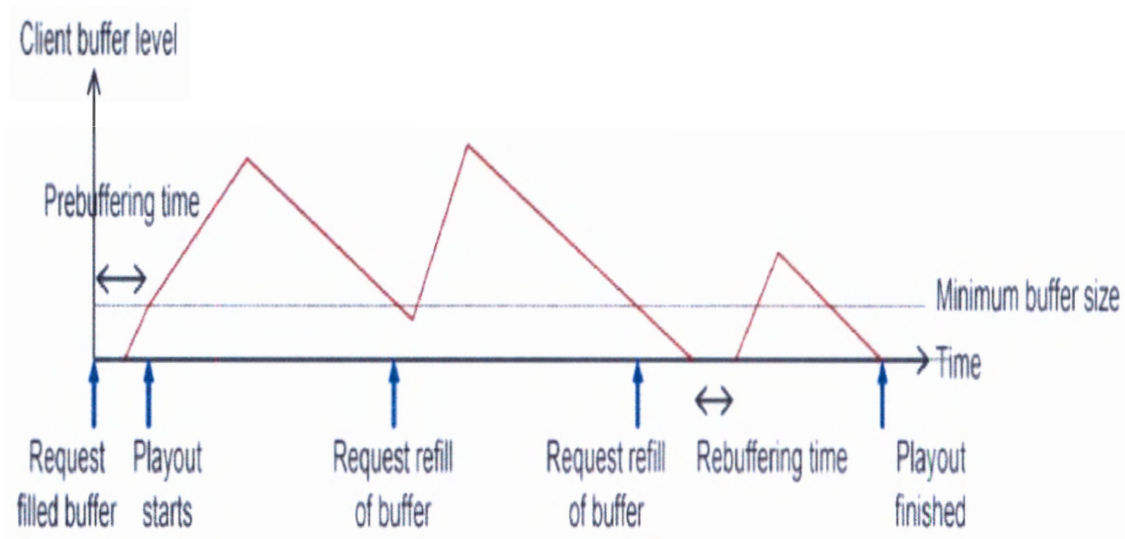


Figure 2-3 A video model based on chunks[30]

The concept of the chunk based video model is shown in Figure 2-3. First, the client sends one request to download the video file to the server via TCP. The client waits for some time before playing the video so that it can buffer the manifest file and also some seconds of media data. This initial waiting period is the pre-buffering time. Refill of the buffer is requested when the data in the local play-out buffer at the client goes below a minimum buffer level. It may happen that the local buffer at the client becomes empty during the video session and the client experiences a frozen image. This situation is called frame freeze. Re-buffering is needed before the client again starts watching the video. The time taken for re- buffering is called as re-buffering time. Too many frame freezes during the video session may cause the user to stop watching the video and leave the system.

The minimum buffer level is configured by the pre buffering time and the buffer multiplier. The buffer multiplier is assigned by the video server. Multiplication of these two parameters gives the minimum buffer level in video play-out seconds. For example, 1 second of the pre buffering time and the buffer multiplier as 3 gives 3 seconds of video play-out seconds as the minimum buffer level. If the video file is encoded with 1 Mbps, then the minimum buffer level is 3 Mb.

In the chunk based video model, server uses frame bundling to form a chunk of data. The server bundles a number of video frames into a media segment, where each bundle consists of a number of consecutive video frames. The client finds out about bundle structure from the information in the manifest file. The server then sends one bundle of data for each request from the client.

The server can configure different bitrates for each bundle, which gives the client opportunity to improve or degrade the video quality in one session. This rate adaptation has proven to be beneficial in current cellular technology, where radio link quality changes continuously.

d) Full file download (FFD) based video model

The concept of FFD is shown in the Figure2-4 with the line marked in red. In case of FFD, first the client sends one request to download the video file to the server and the whole video file is downloaded via TCP. The client then waits for pre buffering time to expire before starting to play the video so that it can buffer manifest file and some seconds of media data. In Figure 2-4, play out starts at t1 and the client completes downloading the whole video file at t2.

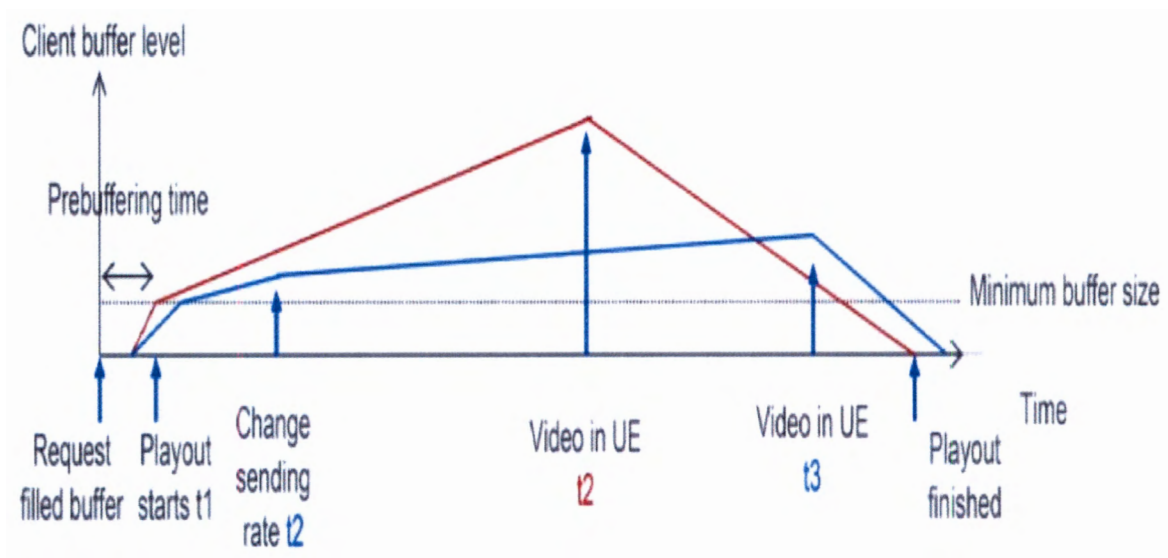


Figure 2-4. The status of the video file buffer in FFD and the Throttling based video model[30]

e) **Throttling based video model**

The concept of the throttling based video model is also shown in Figure 2-4 with the line marked in blue. For the throttling based video model, the client sends only one request to download the whole video as in case of FFD download. However, the server sends it in the beginning at 3 times higher bitrate than the actual video bit rate. After some time, the video server decreases the sending rate to 1.5 times higher in comparison to the actual bitrate. In Figure 2-4, change of bitrate occurs at t_2 and the video file completes download at t_3 .

2.4.2 User Abandonment

A viewer abandons watching a video on a mobile device when the video either has boring and unorganized content, poor quality or frequent frame freeze due to poor received signal power. Visible Measures carried out an online research and found, 20% of the audiences abandon the video within first 10 sec. of playback and 60% of the audiences stop watching the video within first 120 sec., as shown in Figure 2-5[31].

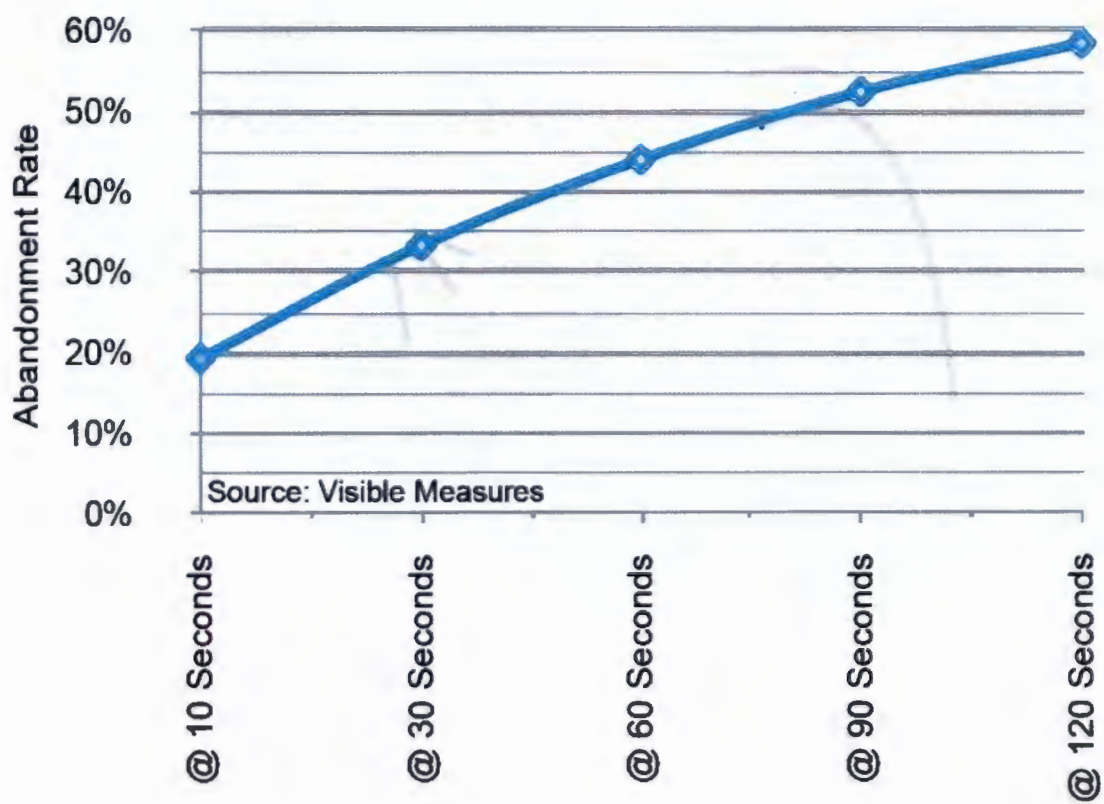


Figure 2-5 Viewer abandonment average rate by viewing time[31]

2.4.3 Video Compression

Compressing a video involves using mathematics based algorithms and theories to compress raw video to low bit rates. Compressing a video remains a key procedure for almost all video applications, for example, digital television, mobile TV, DVD video and internet video streaming [32]. The compression of video can either be lossless or lossy. Compression of a video in a lossless form makes it possible to secure all the information during the compression process. This therefore means that there's a possibility of getting back the original raw video data from the compressed video. When this is put into practice, the reduced data is actually less. Thus , the quality of a video can be well maintained using lossless compression [33]. However, when it comes to video streaming, lossless video compression can't be used since the compressed video data remains large.

In lossy compression on the contrary and as the name indicates is information that is lost with no possibility of recovery[34]. Obviously, the size of the video is reduced and the quality of video is degraded too. This technique is predominantly used for video streaming services. Even though the quality of the video is lost it is still perceivable. Therefore, video compression should be done at an optimum level while simultaneously considering the video quality.

2.4.4 Video Format

The video format consists of two distinct technology concepts: Video Codec and Video Container.

a) Video Codec

The video compression process involves a complementary pair of systems, a compressor (encoder), which converts the source video into compressed form before the transmission or storage and a decompressor (decoder) which converts the compressed data back to the original data. The encoder-decoder pair is referred to as CODEC. Video codecs are software

programs that compress a raw video into a compressed video of smaller data size in a way that the compressed video must play on the computer, since raw or uncompressed data requires a large bit rate, approximately 216 MB per second[32]. Video codecs can only compress a video, similarly audio codecs are used for audio files and played along with video files. Different types of codes are available to compress raw video and the selection of video codec depends on various factors such as size of video, type of video streaming and type of application[35]. There are many video codes available for video compression nowadays, some of them are MPEG-1, MPEG-2, MPEG-3, MPEG-4, H.264, and Vorbis. Among all available video codes H.264 is the most widely used codec. The main features of this codec include, providing better video quality at lower bit-rates, fast encoding speed and other advanced features that make H.264/AVC better than its counterparts[32].

b) Video Container

A video container describes the format in which a video has been compressed and stored [36]. Examples of such formats include: .avi, .mp4, .mov, and .asf.

c) H.264/AVC Codec

H.264 is a method and format for video compression. It was developed based on the concepts of earlier standards such as MPEG-2 and MPEG-4 Visual and provides better compression efficiency[32]. H.264 has the ability to compress video and retain the quality and also it is more flexible as far as transporting and storing of videos is concerned. Another aspect of H.264 is its ability to compress a significant variety of applications ranging from applications with low bit rates to those with high definition broadcast services.

2.4.5 Video Downloading vs. Video Streaming

There have been significant research efforts in delivering multimedia content over IP networks. Ideally, multimedia content such as video and audio are streamed from the server to the client in response to a client request. The client plays the incoming multimedia stream in real time as the data is received. The sections that follow review the most important techniques for delivering multimedia content to users.

a) Video downloading technique

The video downloading technique allows delivering video content through a download process. This technique is called video downloading to discern that the downloaded file is a video and not a generic file. The download is achieved using common protocols like FTP and HTTP. There are many peer-to-peer systems that exist to share video between communities, examples are Kazaa [37], gnutella [38], iMesh [39], and eDonkey [40]. The video file is very large and needs to be downloaded entirely before viewing the video. This technique requires patience from the viewer and also reduces flexibility since neither the quality nor the content can be checked before the end of the download.

The main constraint of the video download approach is related to the size of the video file. Since a video is a very large file, the download could take a long time and large storage space.

b) Video streaming techniques

The video streaming techniques enable simultaneous delivering and playback of the video. This technique overcomes the problems of video delivery via file download. It provides considerable additional features. The principle is as follows. The video is divided into units. Each unit is transmitted over the network. The receiver collects the units, reconstructs the video and begins the decoding. The playback begins without having to wait for the entire

video to be delivered. In general, there are short delays (5 – 15 seconds) between the start of the delivery and the beginning of the playback at the client. This delay, namely pre-roll delay, determines also the buffer size for storing the received units.

The advantages of video streaming are considerable. First, it is not necessary to wait for a long delay before viewing the video. A low storage space is required to store the portions of video to view. These two benefits are determined by the duration of the pre-roll delay.

Many applications are based on video streaming techniques such as Video on Demand (VoD), video conferencing and peer-to-peer video streaming. We discuss a few of them in the following sections.

- *Video on Demand Streaming*

Video on Demand (VoD) streaming enables the distribution of live and recorded audio and video streams across LANs, intranets and the internet after a client request. The client contacts the video server and requests the appropriate stream. Many VoD applications are available such as Microsoft Netshow, RealVideo of Real Networks, StreamWorks of Xing Technology, VDOLive of VDONet, Vosaic, Cisco Precept and PictureTel of Starlight Networks.

- *Video Conferencing*

Video conferencing is an application that uses video to join people in some live interaction. Videoconferencing makes possible interactive meetings with individuals who may be in other countries. The video is delivered to all the participants in real-time. Much as video conferencing comes with its advantages, it is expensive since it requires fixed reception and delivery deployments. The high consumption of bandwidth leads to high transmission costs.

- *Peer-to-peer Video Streaming*

In peer-to-peer computing, every peer becomes a server and a client at the same time. Therefore, playback of multimedia content by any user is shared by other peers. Basically, this sharing happens in the ‘open-after-downloading’ mode, while peer-to-peer video streaming uses the ‘play-while-downloading’ mode. One of the advantages of peer-to-peer video streaming is that peers have direct connection to other peers which makes it impossible to communicate via mediating servers.

A lot of research has been done to address the challenge of efficient multimedia streaming. Both live and on demand peer-to-peer video streaming. Video delivery can be in a multicast [41], [42] model or on demand as presented in [43] and [44].; Peer to peer layered video is also experienced in the work presented in [45].

2.4.6 Types of Video Streaming

Nowadays different types of video streaming applications are in use. Some of them are point-to-point communication, broadcast communication and multicast communication. All these video streaming applications are mostly dependent on two video streaming types, live video streaming and VoD, which is a process of real time transmission of a video over Internet[46]. The streamed video file can be viewed on smart phones, mobiles and personal computers. The video application is mainly used in live sports broadcasting.

VoD video streaming process is the opposite to the previous type. In this video streaming, the selected video file is played whenever the viewer wishes to watch the video. In this type of streaming, one-to-many viewers can view the same or different video[47]. This process is mainly observed in video streaming websites like YouTube, Hulu and DailyMotion.

2.4.7 Video Streaming Protocols

There are quite a number of protocols that make it possible for media streaming to take place.

The main ones include:

- Session Description Protocol (SDP);
- Hypertext Transfer Protocol (HTTP);
- Real-time Streaming Protocol (RTSP);
- Real Data Transport (RDT);
- Real-time Transport Protocol (RTP);
- Real-time Transport Control Protocol (RTCP)[48].

Any of the above protocols is used depending on a particular type of application at a particular point in time.

The other determining factor on the choice of protocol is the requirements of a particular service. The video streaming protocols are either UDP or TCP based. UDP is the preferred base protocol over TCP for video streaming. This is mainly because UDP has the ability to send packet using a constant bit rate and never gets concerned with lost packets in a network. TCP on the other hand is also a base protocol being used in video streaming. This is mainly because it has retransmission capabilities, congestion control and flow control [49]. However, TCP comes with a downside of delay because of the retransmission of lost data, something UDP does not accommodate.

2.4.8 Assessment of Videos

There are mainly two types of methods for video quality assessment. These are objective and subjective assessment.

a) Objective Assessment.

The objective video quality assessment method is based on mathematical models and fast algorithms that produce the results approximately equal to the subjective video quality assessment does not involve any human grading. It is a software program designed to deliver results based on error signal ratio of the original and processed video. The most popular objective methods are:

- Mean Squared Error (MSE);
- Perceptual Evaluation of Video Quality (PEVQ);
- Peak Signal -to- Noise Ratio (PSNR) [50, 51].

This method has two categories, i.e., referred to as Full-Reference (FR) and Reduced-Reference (RR) video quality metrics[51]. Objective data includes measurements of the communication processes and tasks outcomes [52]. PSNR and MSE are currently the most widely used metrics for quality evaluation of video compression techniques. The simplicity of the formulas for PSNR and MSE metrics, (formulas 2.1 and 2.2 respectively) alongside the low computation time they need give to these metrics their popularity. Further on, minimizing MSE is a well understood and interpreted technique from a mathematical point of view [53].

$$P S N R(d) = 20 * \log \frac{2^{25}}{\sqrt{M}} \quad (2.1)$$

$$M S E = \frac{1}{m*n} * \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |X(i,j) - Y(i,j)|^2 \quad (2.2)$$

Where X and Y are two monochrome images.

The only approximate relationship between PSNR and the video quality perceived by human observers, is the byte-to-byte comparison of the data without considering what they actually

represent. However, the authors in [54] show that techniques such as MSE and PSNR are lacking the understanding of the Human Visual System(HVS).

b) Subjective Assessment

The other video assessment includes subjective analysis, which is based on human perception. The subjective video quality assessment is considered an accurate way of measuring quality of video compared to objective assessment. For subjective video quality assessment, a set of videos are given to the subjects for rating the videos on a scale of one to five. The meaning of the grades for quality is given in Table 2.1. The rating given by the subjects are known as Mean opinion Score (MOS). The subjects can include experts and non- expert observers.

Table 2-1 Five level scale for overall video quality rating

MOS RATING	User Opinion
5	Imperceptible
4	Perceptible but annoying.
3	Slightly annoying
2	Annoying
1	Very annoying

2.5 Communication Ecosystem

Service provisioning in user quality satisfaction is the goal of each content and network provider. It is important to analyze all issues that are involved in the service provision chain to improve the delivered service quality. The understanding of the relations between the different issues is necessary and is described by means of a communication ecosystem by Kilkki in [55] and is as shown in Figure 2-6. The ecosystem has several issues, such as technical issues, business models and human behavior. Each of these issues uses its own language to describe the issue's internal relations. Whereas the technical community is only concerned with the performance and QoS of the network, the business community prefers to focus on the revenue each user brings in while the human behavior community concerns itself with happiness and QoE of the users. The identification of dependencies and relationships in the ecosystem enables us to highlight existing drawbacks in the service provisioning chain. As a result, these drawbacks can be selected and improved.



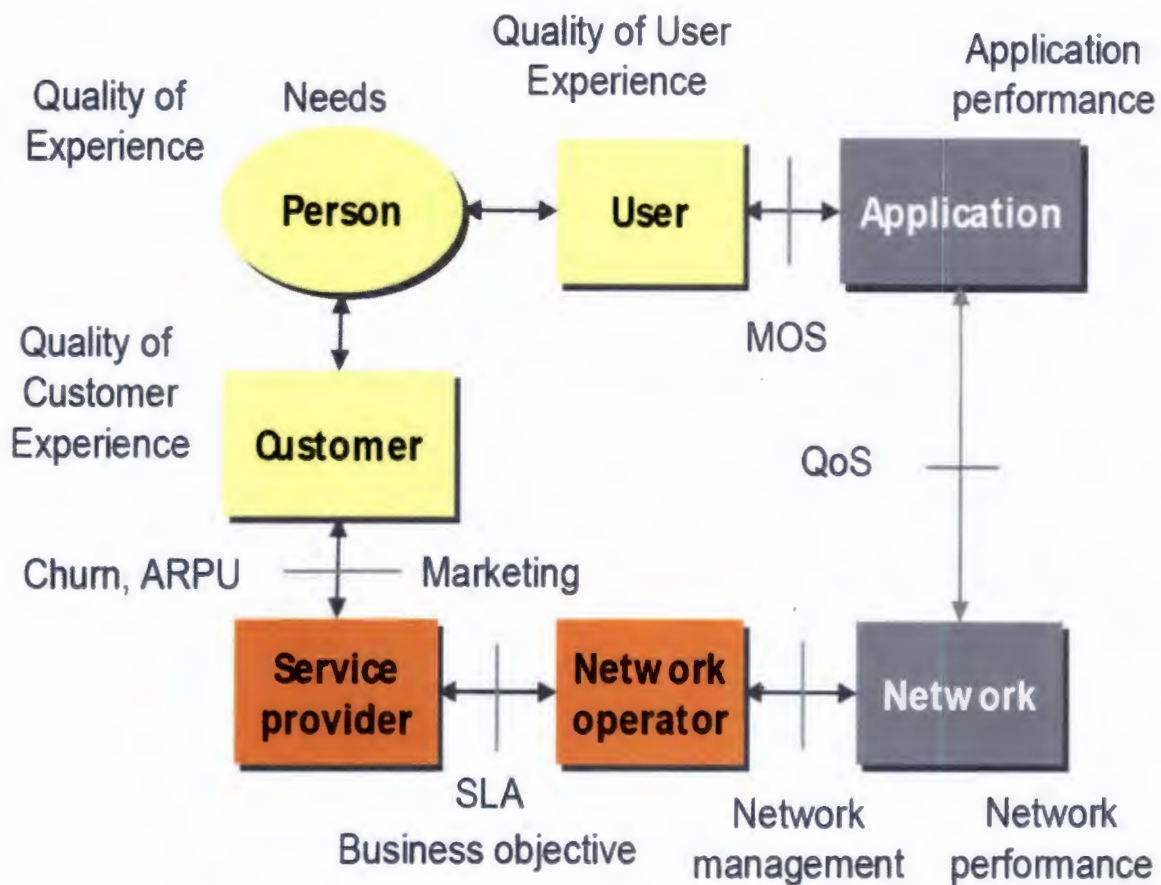


Figure 2-6 Communication Ecosystem Terminology[55]

The service provision chain attaches great importance to the network. In general the success of a network and service provider depends on the satisfaction of the customers. The price of the product such as Internet access can also lead to customer satisfaction. The framework for analyzing the communication ecosystems, shown in Figure 2-6, consists of seven modules. These are the user, the application, the network, the network operator, the service provider, the customer and the person. The person module is used to describe the complete ecosystem and thus, to close the gap between the user and the customer, in the framework the module Person represents a human with its both sides of sensation, such as user regarding service provisioning and as a customer regarding Internet access agreements. The user represents the receiver of a service provided from an application server and the customer represents the contractual partner of a network and service provider. This Person plays an important role within the framework. A purchasing decision of the customer often depends on the quality of experience obtained from service provisioning in the past of the user. Thus, the person defines the needs and the quality level of provided services that satisfy the personal sensation.

Beside these modules the framework in Figure 2-6 also presents the relationships and interfaces among the different modules. The interface between the user and the application is the mean opinion score (MOS) that is used to describe the level of quality of user experience with the provided service as shown in Table 2-1. Moreover, the interface between user and application is the relationship between the areas of human behavior and technology or usability. QoS is the interface between the application and the network that represents the provable performance of a network. This means that the service transmission from application server to the user depends on the level of QoS in the network. Thus, QoS has a direct impact on the provided service quality. The interface between the network and the network provider is similar to the relationship between the area of business and technology that can be described by the network management. The service provider describes the business objectives by means

of service level agreements (SLA) that are used as an interface between the network operator and the service provider. The interface between the service provider and the customer is described by the average revenue per user (ARPU) and churn representing the marketing relationship between business and human behavior.

The interpretation of Figure 2-6 leads to the assumption that improvements of the service provisioning chain need to investigate the network performance and the required application performance.

2.6 Quality of Service(QoS) and Quality of Experience(QoE) in NGMNs

Networks in the telecommunication industry continue to evolve towards multiple domain, service and vendor architectures. All these architectures strive to provide voice, data and video (commonly known as the triple play services) on the same IP network infrastructure along with other wireless 3G networks[56]. As far as end users are concerned, this scenario offers the opportunity of accessing a wide range of services from any wired or wireless terminal, enjoying high personalization levels and always-on availability. From the network operators, the ability to include triple play services and offer all other different services, from web browsing to Voice-over-IP(VoIP) applications, means a substantial reduction of operational costs and thus making full use of a flexible and innovative platform, that relies on the pre-existing backbone infrastructure. However, the deployment of these multiservice networks brings with it challenges of QoS issues and network policy control. In implementing QoS, network traffic has to be prioritized, examination of IP packets has to be done and identifying what specific requirements have to be guaranteed to users in implementing QoE. All these are of paramount importance in network traffic management.

Much as most research work on QoE focuses on assessing and monitoring of QoE [57-60], Management and evaluation of QoE has now picked up more interest from different researchers

around the world. So, various systems and frameworks meant to evaluate and manage QoE can be noticed in these research works. Kafetzakis *et al* in [61] presented a multifaceted framework known as QoE4CLOUD, their framework was meant for cloud computing based networks. The various challenges of managing cloud applications QoE are well presented by Hobfeld *et al* in [62]. As far as managing QoE in NGMNs is concerned. Gómez *et al.* [63] presented a network architecture that implemented QoE management. Fajardo *et al* in [64], proposed an administration system for QoE with regards to VoIP in 3G UMTS offerings. All in all, with the ever evolving technological networking trends in mobility, heterogeneousness of networks and network independent mechanisms and solutions remain inevitable. Mu *et al* in [65] goes on to propose an end-to-end multimedia applications QoE management mechanism they named QoE2M(QoE- aware Real-time Multimedia Management). This mechanism involves controlling the assessment of video and QoS/QoE mapping procedures. The mechanism basically detects network congestion times by using resource allocation controls that depend on collecting network status statistics. In a nutshell the QoE2M mechanism can quickly adjust applications based on the users' gadget means or network status. Much as Mu *et al.* presented a more vigorous way and bringing in the application's QoE features, they fail to present implementation or validation issues [66].

Vakili and Grégoire in [67] introduced a QoE mechanism for H264 video. Vakili and Grégoire make use of subjective simulations in assessing perceived video quality by the end users as a way of finding answers for the various effects of frame rate and quantization on perceived video and video data bit rate. The authors also go on to present a direct way of assessing the least transitional hop's bandwidth on each path at the same time continue to forward video traffic. The only limitation of this mechanism is that it restricts the type of video and size of the videos.

Kim *et al* in [68] introduces an “in-service feedback QoE mechanism”. With this mechanism, feedback from end users is received immediately there’s a bad experience with the service. Basically, the collected feedback data from the poorly satisfied end-users, network routers, and data from application servers is easily analyzed to trace the errors, etc. But again, just like with the QoE2M mechanism in [66], there was no implementation and validation of the findings.

Agboma and Liotta in [31] went on to propose their QoE mechanism with different QoE models for different kinds of multimedia content for basic mobile gadgets. Further on, this mechanism focused on the use of statistical mapping QoS parameters to user perception. In this approach, the vestibule where the service quality in the eyes of the user starts to decline is well defined. The downside of this is that, Agboma and Liotta only focused on ways to maximize end-user quality and they do not consider the impact of the network or how to optimize the usage of network resources [31].

An IPTV QoE mechanism was presented by Lloret *et al* in [69]. This mechanism was implemented between the end user’s network and the Internet Service Provider (ISP) architecture; this mechanism further implements a “process” to collect information from the two architectures. The main purpose of this collected information is to execute the necessary roles within the network to avoid unnecessary network congestion hence avoiding loss of packets. Although this is a significant progress, it doesn’t consider the effect on the network neither does it maximize network resources, this is because , the mechanism in charge of admission control to deliver the right IPTV service to end users of all the channels doesn’t put this consideration into context.

Despite a lot of research work focusing on presenting QoE/QoS mechanisms to manage QoE, literature still exists on scheduling, bandwidth management, admission control and congestion

control. For example, Piamrat *et al.* in [70] presents a QoE-driven admission control framework for IEEE wireless networks. In this work, an average Mean Opinion Score (MOS) on all live connections are calculated. If the result of the MOS is greater than the desired level, then another connection can be allowed, otherwise no connection is allowed.

Another QoE-mechanism was presented by Gallo *et al* in [71]. The main components of this mechanism include the application, user and QoS aspects of the network. The results from Gallo *et al* indicated that their mechanism was able to map QoE to QoS and thus benefit QoE management. However, the mechanism doesn't evaluate QoE since it only focuses on mapping QoE to QoS parameters, leading to the QoE management implementations gap currently in networking research, which is failure to implement and validate the mechanism.

Once again, in this thesis we introduce an independent mechanism that is compatible enough with other technologies. The mechanism monitors and evaluates QoE trends with reference to user perceptions. Using both subjective lab environment simulations and simulated environment testing is key in making sure the mechanism is implemented and validated, an aspect that different research works as discussed above weren't able to perform.

2.6.1 Quality of Service

The concept of QoS is often used with different meanings, such as user perception[55] and network performance parameters[72]. Sometimes QoS can be seen as a quality measure that refers to the level of quality as far as service delivery is concerned, i.e., the guaranteed or achieved service quality. In computer networking, QoS is mainly related to control mechanisms that manage resources in a network other than the attained service quality of that network. With this in mind, QoS involves giving priority to different applications and user traffic flow. Moreover, QoS is based on the ability to assure a particular level of network performance to particular network traffic, For instance, managing of traffic bit rate, packet loss, jitter and delay.

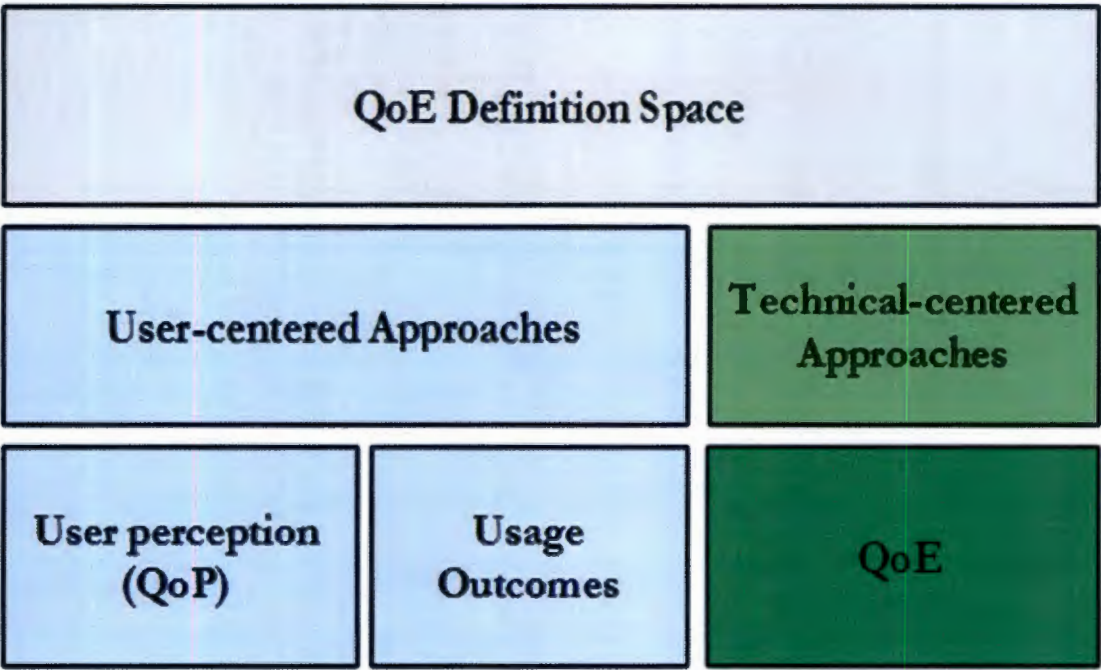


Figure 2-7 Relationship between various QoS and QoE approaches

Kilkki in [73] presents QoS as “a set of attributes that can be used to define the network’s capability to meet the requirements of users and applications”. However, in 2008 Kilkki in [55] tilted the definition a little and removed ‘users’ from the definition as it had been presented in [73]. X. Lu et. al. in [74], defined QoS “as the ability of the network to provide a service at an assured service level and that QoS encompasses all functions, mechanisms and procedures in the cellular network and terminal that ensure the provision of the negotiated service quality between the user equipment (UE) and the core network (CN)”. The authors go ahead to state that QoS is basically a technical concept that can be expressed, understood and measured in terms of networks and their elements.

The absence of a universal QoS definition led to various interpretations and definitions that may not necessarily be compatible or related to each other. Because of this, standardization organizations developed various QoS definitions. Examples of these organizations include ITU, IETF, ATM Forum and OSI. The international telecommunication union – telecommunication standardization sector (ITU-T) fully describes QoS in ITU-T Rec. E.800 standard called “Terms and definitions related to quality of service and network performance including dependability” [75]. What follows are the various definitions and notes on QoS from ITU-T Rec. E.800:

“QoS is collective effect of service performance which determines the degree of satisfaction of a user of the service”.

ITU-T Rec. E.800’s main terminology presents various issues of QoS i.e. Service operability, Service security, support and Server-ability. The Standard presents mainly two issues to note as far as QoS is concerned. These are:

- a) “The quality of service is characterized by the combined aspects of service support performance, service operability performance, Server-ability performance, service security performance and other factors specific to each service.”
- b) “The term “quality of service” is not used to express a degree of excellence in a comparative sense nor is it used in a quantitative sense for technical evaluations. In these cases a qualifying adjective (modifier) should be used.”

ITU-T Rec. E.800 goes ahead to point out that service operability, service security, support and Server-ability rely on network features. The standard defines performance as “the ability of a network or network portion to provide the functions related to communications between users.” The standard takes the following issues into account to fully explain what network performance entails; transmission, charging, dependability and traffic-ability. ITU-T Rec. E.800 goes on to provide four notes to stay aware of, these are:

- a) “Network performance applies to the Network Provider’s planning, development, operations and maintenance and is the detailed technical part of QOS, excluding service support performance and human factors.”
- b) “Network performance is the main influence on Server-ability performance.”
- c) “Network performance measures are meaningful to network providers and are quantifiable at the part of the network to which they apply. Quality of service measures are only quantifiable at a service access point.”
- d) “It is up to the Network Provider to combine the Network Performance parameters in such a way that the economic requirements of the Network Provider, as well as the satisfaction of the User, are both fulfilled.”

Figure 2-8 is a conceptual overview of the ITU-T E.800 framework.

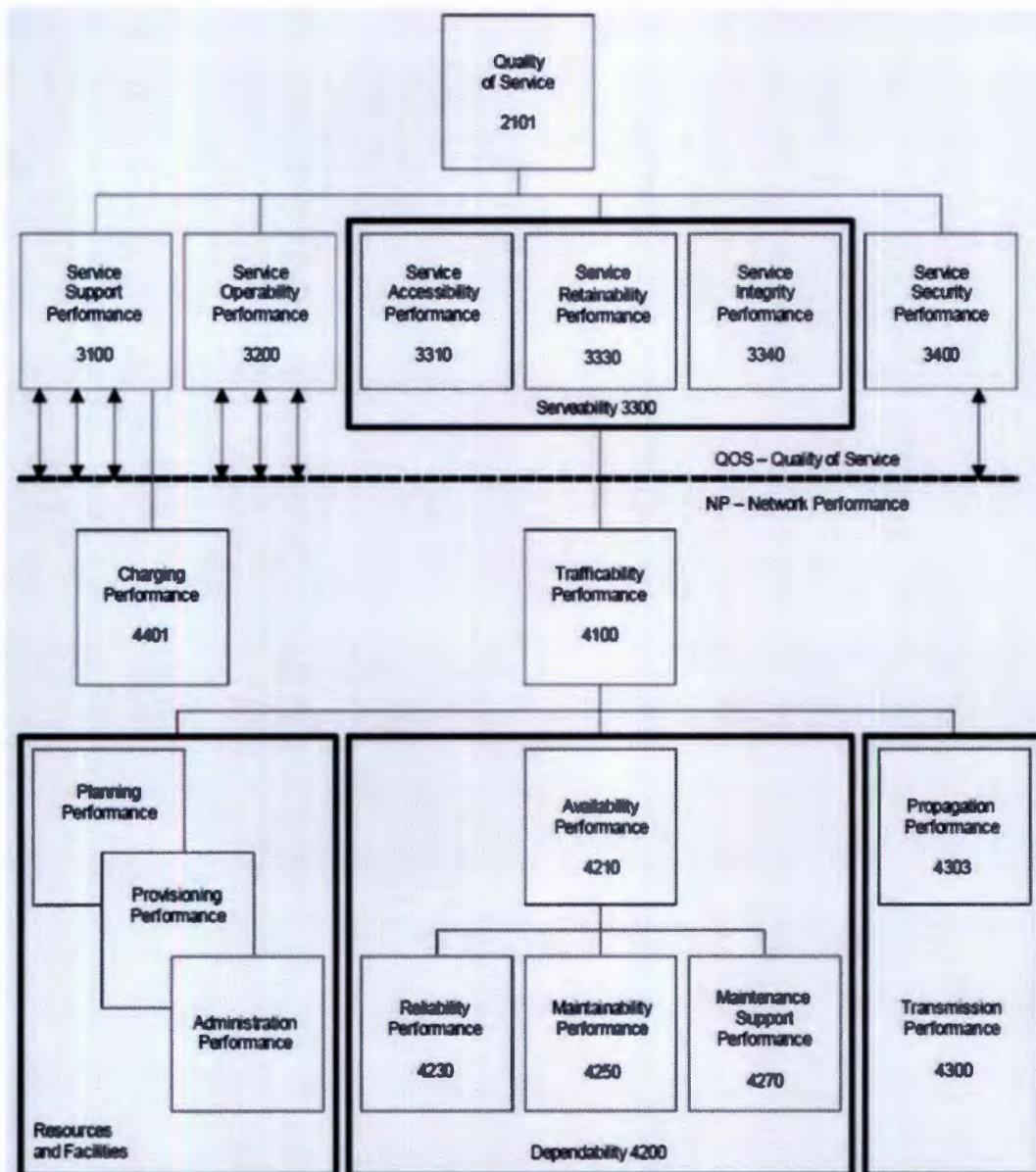


Figure 2-8 ITU_T E.800 framework[76]

The European telecommunications standards institute (ETSI) has debriefed customers with the goal to describe QoS criteria and to derive a uniform QoS matrix, that is described in ETSI ETR 003 [77], this is in order to present a relation between quality and QoS. ITU-T G.1000 [78]. But, QoS requirements for various service types are not categorized in the matrix. ITU-T Rec. G.1010 in [79] presents a user-centric model for categorizing multimedia QoS and service requirements are also defined. The categorization breaks down services in the form of two service types, i.e., delay dependent and packet loss dependent services. Packet loss dependent services are further broken down into two forms of services: error intolerant services and error tolerant. On the other hand, the delay dependent services are decomposed into the following classes: i.e., timely, non-critical, interactive and responsive. Figure 2-9 is a combination of information from two diagrams found in ITU-T Rec. G.1010 such as mapping of user-centric QoS requirements and model for user-centric QoS categories. The two packet loss dependent services and four delay dependent services result in eight QoS categories as shown in Figure 2-9.

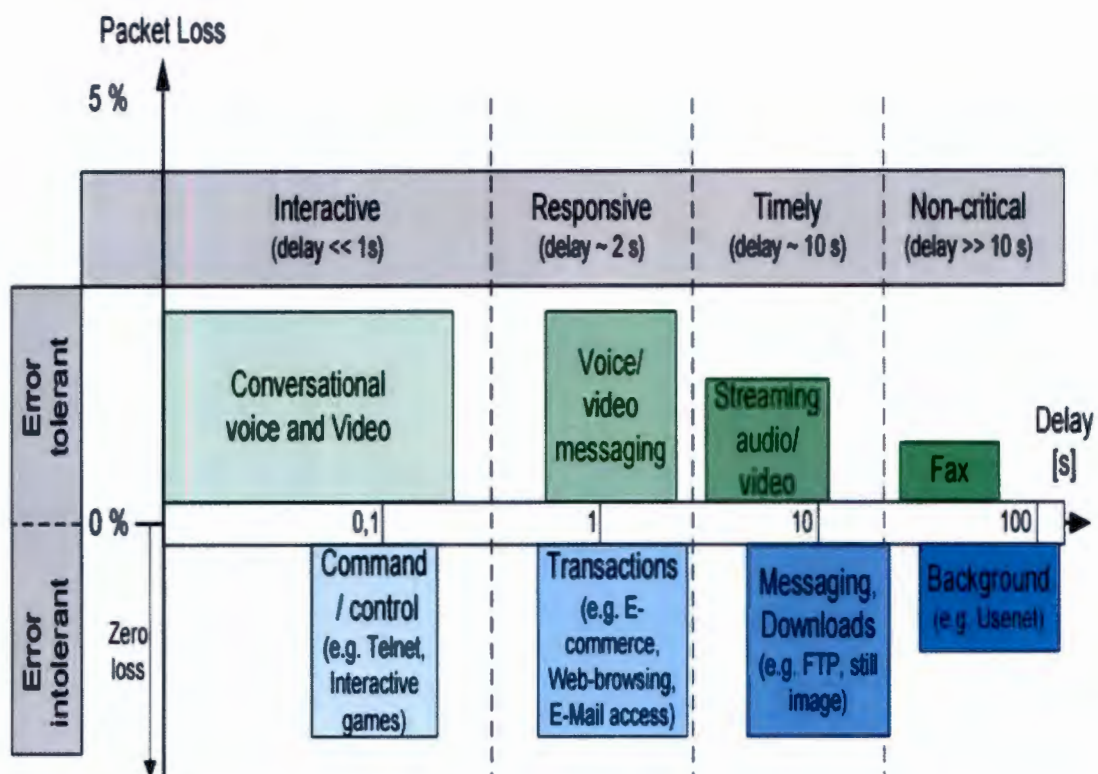


Figure 2-9 Mapping of user centric QoS requirements and model for user centric QoS categories based on ITU-T Rec G 1010[80]

The provided service quality determines the service acceptance of customers. Quality in a service is mainly influenced by jitter, packet loss and network delay. But, the type of service also determines the degree of impact most importantly when it comes to real-time streaming applications, for example, online games, VoIP and IPTV.

However, the degree of impact depends on the type of service. Especially for real-time streaming applications, such as VoIP, online games and IPTV. Whenever network capacity is limited, QoS guarantees become very important. Real-time streaming applications often require a fixed bit rate and are delay sensitive. It is a challenge to deliver reliable network performance to *meet all* the service requirements, this is inevitable where a network's capacity is greatly limited. In Figure 2-9, conversational voice and video require the highest network performance in relationship to delay. The demand for a smaller delay is challenging when customers are on the move and handover processes have to be performed.

In recent years QoS mechanisms have been developed to manage the network performance to fulfil the network requirements of a service. Particularly, prioritizing of conversational voice and video data is focused on, as described in [81].



▪ **Quality of Service Parameters**

QoS parameters provide a means of specifying user requirements that may or may not be supported by underlying networks. QoS in a network is of paramount importance since continuous media streaming applications continue to emerge, because of this, network developers cannot overlook QoS. For example, video streaming and transmission processes need high throughput which automatically dictates high bandwidth. On the other hand, audio transmission doesn't require a lot of bandwidth to be transmitted across the network. The other factors that should be toughly examined are delay variations as these mainly affect time-sensitive traffic. The next four sections here- after are therefore meant to give us an

understanding of the various QoS parameters. They are measured and quantified. This is important because these parameters are very critical in ensuring a congestion free network whenever there is video traffic traversing the network.

Delay

End-to-end transit delay is the elapsed time for a packet to be passed from the sender through the network to the receiver. The higher the delay between the sender and receiver, the more insensitive the feedback loop becomes, and therefore, the protocol becomes less sensitive to short term dynamic changes in the network. There are two ways in which Delay can be analyzed. The variable delays consist of queuing delays at each node or link on the route between source and destination. Figure 2-10 illustrates the factors of packet delay

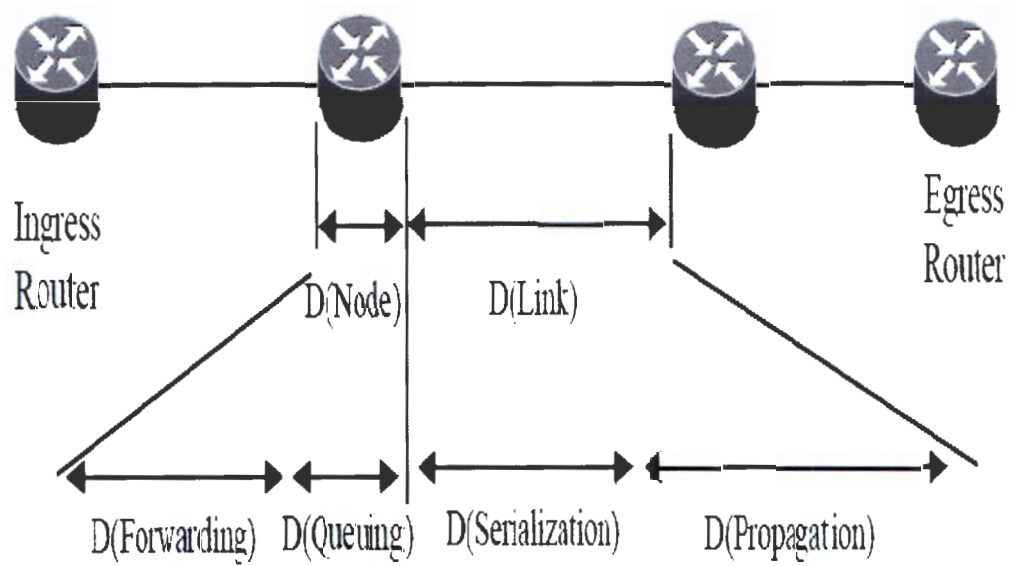


Figure 2-10 End-to-end Delay Calculation

Jitter

If there is a variation in the end-to-end transit delay in a network during video streaming, it qualifies to be called jitter. Most network developers, engineers and researchers also refer to jitter as delay variation.

In packet-switched networks jitter defines the distortion of the inter-packet arrival times compared to the inter-packet times of the packet transmission.

Bandwidth

The maximal data transfer rate that can be sustained between two end points of the network and is defined as the bandwidth of the network link. However, it is equally important to highlight that the bandwidth may not only be limited by the physical network infrastructure of the current paths of the various networks in question.

Loss

The process of quantifying loss is done by calculating the overall loss rate, which is equal to the total amount of lost traffic divided by the total amount of input traffic over a certain period of time. The expression of loss rate is given as;

$$L_r = \frac{L_{lost}}{L_{total}} = \frac{L_{lost}}{L_{total}} \quad (2.3)$$

▪ Basic QoS Operations

The heterogeneous nature of NGNs comes with a variety of applications and their traffic flows, some of these applications have their unique requirements that the network should meet as a guarantee for quality in service. Therefore, without these requirements being met, the overall quality of the application is greatly compromised. Quality of Service operations are the various ways that are used to effectively manipulate QoS parameters

discussed in Section 2.7 to fully implement Quality of Service in a network. This section presents the main QoS operations.

Traffic Classification

Traffic classification is a process of categorizing traffic within a network by determining the different classes of service. It also includes the process of identifying the different applications and protocols that are involved in the network. Traffic classification may take on different actions such as discovering, monitoring and optimization, all of which are meant to achieve the ultimate goal of QoS – effective network performance. Typically, once the packets are classified (identified) as belonging to a particular application or protocol, they are marked or flagged.

Traffic Marking

In Traffic marking, the packets are colored along with a special value depending on how they were originally classified. The main purpose of traffic marking is to enable other network operations to determine the origin of packets and be able to apply the right policy or priority.

Traffic Policing

Traffic policing involves the process of monitoring network traffic for compliance with a traffic contract and taking steps to enforce that contract. Policing uses a token bucket scheme to limit the number of packets that will be transmitted.

Traffic Shaping

Traffic shaping is also known as packet shaping. The latter also uses the philosophy of the token bucket scheme as a way of guaranteeing network performance through increasing the usable bandwidth in the network and improving the latency levels of the

traffic in the network. If a link becomes saturated to the point where there is a significant level of contention (either upstream or downstream) latency can rise substantially.

Traffic Congestion Management

In any given network, congestion is bound to occur. Congestion is a situation where a node or link in a network carries a large chunk of data to the extent of compromising the quality of service of that network. However, there are quite a number of techniques that may be used to manage the traffic flows and avoid network congestion; these are known as queuing techniques. The subsection that follows presents some of these techniques.

Key performance parameters (KPP)

A KPP is a measurable metric which allows characterization of QoS requirements and evaluation of network performance to take place. The main purpose of KPPs is to realize NGMNs QoS objectives. KPPs make it possible to arrive at an objective and quantitative solution when comparing obtained quality performance with the desired quality performance. KPPs can also be employed in classifying traffic and analyzing networks. In [82] KPPs broken down into two service categories. These are:

End-user metrics. These are metrics that give an evaluation of network performance as it is perceived or seen by the network's end-user. Using end user metrics is one way of evaluating the satisfaction level of the network end-user.

Network metrics. These are metrics concerned with the network administrator and the main purpose is to keep track of system performance during the service delivery. Network metrics do not directly present the status of QoS that is being provided by the network but indirectly they do have an effect on the network end-user metrics. However, network administrators can make use of network metrics to manage and control network resources.

2.6.2 Quality of Experience (QoE)

QoE involves getting to clearly understand how a service performs after it has been delivered to a customer. The main purpose of this process is to ensure that the service meets the customer's requirements and expectations. In order for network operators to implement and manage QoS in a network, understanding the experience of the end user is of paramount importance. QoE is mainly measured directly from the user/subscriber's device when a service has been delivered. For this process to be successful, the various factors that influence the experience a user is bound to have when and after using a service have to be analyzed. Figure 2-11 illustrates a summary of all the factors that have to be looked at in relation to QoE. The categorizing of these factors is a way of understanding them systematically.



Figure 2-11 Quality of Experience factors to consider

Of late, end users have become quality meters along with their perception and expectations of a service, application or a product. All these have a great value to the end user.

ITU-T defines QoE as the “overall acceptability of an application or service, as perceived subjectively by the end user”[83], yet ETSI sees QoE as “a measure of user performance based on both objective and subjective psychological measures of using an ICT service or product” [84] and the standard goes on to extend QoE past subjectivity and accommodate the objective psychological measure. Therefore, QoE can be looked at as a multidimensional concept that encompasses both subjective and objective aspects[85]. So, QoE can be related to QoS which is primarily an objective, technical and techno-centric concept [86] that is more or less considered a much more user-centric concept. As it has already been presented, QoS concentrates on how network and application performance is perceived by the user in terms of quality while QoE mainly is all about focusing on how an individual feels or what an individual experiences after accessing the service. The sources of impact on the experience of the individual can include prior experiences, thoughts, feelings, expectations, context, etc.

The authors in [87-91] provided explanations of QoE which leaned towards a tech-centric background and ignore the aspect of human subjectivity thus lacked a much more broad explanation of QoE [92]; that’s to note that, an abrupt switch of products and services as far as the consumer is concerned can be as a result of assuming optimizing of any QoS parameters leads to overall QoE. Nonetheless, QoS can be seen as subdivision of the general QoE sphere. In most cases a high degree in QoS could lead to a high degree in QoE. Furthermore an assumption can be made that service and products which meet end user’s expectations and requirements and also allow them to enjoy high QoE in their personal scope will mostly achieve success compared to services and products that enjoy higher QoS yet they are unsuccessful to satisfy end user’s expectations and demands [93]. Another modern definition by EU Qualinet community in [94] encloses various issues on QoE and goes ahead to define

it as “the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user’s personality and current state. In the context of communication services, QoE is influenced by service, content, device, application, and context of use” [94].

Different applications will have different QoE requirements. To implement QoE for a specific application, understanding and identifying of a variety of factors that affect the application is of paramount importance. This has to be done from the members of the service delivery chain point of view along with the ways they may impact QoE. With advances in technology, the results of QoE modelling determines the parameters that are to be measured and monitored. The eventual outcome of this process has effective QoE optimization strategies. It is for this reason that managing QoE is categorized into 3 phases[62], these are:

- a) QoE modelling;
- b) QoE monitoring and measurements;
- c) QoE optimization and control.

If QoE management is well modelled and implemented with validation, end users are bound to be satisfied because their expectations or requirements shall have been well addressed. This can in turn be an advantage to technology development since the users will comfortably adjust to more complex services. Another advantage of successful QoE management is that, all stakeholders in the service delivery chain such as network developers, network providers, content providers and device manufacturers can have pride in being involved in the service delivery. The more realistic benefit is that, the competitive nature and environment in computer networks means that “price levels are decreasing and pricing schemes are becoming more similar” [95], service availability to end users isn’t enough since these users now have

a wide variety of service providers to select from. Therefore, the stakeholders in the process of service delivery continue to move more to meeting at all costs the expectations and requirements of users through achieving user satisfaction as far as perceived service quality is concerned.

▪ **QoE Measurement**

QoE is best evaluated using living labs[96]. Living labs are mainly user-centered and entail users to dry run newly developed services or applications. An analysis of living labs is further presented in the next section. The other way QoE is measured can be by selecting a panel of representatives with actual users of the service. Both cases involve the use of questionnaires as a tool of evaluating a service after use. These questionnaires’ results can be further use to evaluate QoE by using the Mean Opinion Score (MOS) [97] as already presented in Table 2-1 of Section 2.3. Table 2-2 presents a comparison between objective and subjective methods of measuring QoE.

Table 2-2 Comparison of QoE Measurement approaches

	Data collection	Data transmission	Comparison
Objective Methods	QoS	Involves Bulk data transportation	▪ Wastes less time. ▪ Has got specified QoS (accustomed QoE). ▪ There no need for user feedback. ▪ Mostly inaccurate.
Subjective Methods	Surveyed QoS/QoE	Data transmission is not necessary	▪ Personalized QoE (User-centered). ▪ Wastes a lot of time. ▪ Accurate and dependable QoE. ▪ Mainly depends on user feedback.

2.7 QoE and QoS evaluation in testbed-oriented Living Lab setting

The notion of evaluating QoE and QoS in a mobile, testbed-oriented Living Lab setting realistically does bond well with the current user-centric approach. This approach seeks to close the gap between technical parameters of a given network and human experience factors. With this finding in mind, QoE remains a multi-dimensional phenomenon which demands a perspective from all disciplines. Much as previous research has presented various ways of evaluating perceived QoE, the only setback has been that these works are only inclined towards objective assessment of QoE and do not pay attention to factoring in subjective assessment of end user experiences.



As far as technical testing of mobile devices is concerned, the school of thought on using test bed environments continues to prove its value in recent years. Reference is made to the standardized laboratory environments used for testing new technologies, products and services and protected from the hazards of testing in a live or production environment [98].

Lately though, the authors in [99-101] identify a couple of research works that continue to castigate test bed environments that have a less degree of naturalism. The authors of these works do point out that there's a variety of products that have successfully been developed in these lab environments that are well controlled but upon being deployed in the actual and natural environments, they either underperform or fail to do so after the deployment.

Ponce de Leon, Eriksson *et al.* in [102] look at simulation beds as very important tools to blend innovative components of technology within complicated wireless environments and also bring end users into the mix. Ponce de Leon, Eriksson *et al.* go on to highlight that technology on its own becomes obsolete if the benefits or its usefulness for end users in day to day life isn't proven first to be successful. The outcome of this is that, more scholars and authors have pushed on user-centric approaches in researching on innovative infrastructures meant to assess,

evaluate and manage new applications and service while putting the end user at the center [103-105]. Furthermore, this overture has a link to the appendage of the contained traditional test bed environments towards more natural setups.

Without under looking the above observations, it is admissible to acknowledge the enactment of living labs but with an open and modern criterion. Living labs are “ environments for innovation and development where users are exposed to new ICT solutions in (semi)realistic contexts, as part of medium- or long-term studies targeting evaluation of new ICT solutions and discovery of innovation opportunities” [100]. Opposed to the traditional test beds, living labs strive to yield a more native setting to clearly examine while involving users in the process of innovation. With this paradigm shift in research, research works now make reference to the pervasive computing environment where living labs were established to study ubicomp (Ubiquitous computing) solutions[101, 106]. In addition, in the nineties, living labs were popularized as facilities that were to connect end users to ICT testbed applications.

Etching on the scrutiny of literature of these various innovation avenues, in [100], Folstad categorized living labs based on two facets: i.e., contextualized co-creation and testbed association. Focusing on the latter, the authors in [102, 107] saw traditional testbeds’ extension to living lab environments as a developing trend as far as both technical testing and co-creation research is concerned. However, from a factual point of view, tools that will enable such research in test bed-oriented living labs remain a great need [100].

Albeit technical advancement of mobile computing domains, the latter continue to employ advanced techniques of research in lab testing. There’s a lot that needs to be done as far as the evolution and deployment of the new tools is concerned.

2.8 Beyond State-of-the-Art

Beyond the current 4G LTE/NGMNs, the fifth generation (5G) will become the next innovation of wireless networks, and it is expected to dominate the current generations in the next coming decades.

From the time when the first generation(1G) system was introduced in 1981 by Nordic Mobile Telephone[108], every after 10 years, there has been a release of a new mobile generation. These innovations continue to show interesting and incredible effects on all various wireless communication systems that people continue to use lately. Figure 2-12[109] shows the technology phases that the communication systems have gone through since 2012 and what the future holds.

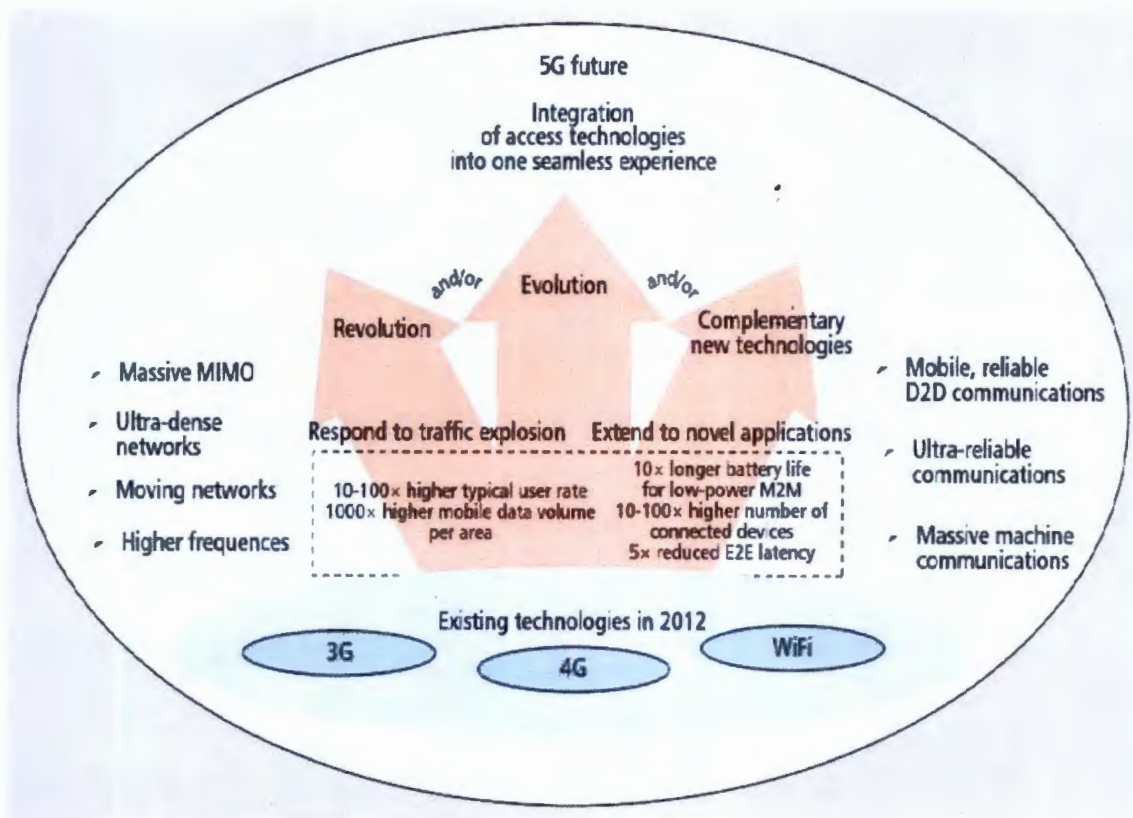


Figure 2-12 New technological advancement towards 5G[109]

The latest 3G and 4G networks are mainly developed with an emphasis on peak rate and improvements in spectral efficiency. As we move towards the 5G era we shall start witnessing dense heterogeneous networks (HetNets) architectures forming the backbone of all 5G systems. HetNets promise to be the most realistic approach to realizing a uniform connectivity user experience and meet the ever growing needs of the telecommunications industry with minimal cost. Figure 2-13 highlights the major developments in the HetNet transformation. The figure also includes small cells and D2D communications o, multiple cell.

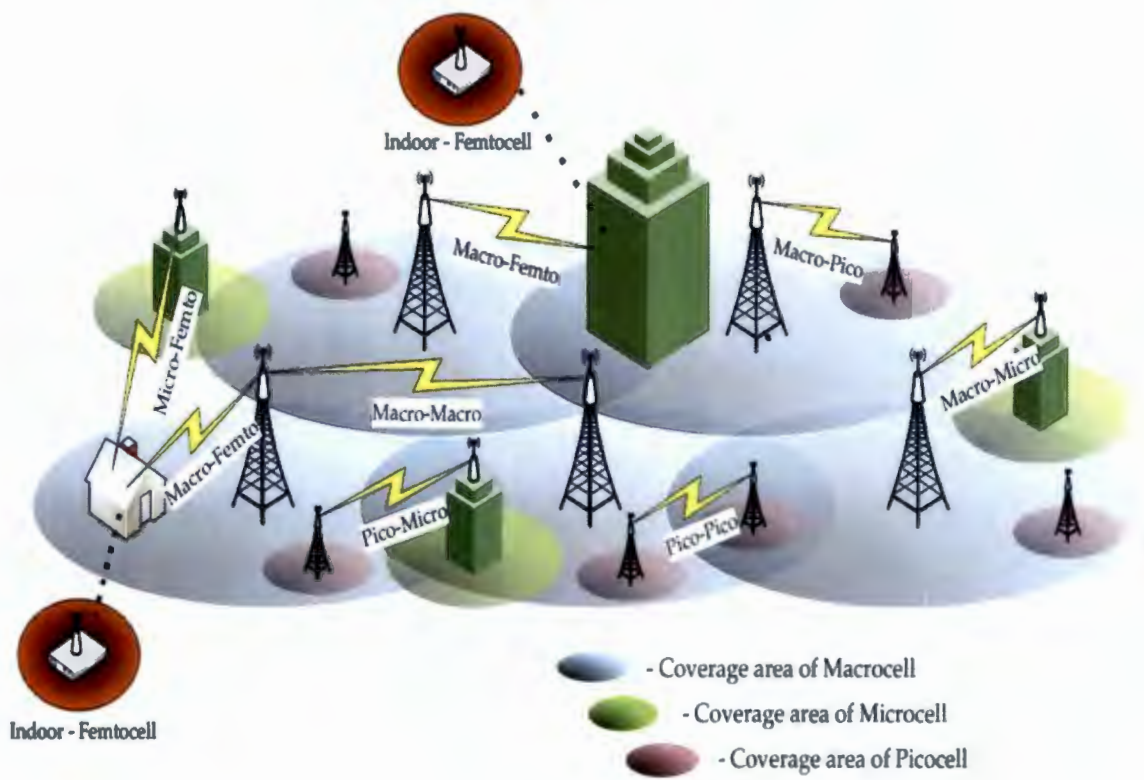


Figure 2-13 HETNETS[110]

HetNets are made up of different categories of small cells which combat the reuse of spectrum resources. HetNets are however to be designed to assimilate the varied set of frequency bands that will range from a variety of network topologies. Small cells will be able to accommodate new higher frequency spectrum such as mm-wave as a way of enabling of ultra-high data rate services. In relation to small cells, user gadgets shall be integral to the 5G network era. Thus both small cells and Device-2-Device (D2D) communication shall be part of the new tier of infrastructure which will complement the deployment of cellular networks while minimizing cost. Minimizing costs will turn out to be an important factor in the 5G systems giving space to better software-based deployments and technologies for virtualization. Basically it will be possible to develop various virtual core networks mainly designed based on the requirements of specific applications. Additionally, when the 5G finally gets here, there will be a need for strong edge nodes to handle traffic offloads coming from the core networks. Firstly, to regulate the efficient flow of data by changing network resources in a dynamic way as a way of ensuring high QoE for the various applications' flows. Secondly, to manage raw information originating from a wide variety of sensors or Internet of Things (IoT) gadgets and equipment. In order to reduce the core network traffic as well as reduce retrieval-latency during peak hours, more content will always be cached on the edge of the network. This process will make it possible for network operators to maximize the use of their network resources through the use of context information.

2.8.1 5G HetNet architecture and enabling technologies.

Societal development will lead to changes in the way mobile and wireless communication systems are used. Essential services such as e-banking, e- learning, and e-health will continue to proliferate and become more mobile. On-demand information and entertainment (e.g., in the form of augmented reality) will progressively be delivered over mobile and wireless communication systems. These developments will lead to an avalanche of mobile and

wireless traffic volume, predicted to increase a thousand-fold over the next decade [111, 112]. Today's dominating scenarios of human-centric communication will, in the future, be complemented by a tremendous increase in the numbers of communicating machines. This so-called Internet of Things (IoT) will make our everyday life more efficient, comfortable, and safe. There are forecasts of a total of 50 billion connected devices by 2020 [113].

The coexistence of human-centric and machine-type applications will lead to a large diversity of communication characteristics. Some of these applications can be supported by today's mobile broadband networks and their future evolution. However, some other applications will impose additional and increasingly diverse requirements on mobile and wireless communication systems that the fifth generation (5G) will have to support: Far more stringent latency and reliability requirements are expected to be necessary to support applications related to health-care, security, logistics, automotive applications and mission-critical control. A wide range of data rates has to be supported, up to multiple gigabits per second, and tens of megabits per second need to be guaranteed with very high availability and reliability. Network scalability and flexibility are required to support a large number of devices with very low complexity and requirements for very long battery lifetimes. One of the main challenges is to satisfy these requirements while at the same time addressing the growing cost pressure. Efficiency and scalability are therefore key design criteria, reflecting the need to respond to the expected explosion of traffic volumes and number of connected devices.

Existing mobile technologies will not be able to provide the capabilities to meet market demands Beyond 2020. Sustained research will be needed to create a high performance 5G environment[114]

Figure 2-14, demonstrates the challenges that 5G is going to offer to the users. These challenges are: capacity (10000 more traffic and also 1-100 more devices); latency (the latency

reduces); energy consumption (10 years battery for M2M / ((J/bit)/10); cost (M2M ultra low cost); user data rates (user data rates will rise) and coverage (100 Mbit/s Low-end data rates. Ultra reliability).

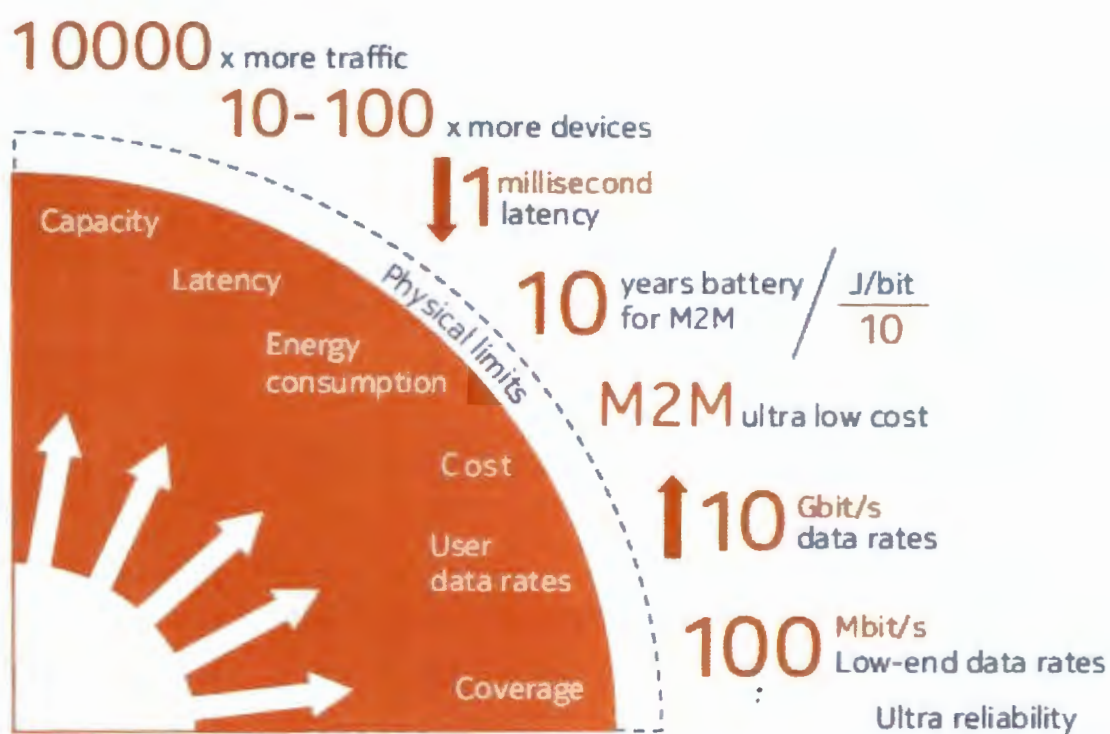


Figure 2-14 challenges that 5G is going to offer to all of the users[115].

5G will enable the long-term networked society and realize the vision of unlimited access to information for anyone and anything. This vision will be achieved by combining evolved versions of today's radio-access technologies (RATs) with complementary RATs for specific use cases, not by replacing existing technologies.

In a nutshell, 5G is going to have a much tougher process behind its development level. The 5th generation will support more connected devices worldwide compared to the generations before and it will also be more efficient compared to them. By making advantage of IP version 6, 5G will have a higher security compared to its previous generation.

In 5G, Network operators will be able to exploit intelligent management of cellular network infrastructures, such as C-RAN and HetNet, to support the increasing number of connected devices and the traffic demand while decreasing their capital and operational costs. These technologies will make it possible to easily upgrade their existing networks and enhance their coverage.

2.9 How mobile networks improve QoS

Different mobile network solutions offer users the capability to improve both network efficiency and Quality of Experience. Mobile networks provide measurable Quality of Experience results. Actual use cases clearly enable operators to provide performance-based Quality of Experience enhancements that:

- a) help reduce lagging and stalling;
- b) provide faster web downloads;
- c) accelerate performance of mobile apps;
- d) promote subscriber roaming;
- e) deliver HD YouTube videos;
- f) mitigate network resource monopolization;

▪ **Stalling of Mobile Video**

Depending on system conditions and time of day, live features such as video streaming slows down between 5% and 35% of the time. Even 4G/LTE systems can experience slowing down and can also get congested. This stalling process can last from seconds to minutes and lead users to close their ongoing sessions, causing loss of interest and dissatisfaction among subscribers, as well as decrease in data consumption and eventually increase in lagging. Activity administration arrangements have demonstrated effectiveness with controlling application movement; however, this does not work well with video streaming because these solutions are normally implemented at the application layer through packet blocking and shaping. Video streaming is based on bit rate which varies its maximum rates that could exceed the shaped bandwidth of traditional traffic management solutions, leading to stalling, clipping and a very low mobile video experience.

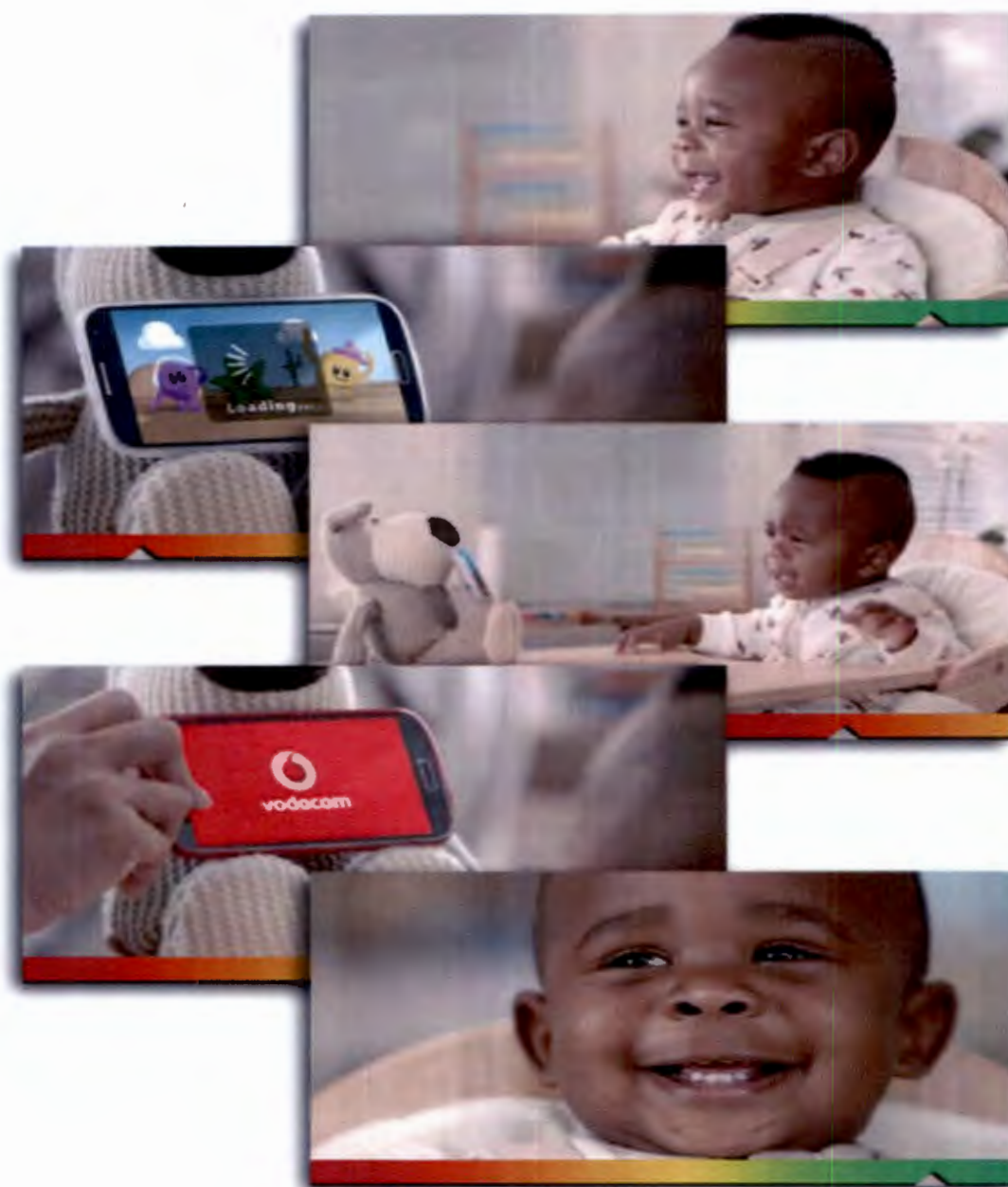


Figure 2-15 User satisfaction is easily affected by mobile video QoE[116]

Figure 2-15 shows a snapshot from an experiment conducted that involved a child engaging with an animated video on a typical smartphone. The child kept on becoming upset each time the video stalled. However, the child started to enjoy and laugh again as soon as the connection was restored. During the long stalls, the child started crying so loudly until the phone was replaced with the one having a better connection, which ran the video seamlessly. This is a clear demonstration of how mobile QoS is seen by the subscriber.

NGMNs reduce video stalling by 40% to 60%. To achieve this NGMNs have applied a number of techniques which helps in reducing the consumed amount of video bandwidth. It also helps create a headroom especially when mobile networks are under huge amount of request burden. NGMNs are implemented with just-in-time video pacing which eliminates the useless and wasted downloads. NGMNs use quality-aware transcoding techniques which are designed to compress videos to certain levels ranging from barely watchable, to nearly watchable, and then to visually apparent. NGMNs also apply dynamic bandwidth when a user's video session is at risk of lagging, stalling, adjusting the compression level in order to match the available network capacity which is available at the time, to make sure that the video being delivered to the user can continue uninterrupted. These capabilities all highly reduce lagging while keeping high QoE as shown in Figure 2-16.

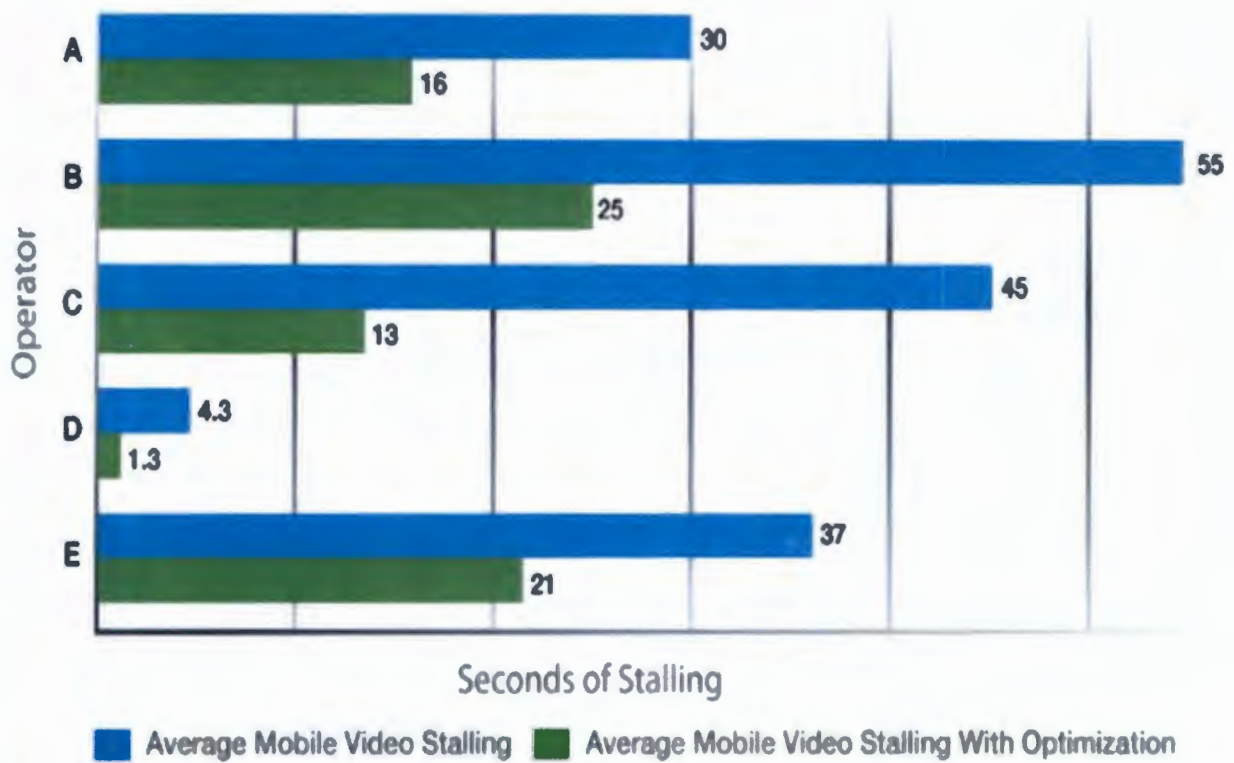


Figure 2-16 Improvement in mobile video QoE across 5 different mobile networks[117]

▪ **Faster Web Downloads**

Different studies related to QoS in NGMNs that were conducted showed that hundreds of mobile web pages are downloaded worldwide across different mobile operators. The results in these studies showed that around 60% of web pages required more than 9 seconds to load and more than 20% of them took 21 seconds or more to load. However smartphone users do not wait 9 seconds and prefer giving up on a download in 8 seconds or less, it became clear that NGMNs should provide faster webpage downloads in order to improve mobile users QoE. NGMNs allow operators to provide faster web downloads. Now that all major mobile applications are web-based, NGMN solutions also increase mobile applications without requiring any further information from the developer or operator. Improving video streaming quality and mobile data quality experience is very important. Simulations were conducted to measure data transfer performance so that cellular networks can efficiently evaluate activities related to NGMN service providers. For example, a benchmark simulation was conducted by P3 Group in 2012 in Germany.

▪ **Quality of Experience Improvement**

Even in a tiny cell network, conditions are constantly fluctuating; while subscribers' experiences fluctuate too, some have good experiences, though some might not be so lucky. Dynamically network changing also affects the subscribers experience related to live video streaming. NGMNs make sure that users face low poor video experience so that they can enjoy a higher Quality of Experiences of video streaming with low video lagging and also improved web pages downloads. Fluctuating network conditions results into differing QoE within the same cell mobile data usage when improved data usage also increases along with video streaming Quality of experience. For example, an operator observed a 29% increase in web browsing after Next Generation Mobile Networks optimization was applied.



Figure 2-17 Improved Quality of Experience with Smart Capacity Optimization[117]

- **Manages Quota Breaches**

Quota breaches should effectively be managed in order to achieve an improved Quality of Experience for smartphone subscribers. Optimization is the most important technique used to handle quota breaches in different ways. For example, using shaping techniques or aggressive blocking leads to unviable mobile video streaming. Multiple companies especially in Europe have been offering continued video streaming consumption to their users who exceed their monthly data usage limits. A message is displayed on subscriber's laptops and smartphones which displays their data usage. It also displays upgrade of monthly data plans, while consumers continue to consume their data. Once the limit is exceeded, any further videos delivered to the subscriber are optimized.

- **Subscriber Usage While Roaming**

Users ordinarily consume less video while browsing. This might be because of overbill shock, exceeding limits, poor down speeds and poor live video streaming quality. This can result in decreased Quality of Experience for the subscriber.

- **Monopolization of Network Resources**

Feature content suppliers, for example, Hulu, Netflix, Apple and others utilize Adaptive Bit Rate (ABR) protocols used for streaming for prime users. There are a few executions of ABR protocols from Apple, Adobe and Microsoft and all share a typical methodology. Adequately, they adjust continuously to network output and gadget CPU abilities by alterably expanding or diminishing the determination of the feature.

A key conduct of Adaptive Bit Rate gushing conventions is that they expend as much accessible system limit as the substance server can use and convey the HQ that the present system association can utilize for a live stream.

- **Quality-Aware Optimization**

HD features drive a lopsided rate of general remote system activity. Case in point, iPhone supporters ask for high-determination features 27% of the time; in any case, these high-determination features create just about 80% of the portable feature activity on the system.

Remote systems that backing higher-determination features convey a superior visual QoE to their supporters, however these features must be viably upgraded to guarantee that general endorser QoE is not contrarily affected.

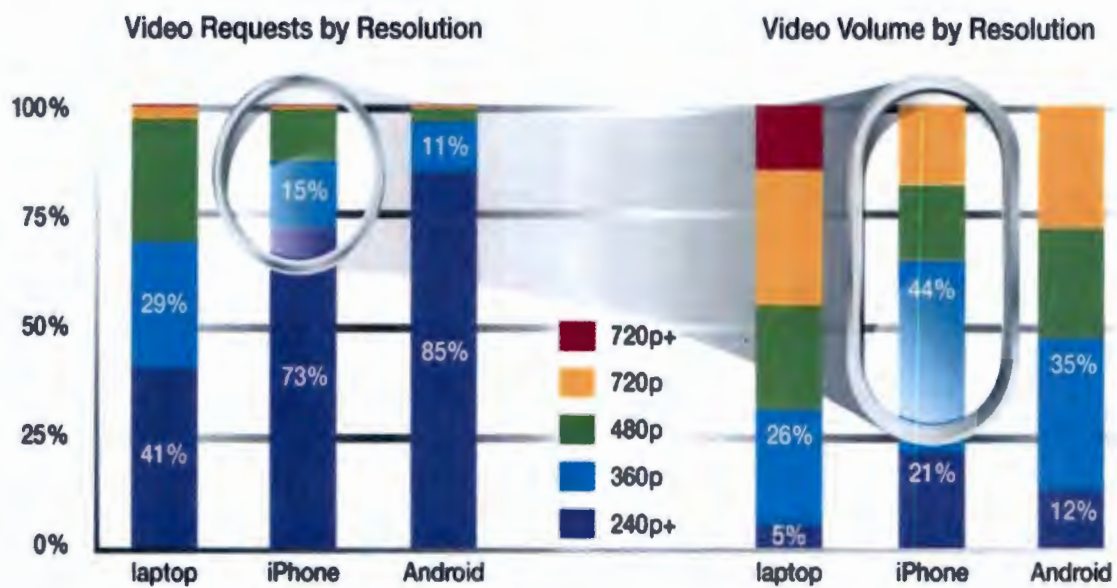


Figure 2-18 Higher resolution videos create the most traffic volume on networks[117]

Numerous 'right now accessible streamlining systems and advanced features' aimlessly result into either excessively traditionalist executions that leave space for extra enhancement or excessively forceful improvement that makes decreased feature quality. This translates into poor support for QoE. Cutting edge Mobile networks apply feature advancement that is outwardly lossless, giving the ideal Quality of Experience to every supporter. Improvement is connected taking into account numerous variables, for example, the composition of the first feature and the codec being used. As opposed to just utilizing pressure, NGMNs apply refined advancement methods to boost the endorser experience; for instance, evacuating casings that can't be seen or specifically decreasing feature quality for distinctive parts of a frame. The NGMNs give an industry-driving versatile administration activity and enhancement innovation that allows administrators to increase users Quality of Experience and amplify system proficiency. Impressive data bundles and deals with all information activity in an administrator's system along with other features such as coordinated application recognition, versatile activity administration and streamlining administrations change in accordance with system conditions progressively. This guarantees the best Quality of Experience conceivable. Therefore, operators can help versatile information utilization, separate their systems, draw in new clients and decrease churn.

2.9.1 QoE and the Techno Centric Approach

The adjustment and evaluation of network parameters in the user techno centric approach provide the description and determination relationship between the telephone circuit loads and the equivalent traffic delays within the network, it also provides the ease with which delays relate with the service quality from the consumer viewpoint.

According to Pocock's work on QoS [118], the monetary impact of the administrative quality provided by the telecom administrators endeavors the discourse about the general quality that the user experiences focusing only on the essential changes in quality transmission.

The advances in research provide the evaluation of the network QOE using the QOS metrics, feedback of the network and the requirements of the users. Siller and Woods in [119] mention that QoE has the effect on the application line layer over the user's observation. QoE is the client's apparent experience with what the application layer displays, where the application layer is the front end client interface that shows the general effect of the individual quality of services.

The advancement of QoS directly links the application layer with the QoE evaluation layer.

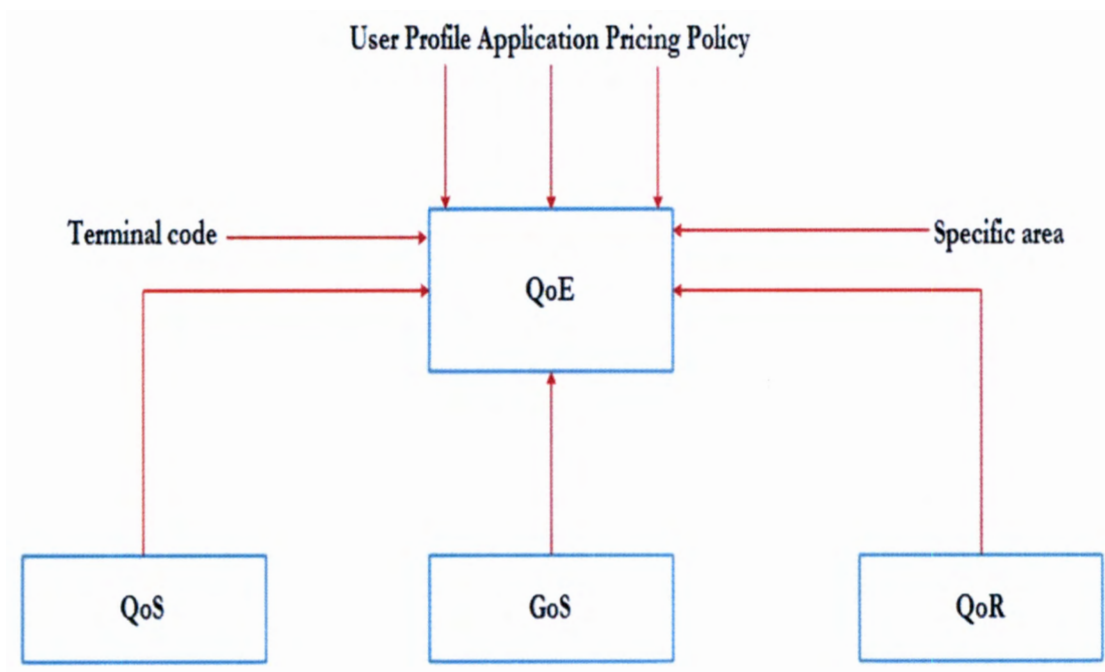


Figure 2-19 Elements that influence QoE

2.9.2 QoE and the User Centric Approach

According to the ITU, QoE describes the general acceptability of an application or administration, as subjectively seen by the end-client. It usually includes the broad end to end system effects such as the terminal, client, services, network, and infrastructure. On the other hand, QoS describes the totality of qualities of information transfer benefit that bears on its capacity to fulfill expressed and suggested needs of the client of the administration. The QoE provides the effect of application and services for the user to also consider the external factors as seen in Figure 2-19.

▪ **Difference between User Centric and Techno Centric Approach**

The differences between the techno centric and user centric approaches are considered as good forecasting for the user’s perception of a service. A summary of how these approaches differ is presented in Table 2-3.

Table 2-3 Difference between User Centric and Techno Centric Approach

Techno Centric Approach	User Centric Approach
a) Evaluation is totally based on the performance of the network.	a) Evaluation is totally based on human factors.
b) Considered as one dimension for the consumer's observation evaluation, i.e. based on low, fair, good and excellent perceptions.	b) Has interdisciplinary attitudes incorporating the elements like fun , beauty etc.
c) Users' observations are mainly on results of traffic load, scheduling and routing techniques.	c) Users' observations are based on the thoughts and needs of the human.
d) Has a large interest on laboratory simulation results and network performance monitoring.	d) Has major interest based on the practical and field studies that gather the natural and real user data.

2.10 Techno centric approach in advanced mobile networks.

The current advancement in mobile broadband systems must have the capacity to support several audio and video usages as well as its single IP based network framework information. In this way mobile service providers hope to offer separated versatile administration bundles. One methodology to recognize the activity from distinctive applications is to offer to a suitable QoS scope for the measurement of network service quality. This strategy is intended to guarantee the quality of the application and permit the providers to offer various administrations to clients and oversee system clogging. Thus, 3GPP built up QoS, charging and arrangement control structure to make a productive apportioning of the accessible remote system properties.

The major goal in providing the 3GPP services is to effectively access the indefinite network control framework policies with the major objective of regulating the service quality and policy mechanisms that can provide the services and subscriber variations to support the network operators effectively. The 3GPP standards mainly disclose how to construct the transmission ways between the client equipment and the external information system packets with all characterization of QoS. Recently, 3GPP services have been characterized broadly as the "carrier model" to execute the network QoS. The carrier division can be effectively predicted by the 3GPP network. A transmission is an activity detachment component that empowers separated treatment of movement in view of its QoS necessities, and gives a legitimate way between the portable terminal or client hardware and a door. Every network carrier is always linked with QoS parameters that depict the properties of the vehicle channel, including the bundle delay, packet misfortune, bit rates, planning strategy and bit slip rate in the radio base station. All streams mapped to a solitary carrier get the same package sending treatment e.g. booking, line administration, rate forming, connection layer arrangement and so forth. The data

modulator has two to four QoS parameters, contingent upon whether it is constant or best exertion administration.

2.11 Quality of Service Class Pointer

The QoS Class Indicator (QCI) provides the data sending action i.e. weighted data planning, data validation limits, data line management and data connection layer setup that the client usually plans actively gets between the UE and the PDN Gateway passage.

2.12 Allocation and Retention Priority (ARP)

In 3GPP standards, the allocation and retention priority method comprises of the information related to its significance level. The anticipation capability and the prevention susceptibility provide the priority level that has comparative importance with respect to the appeal resources. This permits determining whether the carrier formation or adjustment request can be acknowledged or the requirements can be precluded during the resource limitations process. The ARP parameters for mobile subscribers are very important because according to this the evaluation and congestion of the mobile subscriber occur.

2.13 Guaranteed Bit Rate (GBR)

The guaranteed bit rate is defined as the minimum bit rate that is expected to be available to the PDP (Packet Data Protocol) context when required[120]. The proper implementation of GBR provides the network with reliability that no data packet will be lost during the IP network and radio link transmission due to network congestion. The GBR carriers are used to provide less latency and jitter tolerances for the real time services. The GBR also aids the consumer with the security of the data packet so that if the data is missed during the transmission, then it can be recovered by detecting the data packet. If there is any noise during the transmission then we can use the repeater or filter by considering its magnitude level and suppress the noise.

2.14 Maximum Bit Rate (MBR)

MBR is used to make sure that the data traffic does not exceed the network data rate. The transmission data rate does not exceed the network rate. The MBR is usually appropriate for the real time services. The MBR has to be equal to GBR for efficient real time services. The information regarding the controlling of the network operator should require that the data packets flow be mapped onto a dedicated network carrier. At the receiver end, the quality of service level should also have to be determined through the network policies. This combination process usually occurs at the UE (User Equipment) at the packet data network gateway (PDN-GW) layer by using the traffic flow templates (TFT). The traffic flow templates usually comprise of the information about the filtering of data packet for mapping and identification of the user data packets for the particular bearers. The gateway implemented policy controller should have to adopt to the provision of network strategies and resource charging function. These are defined as network policy controllers and are usually executed on the gateway level.

2.15 Parameters for the Network Policy Controller

The network policy controller comprises five parameters for the filtration of data packet flows which are also defined as the Internet Protocol Stacks. These are:

- a) source IP address;
- b) destination IP address;
- c) source Port number;
- d) destination Port number;
- e) Network identification protocols (UDP or TCP).

The network policy controller parameters mainly control the network QoS necessities by taking in consideration the three network criteria. namely:

- a) Application requirement;
- b) Subscription information;

c) Network operator policies.

According to the network policies, controlling the decision for dedicated network for data transmission can be established. The periodic service level policies should have to adapt accordingly with the policy decisions for the existing information network and operator designed policy decisions. These policy control regulations are usually developed for the policy and charging enforcement function (PCEF) located in the gateway. The PCEF enforces the policy decisions by the founding bearers, mapping the data flows service for bearers and execution of traffic policing and shaping.



The transport network mainly carries the gateway maps. These are based on the Ethernet technology and use the multiprotocol label switching network (MPLS). The MPLS protocol is defined as the accessible, independent network transport protocol with labeling feature for data packets. The labeling of data packets allows the tracing of data and if the data packet is missed during its transmission it can be recovered easily. Therefore, it can provide the data reliability as well as security. The decision for data packet forwarding is exclusively based on the labeling contents, it does not require the examination of the data packet itself.

The gateway mapping technique mainly carries the core transport network. The transport network can be either Ethernet or MPLS network depending on the network protocol. The network transport protocols usually has no information about the carrier and requires the use of standard IP based QoS techniques like the differentiated services. The differentiated services (DiffServ) is based on the networking architecture of the computer that requires a modest, accessible and highly integrated mechanism for organizing and handling of the network traffic. It is also capable of delivering the essential QoS on the moder IP based networks.

2.16 Uplink and Downlink Base Station Policies

The data transmission from ground to base transceiver station (BTS) is known as the Uplink while that from BTS to the ground station is known as the downlink. This two way data rate communication policy requires the scheduling of the radio frequency (RF) resources. The allocation of bearer resources requires the address resolution protocol (ARP) to resolve the network layer addresses into the data link layer to provide the multiple network access. The RF resource scheduling algorithms for base station can provide the effective influence on the service, quality and the total network performance. The IP network transportation or User Equipment (UE) Protocol can also provide the initial data mapping service from ground station to base station. This implementation of the data network function requires end to end network QoS for data processing.

2.17 Identification and Designing of QoS Networks

The identification and marking of certain data traffic, requires the transmission through data packets. This is essential for the end to end synchronization of QoS through the network. The effective method to perform the identification of data flows through the network contains the data routing tables which depend on the policy based routing tables which are defined by routing and data access control lists. The following network policies spell out the characteristics of data flow through the network.

2.17.1 Policing

The policing function is used to ensure that the data flow through the network should not exceed the agreed bit rate of the network. The network policing method provides the guarantee to share the minimum network services. It decreases the network traffic during the peak time and also provides for the efficient use of network resources by using the time and frequency division multiplexing, in which data packets are transmitted precisely to use the overall bandwidth of the network efficiently. The network policies also provide the delayed or dropping capability

for data in case of non-acquiescent traffic. In mobile network the transmission network bandwidth capability is usually based on the customer's quantity within the cluster. If an anonymous customer enters in the region then either the network resource roaming, data delayed function or dropping function is performed.

2.17.2 Traffic modelling and management.

The traffic modeling and management function provides the bandwidth management of the network. It provides the mission critical based use of network bandwidth to accommodate the external customer into the network cluster during the peak use of network resource. Traffic management can provides the overall bandwidth management of the network so that the exchanged bandwidth necessities for the different network services can be satisfied for all the users. It can provide the assurance of guaranteed bandwidth for VoIP and control traffic applications with the requirement of minimum bandwidth for data packet transmission.

Traffic management algorithms also provide the early random weighted detection and random early detection features for the use of congestion observation. The function can provide the network shaping function for the limited use for the network bandwidth flow and also avoiding overflow situations.

2.17.3 Administration of QoS policy and management

The administration of QoS policy provides the functions for accounting, provisioning as well as algorithms in the domain of network administration, operations and management.

2.17.4 Data planning and scheduling

The data planning and scheduling functions are performed to confirm the dropping of data traffic in heavily oversubscribed events. The delivery of data is usually based on the IP network that is usually supported over the Ethernet core in the LTE networks. The differentiated services protocol's Ethernet feature is usually used to offer the best network quality of service. The differentiated services protocol is applicable for up to six data queues and the IP traffic is

mapped for the differentiated services classes. Differentiated services classes can also be controlled through the differentiated services code point (DSCP) field in the IP layer of the network in the channel headers. According to LG Martinez Ballesteros *et. al.* in [120], the data planning and scheduling feature is in most cases seen in a similar fashion with the particular priority based network without the following features: data avoidance, weighted round robin, weighted fair queuing and deficit weighted round robin or smoothed deficit weighted round robin. The authors in [120] go on to emphasize that these functions have been made more dominant by the techno centric analyses of network QoE. This maintains network optimization of QoS, which in the long run positively impacts user QoE. Despite this positive impact, certain cases exist where perfect QoS is implemented but poor network QoE remains prevalent.

Overall, QoE in networks is mainly evaluated with reference to technical QoS metrics. There is negligence of the main aim of the network which is to deliver a perfect service that intends to yield the best user QoE other than delivering state of the art user applications with innovative architecture and designs. This gap in network design and management comes as a result of the failure to acknowledge the importance of users during network frameworks design and implementation. Luckily, there is now recognizable progress in the creation of HCI tools that are mainly focused on the user and application interfaces. All these is to foster QoE mechanism designs and implementations that are able to involve users during the design processes.

2.18 Interrelating network QoS to User QoE

Measuring of network quality previously has always been done objectively- a phenomenon known as QoS of a network. QoS involves the various ways a network achieves a more desirable and expected behavior, i.e., transportation of data with minimal bandwidth, delay and packet loss. However, it is a mistake not to note that QoS as a whole fails to factor in the perceptions of users. QoE is the other phenomenon that indeed puts user perceptions into context. QoS involves subjectively factoring in human dimensions into network management.

This is done through considering perceptions of users, their experiences and expectations as far as an application and its performance is concerned. QoE is turning out to be vital research niche area. This is because, users are negatively affected by low quality services and during this time, service providers have little or no time at all to wait and address user quality complaints.

In [121], Accenture carried a survey that found out that close to 90% of end users of a service may not necessarily moan about the low quality in using the service, they straight away look for another service provider. Service providers therefore should have a criteria of monitoring and assessing QoE and immediately better on it.

Video streaming QoS over NGMNs is determined by a number of related parameters. Among these parameters, video resolution and bandwidth are easily adjustable unlike delay and packet loss yet these are the parameters that have to be addressed in order to achieve user satisfaction. But again, satisfying a user doesn't only involve getting QoS parameters right but also QoE subjective factors such as the experience the user gets, their interest, and what they expect to get have to be considered as well.

2.19 Quality of Service and Quality of Experience Layers

The research community in computer networking has proposed a number of possible QoS solutions that can be implemented within the seven layers of the Open Systems Interconnection (OSI) model[71].

The application of the OSI mainly deals with application services that implement QoS. QoS at this layer is responsible for handling different parameters like resolution, audio and video codec types and frame rates among others. Furthermore, the services on the network layer are administered by routers and switches. In terms of parameters, the network layer handles packet loss, delay and jitter, etc. The authors in [71] recommend a visceral virtual layer above these

two layers that is envisaged to mainly focus on QoE. The other scholars in [122] think this virtual layer should be an addendum to the application layer, yet the others in [123] envisage QoE as an addition of the traditional QoS since QoE is all about acquiring user feedback on consumed services.

Human perception is considered the main driver of QoS at the application layer. From a user point of view as far as video services are concerned, there are two characteristics that have to be looked at; a temporal viewpoint and a spatial viewpoint. When it comes to achieving compression during video coding, Entropy, Interframe and Intraframe are the techniques that are used. QoS at layer three can be categorized as either resource reservation or prioritization QoS. There are various ways that QoS can be implemented at the network layer, these include: Multi-Protocol Label Switching (MPLS)[124], Integrated Services [125] and Differentiated Services (DiffServ)[126].

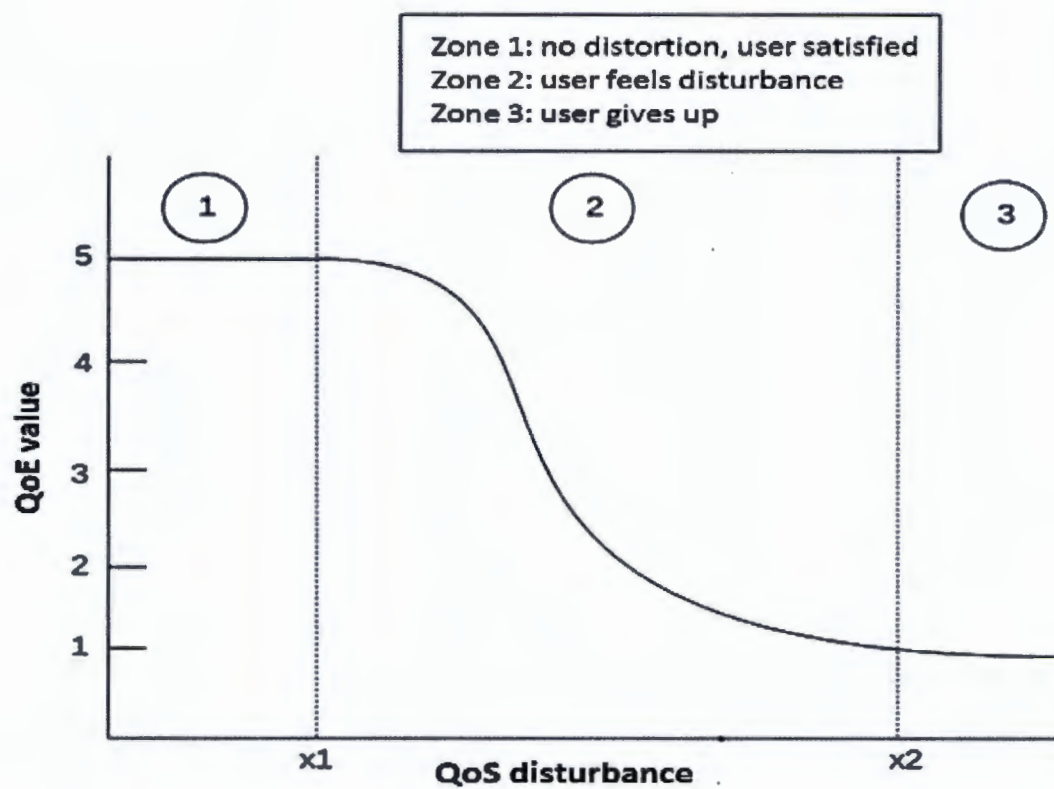


Figure 2-20 QoE and QoS Mapping Curve[127]

In Figure 2-20 above, a delineative depiction of how QoS and QoE are related is presented. The Figure 2-20 is apportioned based on three zones. If there is an interruption or disturbance in QoS and the latter falls below the first zone, a high value of QoE is expected, thus the user's experience isn't affected in any way. If and when an interruption in QoS drops to zone 2, then QoE will drop too. Lastly, when a disturbance or interruption reaches the third zone, QoE is bound to fall, thus a user of the service in question may negatively be affected and they may decide not to continue using the service. Therefore, in most cases whenever there is an increase in disturbances or interruptions on QoS, QoE is bound to be affected and the way a user perceives the quality of a service is affected negatively.

2.20 Ways of Measuring QoE

Measurement of QoE can either be subjective or objective as previously discussed. The process of QoE measurement especially in video applications mainly leans towards the subjective approach. The mean opinion score (MOS) in this case is the most recognized method for subjectively measuring QoE. ITU-T recommendations in [128] standardize MOS, ITU-T in its recommendations defines MOS as the arithmetic value that starts from 1 to signify "poor" up to 5 that signifies "excellent". However, this method of measuring QoE comes with a couple of challenges such as, time wasting, difficulty of implementation in real time scenarios and it is very costly. These challenges have motivated developers and researchers to come up with objective solutions, mechanisms and tools that can still also use MOS in order to envision subjective quality entirely from the physical attributes.

The objective way of measuring QoE solely depends on the use of comparative and mathematical techniques or algorithms that are clearly meant to bring about a quantitative dimension in assessing 'one-way' video quality. This objective approach is thus very important in terms of monitoring service quality, and also optimizing and selecting codecs. Objectivity

can either be intrusive thus it involves use of signals or non-intrusive that involves the use of application or network parameters. This means that intrusive objectivity is more accurate though impregnable in terms of live or real traffic monitoring because of the need of the original sequence of the traffic from the source. Contrary to this, non-intrusive objectivity seldom needs the original sequence of traffic. Therefore objective approaches under ordinary conditions have the ability to overlook the type of content and the way humans are bound to feel [54]. To give an example, there are various objective approaches and solutions that can analyze pixel by pixel and the original signals sent and those that are received to reveal distortions in the signal like Peak Signal to Noise Ratio (PSNR)[51].

An amalgamation of both subjective and objective solutions can be implemented to solve the challenges each solution is bound to have. In terms of the network conditions that affect video streaming QoE, mapping PSNR-to-MOS is the most used mechanism. However, scholars in[129, 130] deemed the PSNR-to-MOS mechanism defective as far as linking the latter to the perceived visual quality is concerned. With these findings, a number of changes were suggested to boost the accuracy of the estimates. In [57] and [131] the improved PSNR-to-MOS mechanism was fully accredited.

Peter and Bjorn in [121], categorized the various solutions of assessing the network services from the user's point of view in 3 ways. As shown in Figure 4-2, these include, Surveying Subjective QoE (SSQ), Testing User Perceived QoS (TUQ) and Modelling Media Quality (MMQ). SSQ and TUQ mainly focus on gathering subjective feedback from the users while MMQ deals with the objective and technical metrics of the network. Figure 2-21 is a summarized analysis of the discussed ways of assessing network quality of service.

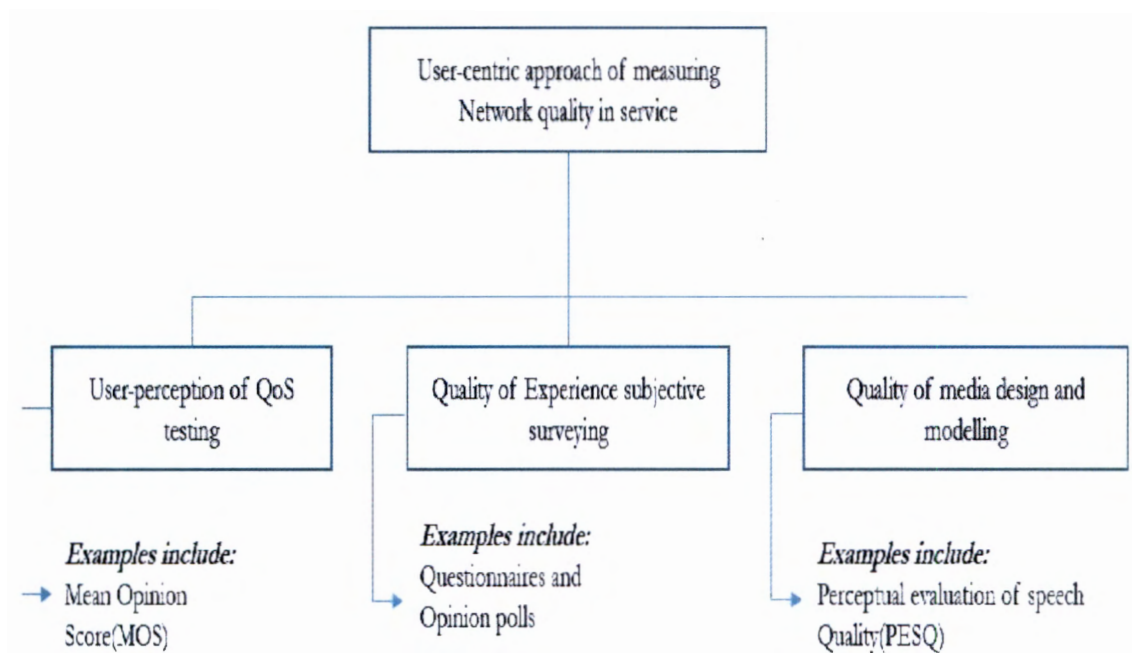


Figure 2-21 User-centric ways of measuring the quality of network services

2.21 Objective Quality Assessment Modelling

The investigation of user centric QoE versus network centric QoS is still of paramount importance. This observation leads the research community to continue pursuing and studying the various ways of how QoS metrics of a network affect the numerical sections of QoE. It is indeed probable to assess and measure QoE and afterwards develop a relationship of network QoS parameters with QoE metrics. Further on, it becomes a realistic target to design and model QoE-aware QoS mechanisms. Because of this progress, a variety of mechanisms have been designed to estimate QoE. In [132], ITU standardized the mechanisms in the form of specific classifications relying mainly on the individual model type. Much as most research work on QoE focuses on assessing and monitoring of QoE [57-60], Management and evaluation of QoE has now picked up more interest from different researchers around the world. So, various systems and frameworks meant to evaluate and manage QoE can be noticed in these research works. Kafetzakis *et al* in [61] presented a multifaceted framework known as QoE4CLOUD, their framework was meant for cloud computing based networks. The various challenges of managing cloud applications QoE are well presented by Hobfeld *et al* in [62]. As far as managing QoE in NGMNs is concerned, Gómez *et al.* [63] presented a network architecture that implemented QoE management. Fajardo *et al* in [64], proposed an administration system for QoE with regards to VoIP in 3G UMTS offerings. All in all, with the ever evolving technological networking trends in mobility, heterogeneousness of networks and network independent mechanisms and solutions remain inevitable. Mu *et al* in [65] goes on to propose an end-to-end multimedia applications QoE management mechanism they named QoE2M(QoE- aware Real-time Multimedia Management). This mechanism involves controlling the assessment of video and QoS/QoE mapping procedures. The mechanism basically detects network congestion times by using resource allocation controls that depend on collecting network status statistics. In a nutshell the QoE2M mechanism can quickly adjust

applications based on the users' gadget means or network status. Much as Mu *et al.* presented a more vigorous way and bringing in the application's QoE features, they fail to present implementation or validation issues [66].

Vakili and Grégoire in [67] introduced a QoE mechanism for H264 video. Vakili and Grégoire make use of subjective simulations in assessing perceived video quality by the end users as a way of finding answers for the various effects of frame rate and quantization on perceived video and video data bit rate. The authors also go on to present a direct way of assessing the least transitional hop's bandwidth on each path at the same time continue to forward video traffic. The only limitation of this mechanism is that it restricts the type of video and size of the videos.

Kim *et al* in [68] introduces an "in-service feedback QoE mechanism". With this mechanism, feedback from end users is received immediately there's a bad experience with the service. Basically, the collected feedback data from the poorly satisfied end-users, network routers, and data from application servers is easily analyzed to trace the errors. etc. But again, just like with the QoE2M mechanism in [66], there was no implementation and validation of the findings.

Agboma and Liotta in [31] went on to propose their QoE mechanism with different QoE models for different kinds of multimedia content for basic mobile gadgets. Further on, this mechanism focused on the use of statistical mapping QoS parameters to user perception. In this approach, the vestibule where the service quality in the eyes of the user starts to decline is well defined. The downside of this is that, Agboma and Liotta only focused on ways to maximize end-user quality and they do not consider the impact of the network or how to optimize the usage of network resources [31].

An IPTV QoE mechanism was presented by Lloret *et al* in [69]. This mechanism was implemented between the end user's network and the Internet Service Provider (ISP) architecture; this mechanism further implements a "process" to collect information from the two architectures. The main purpose of this collected information is to execute the necessary roles within the network to avoid unnecessary network congestion hence avoiding loss of packets. Although this is a significant progress, it doesn't consider the effect on the network neither does it maximize network resources, this is because , the mechanism in charge of admission control to deliver the right IPTV service to end users of all the channels doesn't put this consideration into context.

Despite a lot of research work focusing on presenting QoE/QoS mechanisms to manage QoE, literature still exists on scheduling, bandwidth management, admission control and congestion control. For example, Piamrat *et al.* in [70] presents a QoE-driven admission control framework for IEEE wireless networks. In this work, an average Mean Opinion Score (MOS) on all live connections are calculated. If the result of the MOS is greater than the desired level, then another connection can be allowed, otherwise no connection is allowed.

Another QoE-mechanism was presented by Gallo *et al* in [71]. The main components of this mechanism include the application. user and QoS aspects of the network. The results from Gallo *et al* indicated that their mechanism was able to map QoE to QoS and thus benefit QoE management. However, the mechanism doesn't evaluate QoE since it only focuses on mapping QoE to QoS parameters, leading to the QoE management implementations gap currently in networking research, which is failure to implement and validate the mechanism.

Once again, in this thesis we introduce an independent mechanism that is compatible enough with other technologies. The mechanism monitors and evaluates QoE trends with reference to user perceptions. Using both subjective lab environment simulations and simulated

environment testing is key in making sure the mechanism is implemented and validated, an aspect that different research works as discussed above weren't able to perform.

The next section presents the five models used in accessing objective network quality.

2.21.1 Bit-stream model

The Bit-stream model is a new model that bridges the gap between the media-layer and the packet layer models. This model uses the various features of a bit-stream to derive quality through extracting and analyzing the bit-stream codes.

2.21.2 Parametric planning model

This model uses the network metrics plans as the input, however for this model to successfully work, it needs to have all the previous information about the system being tested.

2.21.3 Parametric packet-layer model

The parametric packet-layer model uses packet header information to anticipate the probable effects on QoE. The model doesn't require the media signal to do this. Thus, the model doesn't pay attention to the information payload. The main short-coming of this is that the model may experience a sizeable amount of challenges in evaluating QoE of any content.

2.21.4 Hybrid model

This model uses a mixture of the other models for QoE prediction. It is considered to be quite effective since it uses a lot of information during QoE prediction.

2.21.5 Media layer model

The media layer model depends on the analysis of the signals of media by the help of the Human Visual System. However, the absence of media signals can easily render this model irrelevant for use.

2.22 Mapping subjective scores to objective indices

Mapping subjective scores to objective scores is one in which a relationship between QoE and QoS can be established. There are mainly two ways in which QoE can be mapped to QoS. These are: the bottom-up method and top-down method. The bottom-up method is when mapping is done starting from the network itself (QoS) and the top-down method involves beginning the mapping from the user's side (QoE). These two methods can therefore be merged as shown in Figure 2-22.

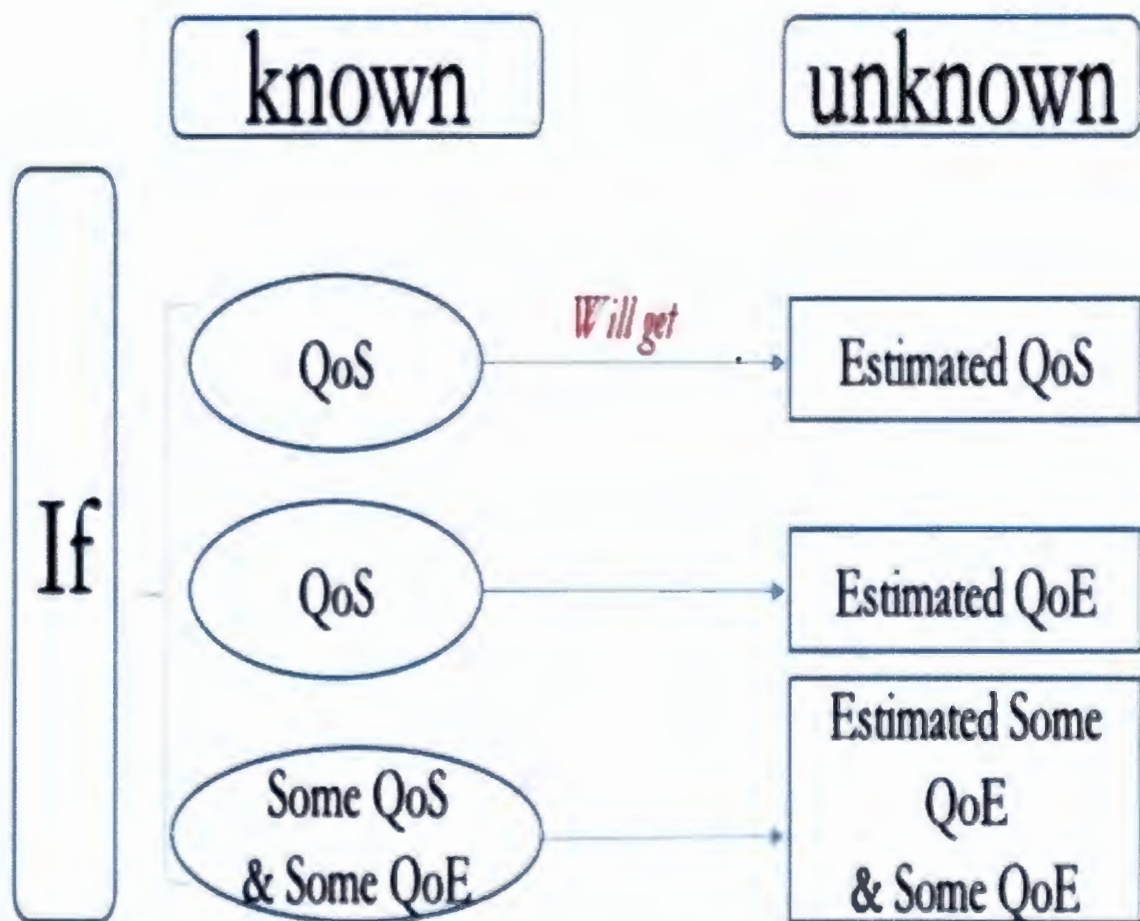


Figure 2-22. The methods of mapping QoE to QoS

As far as mapping the parameters is concerned, models that will always use optimization techniques will always outperform the other models. The literature to be reviewed in this section will prove this. Another aspect to be understood is that mapping functions become inevitable whenever objective quality of video (VQ) has to be linked to subjective QoE mean opinion score (MOSp). This is mainly because the ranges of objective indices of quality and the subjective QoE scores are always different. The process of mapping VQ to MOSp can either be linear or non-linear.

2.22.1 Linear mapping function

This function is only used when the objective indices of quality and subjective QoE scores are uniformly scaled; thus a result of the same arithmetic difference correlates to the equally perceived difference in quality above the aggregate range [133].

$$M_{\text{p}} = a_1 + a_2 * V$$

Where VQ is the objective quality of video;

MOSp is QoE mean opinion score;

a1 and a2 are parameters. (2)

In equation 2, a linear point in between the VQ values and MOS scores is used to derive the parameters a1 and a2. This makes it possible to assess the objective metric. The MOSp results can now be realized, and lastly the scores that are predicted can also be linked to authentic scores. However, the uniformity of the objective quality scales remains questionable. Because of this cloud of doubt, the linear mapping function casts a despondent reflection on quality performance.

2.22.2 Non-linear mapping functions

In order to overcome the setback presented on the linear mapping functions in Section 4.6.1 above, non-linear mapping functions are used. Basically, the latter produces more

interrelations compared to the linear mapping functions. The research community and network developers currently make use of the following mapping functions;

- i. Logistic functions represented as shown in Equation 3:

$$M_{Sp} = \frac{a1}{1+e^{-a2*(V-a3)}} \quad (3)$$

- ii. Cubic functions represented as shown in Equation 4:

$$M_{Sp} = a1 + a2 * V + a3 * VQ^3 \quad (4)$$

- iii. Exponential functions represented as shown in Equation 5:

$$M_{Sp} = e^{(a2 * V)} * a1 + \exp(a4 * V) * a3 \quad (5)$$

- iv. Logarithmic functions represented as shown in Equation 6:

$$M_{Sp} = a1 - a2 |\log(V)| \quad (6)$$

- v. Power functions as shown in Equation 7:

$$M_{Sp} = a1 * VQ^{a2} + a3 \quad (7)$$

The various mapping functions compare quite well with the different measurements of QoE and QoS [18].

2.23 Different approaches in interrelating QoS to QoE

In order to answer research question 3 (RQ3) and realize the research objective 3 (RO3) of this work, and to generally achieve a further broader understanding of our research topic, we identified a few modelling approaches from the various works of other scholars. This section therefore analyses and reviews these approaches.

2.23.1 VQM-BASED Mapping Model

In [127], the authors did suggest a multidimensional hierarchy of mapping QoS to QoE; in this approach, Video Quality Metric(VQM) represented QoE [134]. The authors in this work

present VQM as a function of QoS in n-dimensions, where n in this case represents the number of the various parameters of QoS and it is implemented using Equation 8:

$$VQM = f(QoS_1, QoS_2, \dots, QoS_n) \text{ (where } QoS \text{ is QoS parameter)}. \quad (8)$$

Delay, packet loss and jitter were the three parameters of QoS that were considered in the work. The VQM scores were produced using the NTIA General Model [135]. In order to forecast the QoE of streaming video from the QoS conditions, a formula was derived using the analyzed emulation results. With the help of the curve fitting tool known as IstOpt [127] in which $n = 2$, Equation 8 was derived to give the mathematical expression of the VQM function. Equation 9 generated the final result.

$$VQM = \frac{P1 + P2 * X1 + P3 * X1^2 + P4 * X2 + P5 * X2^2}{1 + P6 * X1 + P7 * X2 + P9 * X2^2} \text{ s.t } 0 \leq X1 \leq 5, 0 \leq X2 \leq 5 \quad (9)$$

In the Equation 9, $X1$ represented jitter and $X2$ represented the percentage of packet loss. Four video samples were used to obtain the VQM score variation with the packet loss ratio. The results showed that packet loss increased along with the VQM score. This portrayed an occurrence of negative QoE. Likewise, the four video samples had their VQM score dependence on jitter and delay recorded. Equation 9 was verified using the other emulation results coming from the other samples of video. Nonetheless, question marks remain on Equation 9. This is because there is no sufficient proof that QoS-to-QoE mapping takes place. It is because of this finding that further conglomerate curve fitting solutions are still required. The study mainly concentrated on the network layer QoS parameters.

The authors in [28] did also suggest looking at the interrelation between QoS and VQM as a quality metric, of which it is also seen as a QoE indicator.

2.23.2 Statistical Analysis QoE Modelling

The authors in [136] made an effort to cut down on the demand to carry out subjective studies by suggesting and presenting a model that would greatly rely on limited subjective testing. In this study the authors issued different users with a similar video either in ascending or descending quality order. By using the method of limits, [137] each user/viewer identified the stage where a distortion in quality was worth noting. To forecast group memberships out of a set of quantitative variables, discriminate analysis(DA)[138] was used. By using arithmetic equations and derivations, group memberships were separated. The equations that were derived are termed as discriminant functions, these are mainly used for purposes of prediction. Equation 10 gives the general function formula:

$$f_k = u_0 + u_1 + X_{1k} + u_2 X_{2k} + \dots + u_p X_{pk} ; \quad (10)$$

Where, f_k = forecast discriminant value of m in category k ;

X_{ik} = forecaster's quantitative value of m in category k :

u_i = variable i , for $i = 1, \dots, p$ Coefficients.



In the research work, the bitrate and frame rate were used as video parameters for six video content categories and three exclusive nodes. Authors of the research work went ahead to elaborate that, in order to benefit from efficient network utilization and maximization of user perceived quality, all other issues and factors relating to the video coding parameters and content have to be recognized. The model's accuracy on each node yielded 83.9% for laptops, 76.9% for mobile phones and 86.6% for PDAs. Nonetheless, because of the statistical method used to construct these prediction models, the model lacked a favorable degree of accuracy. Furthermore, during implementation, there was no consideration for QoS parameters at the

Network Layer. The work presented in [139] also employed the statistical analysis model as a gateway for QoE optimization during video streaming.

2.23.3 Machine Learning based QoE models

Research work in [136] which was also extended further in [140], prediction models on QoE were implemented through the use of machine learning methods, thus there is recognizable use of decision trees[141] and support vector machines(SVM)[142]. The various models that were implemented in these works performed with an impressive accuracy of 90% and above. The estimates in these works used cross-validation technique[143]. Through validating the Decision Tree (DT) model's accuracy of three nodes, mobile phones achieved a 93.55%, laptops achieved 95.46% accuracy and PDAs – 90.29%. As far as the SVM model was concerned, mobile phones yielded an accuracy of 88.95%, laptops-91.45% and PDAs-89.38%. An observation on the two methods was made and the authors concluded that the two methods outperformed the method used in [136] i.e., Discriminate Analysis (DA). Nonetheless , in [144] it is indicated that the error percentage of the models falls between 10% and 20%. The authors in [145] likewise implemented an objective QoE model using SVM and DT. They went ahead to compare the models with other various Machine Learning techniques, these included Random Forest(RF), Neural Networks (NNet), Naive Bayes (NB) and kNearest Neighbors (k-NN). The results of the research work showed that RF's performance was more recognizable than the other techniques studied.

Another study that was presented by the authors in [144] presented a technique that could straight away be linked the QoE scores to QoS parameters based on the analogous QoE level through a WiMAX network. A Multiplier Artificial Neural Network (ANN) was used to estimate QoE. Network QoS metrics were then identified and chosen to be the input layer, and output layer was made up of VQM, PSNR, MOS and SSIM. The ANN technique was

configured and tested to generate the accurate weights. A database of video was used to estimate and gather results. The database was further broken down into three smaller groups: these included, 70% training for ANN, 15% testing for the database and another 15% to certify the process of training of the database. With the completion of training of the ANN model, there was an establishment of the relationship between the input layer and the output layer. The findings and results of the work indicated that the model proposed is able to maintain the accuracy in prediction. The implemented model also had the ability to change the network parameters and receive an optimal output as a way of satisfying the user's need. Nonetheless, a bigger database is required to be the input for the neural network while coping with the other scenarios. Furthermore, the suggested model did not depend on any human interactions thus there was no wastage of time.

The authors in [146], [147] and [148] also employed the ANN method to alter the input parameters as a way of acquiring the desirable output for user needs. In essence, for the ANN method to be successful, the model needs to know the non-linear links that exist between QoS and QoE.

In [149], the authors presented two different models that use MOS to anticipate video quality. An Adaptive Fuzzy Inference System (ANFIS) was chosen as the benchmark for the first model and the other model was benchmarked with non-linear regression analysis. The authors went ahead to inspect how QoS affects video end-to-end quality of the encoded H.264. They also looked into how models with radio link loss are affected over UMTS networks. During their experiments, the authors combined both application and physical layer parameters to have the models adapt to the network; an unseen dataset was then used to validate the models. The results generated from these experiments showed that both models had the ability to accurately predict outcomes. The conclusion that the authors arrived at was that the selection process of parameters is very essential in achieving a good reliable and accurate result. Therefore the

models discussed in this section still have to be validated using subjective testing. The other scholars especially in [6] also employed the ANFIS technique in order to ascertain that there's a normal relationship between QoS parameters that have an effect on QoE along with the perceived QoE.

2.23.4 Crowdsourcing for subjective simulations as a QoE model

QoE modelling and assessment for online video streaming was a challenge for the applications that depend on the Transmission Control Protocol (TCP), e.g., YouTube. The authors in [150] did find the solution to this challenge. In their work, they implemented a model of QoE that mapped video breaking and stalling situations to ratings of the users. The breaks and stallings were mainly caused by various bottlenecks within the network. The authors went ahead to further propose a subjective assessment method for evaluating QoE based on the principle of crowdsourcing. Crowdsourcing involves obtaining a service, an idea or feedback using open calls to a big group of randomly selected people online. To carry out the different online user surveys, the platform of microworkers1 crowdsourcing [151] was used. In order to refine and analyze the diversity in users, the SOS hypothesis was employed [152]. Examples of diversity issues that were considered prior to investigating the various factors that affect YouTube's QoE included, rating uncertainties, impairment sensitivity, the quality expectations of the individual users and types of users. The findings and results of this study showed that video stalling situations indeed influence QoE. However, there were no findings on recognizable effects by the other factors such as age, the level of using the internet and type of video content. The authors observed that various users will be able to accommodate at least a single stalling or break in video for about three seconds on each clip but eventually if these breaks increase, the dissatisfaction levels will also increase. Nonetheless, because of the anonymous nature of the users in the platform of crowdsourcing and general control, it is only fair to assume a realistic level of reliability. Among these users, there could be a couple who may hand in low quality

work as a way of maximizing payments while at the same time not putting in much effort. It is for this reason that it becomes inevitable to design and develop a more practical and sophisticated model that will be able to assess the legitimacy, trustworthiness and quality of the users being used in the experiments. Crowdsourcing as an approach was also used in [145] to gather QoE data in order to assess the quality of network services.

2.23.5 Resource Arbitration based QoE modelling

In[119], the authors presented a framework that estimates QoE based on the thesis that in order to accomplish improved QoE, all parameters of QoS should be well reconciled and consolidated rather than focusing on each independently. Application and Network QoS metrics were weighted and defined by a standalone QoE factor. The authors argued that QoE should be measured as a function of Application Quality of Service (AQoS) and Network Quality of Service (NQoS), as shown in Equation 11.

$$Q = f(A, N) \quad (11)$$

Based on the above thesis, the authors in the study went ahead to design a resource arbitration mechanism. Further experiments were carried out using a network enabled QoS and user replies from software representatives where the users physically altered values on the network layer and application layers in order to implement QoE. To do all these, the authors used the DSCQS tool [153] to engage with users to poll their satisfaction or dissatisfaction of the service. After this, PSNRs were then correlated using prioritization mechanisms. Jitter, packet loss and delay were the chosen metrics which were dynamically altered in order to increase QoE. The results of this work showed that QoE can be improved by network arbitration, signifying that the synergy between Application and Network layer arbitration can successfully administer improved QoE. Nonetheless, if at all there was any observation of a positive change

in video quality, this was entirely because of a drop in the quality of the other media services. The model presented in the work however targets only the NQoS parameters.

2.23.6 Consideration of environment and equipment factors in QoE modelling.

In order to make the use of network resources more profitable, use of scheduling algorithms in QoE modelling is gaining essence in the research community. The authors in [154], presented a scheduling algorithm known as QoE-Based Scheduling(QBS). The idea involving both equipment and environmental factors, was to make the development of this Video QoE model a successful endeavor. These factors greatly impacted the QoE of a user absolutely with reference to network QoS. In this work, bit rate and throughput were the main QoS parameters that were focused on. The main task at hand for the authors was to examine and assess the hardware and the parameters of the environment using the user's equipment. Subsequently, the quality of the network signal would then be acclimatized based on the results from the simulations as a way of adhering to the needs of the user. After examining the results, the authors observed that, by applying the QBS algorithm, user satisfaction greatly improved more especially for those users who had limited resources. They argued that, agents in the surrounding environment such as shaking, noise, light, etc., resulted in reflecting various levels of QoE of the user. In one of their concluding remarks, the authors attest that, recognizable QoE is indeed achievable through the provision of signals with a high quality towards sophisticated devices with a profound interference from surroundings and the reverse is true. Nonetheless, the authors went further to recommend that the QBS algorithm requires more optimization, and that deeper user simulations on QoE are needed with a variety of scenarios while including more QoS parameters.

2.23.7 Quantitative and Qualitative Assessment based QoE modelling

Qualitative and quantitative assessment approaches are yet the other ways of modeling QoE. In [155], a merger of the Rough Set Theory(RST) [156] and (Catalog, Categorize and Analyze)-qualitative and quantitative approaches respectively to assess network quality ,is witnessed. The impact combining NQoS parameters with AQoS parameters along with the content features over QoE were also emulated and evaluated by the authors. With reference to the quantitative assessment, the authors observed that the various categories of video content need separate QoS support levels and that NQoS and AQoS parameters do also have their own levels of impact on QoE. On the other hand, during qualitative assessment, the authors observed that video clips that were relatively slow received few adverse remarks than the fast-paced video clips. Likewise, more adverse remarks were received for the two types of clips as NQoS kept on varying. However, there were no variations in the bit rate of video reported. Therefore with the presented results, it was concluded that bearings within the comments qualitative in nature, mostly matched the quantitative data evaluations.

This study only presented four QoS parameters as far as the implementation of this model was concerned; this made the findings look slightly perceptive. Nonetheless, within an actual environment, the influencing factors are bound to increase which ultimately makes the striking of a link or interdependence amongst them a challenging task.

2.24 Summarized findings of studies that confirm links between QoS and QoE

Fiedler *et al* [157], confirmed that there is indeed an exponential interrelation between QoS and QoE. In this work, the authors presented an exponential interdependency of QoE and QoS (IQX) hypothesis. The formula used in this work mainly related to the various changes in QoE in relationship to QoS up to the present stage of QoE. According to the results, the approximations presented turned out to be of more quality compared to the logarithmic

approximations. The authors recommended the use of the IQX hypothesis whenever investigating the links between QoE and QoS parameters.

Siller *et al* in [119] made a conclusion that through the use of prioritization mechanisms, network arbitration can greatly improve Network QoE, thus the authors conclude in their work emphasizing that QoE can get much better whenever the parameters of QoS are conciliated in an integral manner compared to doing it independently.

The authors in[127], sought to propose a mechanism that would use an objective approach to mapping QoE to QoS where Video Quality Metric(VQM) would be used to indicate the status of QoE. The mechanism had the ability to source the VQM scores and use them to map QoS parameters to QoE as far as video quality is concerned.

The model that was developed in [136] by Agboma *et al* sought out to explain how QoE metrics are captured. The authors used a statistical modelling technique to interrelate QoS parameters with QoE and perception estimates. The results in this work indicated that with the implementation of the model, PDAs had an accuracy of 86.6%, Laptops: 83.9% and Mobile phones 76.9%. Their conclusion was that it is indeed possible for network administrators and operators to foresee QoE and be able to apportion the available network QoS resources appropriately.

Menkovski *et al*[140] proposed a novel approach to implement an adaptive and accurate prediction model for QoE using machine learning classification algorithms mainly based on subjective data. The novel models developed are mainly used for real world applications. The results indicate a success rate in accuracy of above 90%, thus the “J48_model” produced 93.55% for mobile phones, 95.46% for Laptops and 90.29% for PDAs. “The SMO” model produced 88.59%, 89.38% for PDAs and 91.45% for Laptops.

The model presented by Kim *et al* in [158] evaluated IPTV QoE since network providers of IPTV require methods of controlling QoE whenever users become dissatisfied with the quality of video streaming. To achieve this, the authors had to first thoroughly analyze the QoS parameters of jitter, delay, throughput and packet loss which influencing QoE. The model presented was able to predict user QoE within the network environment provided and also evaluate the service itself.

The authors in [149] proved that a better prediction accuracy in video quality is achievable but only when the appropriate parameters is selected. The authors contend that NQoS parameters are bound to sensitively affect video quality compared to the AQoS. The authors went on to conclude that the models developed would be in position to assist in the development of a reference-free video prediction model and control methods for QoS for streaming video over UMTS networks.

It is of great importance to clearly understand the relationship between QoS and QoE. The authors in [6] ran simulations of video streaming in a typical broadband network with alternating QoS parameters such as packet loss, packet reordering and bit rates. Sample users were then required to give quantitative feedback. The results showed an exponential link between Perceived Video Quality (PVQ) and the QoS parameters. The authors confirmed that QoS management, PVC and network performance are indeed linked.

Han *et al* in [154] , identified a lack of network optimization past implementation of QoS leading to unsatisfying results especially from the user's side. Their solution was to find means of amending resource allocation in wireless networks by intensifying on user QoS in order to further intensify user QoE. The simulation results harvested indicated that there was an improvement in user satisfaction after QBS algorithms were implemented most especially as far as the users with limited resources were concerned. The authors concluded that, in order to

realize an efficient network resource utilization and achieve user satisfaction, the latter's hardware, equipment and environmental concerns had to be put into context.

The authors in [155] came to a conclusion that firstly, different content types needed different QoS support levels. Secondly, AQoS and NQoS parameters affected QoE differently. This conclusion was arrived at after the authors used the Rough Set Theory to quantitatively assess and evaluate the impact of QoS parameters and content features on QoE.

Koumaras *et al* in [159] presented a framework to predict MPEG-based video sequences end-to-end quality. In this work, the authors implemented two discrete models; the first model was to predict video quality basing on signal encoding and the second model was to map QoS parameters to the degeneration in the quality of video. From the experiments in this work, the authors achieved their goal because the presented framework overcame the various predicaments of QoS parameter mapping to video quality degeneration.

The authors in [147] , presented and implemented a mechanism to evaluate PVQ streams in packet networks by using artificial neural networks. The tool used in the evaluation was able to accurately predict MOS results. After analysis of the results, the authors realized that the conventional QoS parameters which are controlled by MPEG-4 codecs showed weak consequences on the PVQ streams. This is contrary to the recommendation that effective network development requires the protection of video flows. This yet another observation that QoS and QoE are interrelated.

Network bottlenecks in real time streaming applications have become prevalent lately. The authors in [148], proposed and presented psych-acoustic-visual models that predict multimedia application QoE. The results presented in this work show that the models presented are indeed well placed to monitor and allocate resources (QoS parameter based) within content delivery networks.

Mok *et al* in [160], investigated the interrelation amongst Http video streaming QoS, NQoS, AQoS and QoE. The authors used empirical evaluation and analytical models to establish the link between AQoS and NQoS. Secondly, they carried out subjective experiments to assess AQoS and AQoE. From the results of the experiments, the authors concluded that; the most critical factor in QoE variations is the rate and frequency of rebuffering.

In terms of network flow management, Elkotob *et al* in [139] presented a mechanism that permitted a mobile device to always be extremely alert of the finest available access network with the best resources for the ensuing available interval. The authors recommended the mechanism to network operators who may have that need to maintain user QoE while optimizing network QoS resources.

The authors in [145], investigated the various factors that affect QoE in cloud based video streaming applications. Network parameters (QoS), video features, terminal features and used types were identified as the main factors. From the described methods of collecting QoE data (MOS) and the use of machine learning (ML) methods to categorize the QoE data, the authors observed Decision Tree (DT) as the method with the desired performance in terms of the Mean Absolute Deviation (MAD) rate in comparison with the other algorithms. However, upon the authors carrying a statistical classification analysis, the results reflected the Random Forest (RF) method outperforming the DT method.

2.25 Conclusion

Having carefully reviewed in detail the literature surrounding the relationships between QoS and QoE through the use of interrelation models presented by different scholars, we can conclude that the main aim of all these models are twofold: first, to realize QoE based on QoS and secondly to implement the relevant QoS controls with the expected QoE. As observed in this chapter, numerous researchers presented different models meant for different multimedia

and content types. It was seen that different measurement mechanisms exist for the different media types and possess different operational and computational requirements. Based on this observation, it remains challenging to present and discuss the level of efficiency of the various models mainly because there is no standard to follow. Further on, QoE isn't adequately described anywhere in these works. Much as relationships exist between QoS parameters and QoE, the effects QoS has on QoE is yet to be clearly understood and the failure to have a standard method to quantitatively link QoS to QoE remains absent. The majority of the current mapping models are mainly skewed towards predicting QoE. Therefore, it is this observation that continues to motivate more researchers to try and present methodologies that will be efficient in interrelating QoS to QoE basing their models and mechanisms on statistical, quantitative and subjective metrics.

The literature reviewed in this section also showed that the simulated optimization techniques in the various research works had the ability to adapt to the models and realize an optimal relationship between QoE metrics and QoS parameters. The degree of success of the optimization techniques solely depended on the ability of the various models to adapt to the non-linear links of QoS to QoE.

Having reviewed these works, interrelating challenges of QoS to QoE remain. Therefore there is still a desire to clearly pick out and understand the various QoE factors that impact on a service and how they may affect each other. Caution however remains on those factors that may be complicated to measure. For example, context information on the user could go a long way in contextualizing QoE, however there could be a price to pay in terms of privacy and security issues. Lastly, it is recommended that quantifiable subjective metrics be used in modelling of user behavior. Then, the results from any simulations have to be carefully evaluated to make it possible to identify the pivotal factors that may influence QoE and

desirable QoS parameters. All this will in turn be used as input in achieving admissible optimization strategies for implementing QoE.

2.26 Chapter Review

In this chapter, the most recent implementations and innovations from various authors on NGMNs were presented. The chapter presented the various challenges that NGMNs experience, it is also in this chapter that RQ1, RQ2, RO3 and RO1, RO2, RO3 respectively were all addressed. Video Streaming is one of the many innovations that has benefited from the recent advancement of NGMNs, the chapter presented different works on the various video models that exist, video compression, video formats, types of video streaming among others. The chapter also went ahead to present the communication ecosystem as a key component of all innovations around NGMNs. The chapter then presented works from various authors concerning QoE and QoS testing in test-bed living lab settings. It also presented what it takes to improve QoS in mobile networks. The discussion mainly centered on improving faster web downloads, avoiding video stalling during streaming and subscriber usage of roaming services, among others. A comparison between User centric and Techno centric QoE approaches was also done in this chapter. The chapter then went ahead to give a detailed discussion on the previous work done on QoE layers, various approaches of QoE measurement, how the models of objective quality assessment are classified and then how subjective QoE scores can be mapped to objective QoS scores. During this phase, work by other authors that was done to establish relationships between QoS and QoE was also evaluated and we came to a conclusion, the task of mapping QoS to QoE remains a great challenge and yet this task is used as a yard stick in assessing service quality as far as multimedia applications and networks are concerned. From the literature reviewed, majority of the analysed models did source their subjective metrics from the users by using MOS. It is with this evidence that much as QoS-to-QoE interrelating models exist now, the majority of these models merely try to predict QoE.

Because of this, these models remain application-specific on top of having contrasting operational and computational requirements. This is the major motivation behind the ever growing need to focus investigations on user-centric QoE vis-à-vis net centric QoS parameters while establishing mapping functions between the two. Therefore, the chapter concluded by presenting the answer to RQ3 by confirming that indeed the link between QoS parameters and QoE metrics exists, however based on the literature reviewed, there is still a lot that needs to be done to achieve more network performance in terms of resource optimization and user satisfaction. The next chapter introduces the research methodology that was used to answer RQ4 and achieve RO4.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Chapter Overview

The main contribution of this work mainly includes implementing a successful independent user centered QoE and QoS monitoring and evaluation mechanism to monitor video streaming in Next Generation Mobile Networks. This chapter concentrates on introducing the various concepts behind the modelling, testing and implementation of the mechanism mentioned in RO4 in chapter 1. The chapter begins by introducing the research methodologies that exist to execute scientific research. It then introduces the methodology that was used in this thesis. This is done through describing a QoS/QoE administration system that mainly has three layers of operation. The description of these layers sets the scene for the deployment of the mechanism.

3.2 Introduction

A research methodology is a way of gathering information that will help in making specific decisions. The process of gathering this information may involve the use of both primary and secondary research methods or also known as qualitative and quantitative research methods[161]. The two are briefly discussed below.

3.2.1 Quantitative Research Methods

In 2007, Holland and Brook[162] presented quantitative research methods as ways of executing experiment-based studies or investigations. In this method of research execution, opinions and views of various scholars, academics and authors on a particular area of research where research information is collected, analyzed and findings discussed. As already hinted, articles, papers, case studies, journals, book chapters, theses and web documents are some of the different types of secondary data sources that are used in carrying out quantitative research.

3.2.2 Qualitative Research Method

Collis and Hussey in [163], emphasized that actually the various ways in which humans behave and act in different situations play a crucial role in executing qualitative research. In comparison with quantitative research methods, the sources of data when out carrying quantitative research are much more reliable and spot on since they may include views and opinions from various research candidates, basically this impetuous information that has no authors' opinions. Qualitative research mainly depends on research interviews, surveys and focus groups as the primary sources of data. Thus, questionnaires in this method play a key role since questions of research to be answered can be pre-fixed and analyzed when designing the questionnaires.

3.2.3 Mixed Research Methods

A mixed research method involves using both qualitative and quantitative methods to execute a research project. According to Gibbs [164], all sources of data are employed in drawing conclusions about a research problem. This method is mainly used if and when there's a need to get authors' opinions and human opinions on a key issue being investigated. This method can employ any of the following techniques to execute the research: Research interviews; research surveys; research questionnaires; focus groups; scholarly articles; individual and company case studies; peer reviewed journals and expert reviewed web documents. Thus in our work we employed the mixed research method since it both involves experiments and collecting of user opinions through using Mean Opinion Score (MOS) tool. We later elaborate on how the tool is used.

3.3 Methodology used in the thesis

To further execute the research successfully and achieve the objectives and the main goal of this research, the main research techniques used are further presented in this section:

3.3.1 NGMNs State-of-the Art review

An in depth state-of-the art review on the key areas of NGMN research work was done in Chapter 2. In order to find the answers to RQ1, RQ2 and RQ3 and also achieve research objectives RO1, RO2 and RO3.

3.3.2 QoS/QoE Administration Mechanism

The QoS/QoE administration mechanism presented in this section is a portion of the complete customer experience administration mechanism (CEAM) that is illustrated in Figure 3.1. This CEAM has three layers: data acquisition, monitoring and control. In order to fully define the CEAM, the presence of the data acquisition and monitoring levels is very essential. We briefly describe these two in our work since the main focus of our research contribution centres on the control layer. The data acquisition layer is mainly in charge of collecting all raw data using active probes or passive probes which are nonintrusive network monitoring agents. The passive probes collect data without causing additional traffic overhead on the network and usually they are transparent to the other components of the network. On the other hand, active probes send dedicated measurement packets to benchmark the network performance. A measurement layout usually consists of two or more probes performing independent measurements and possibly exchanging information among themselves. Nonetheless, layouts with a single point measure are achievable. This, in most cases, leads to measurable parameters which are smaller. For example, accurately estimating one-way delay with a single measurement point is not feasible in general.

On the monitoring level, the raw data produced by the data acquisition layer is processed into knowledge about the state of the network, which is in turn passed to the control level. The control level then performs actions on the network based on the collected knowledge. At this

level, we have three other essential components: the control manager, intervention manager and policies. These components are further described in subsections that follow in this section.



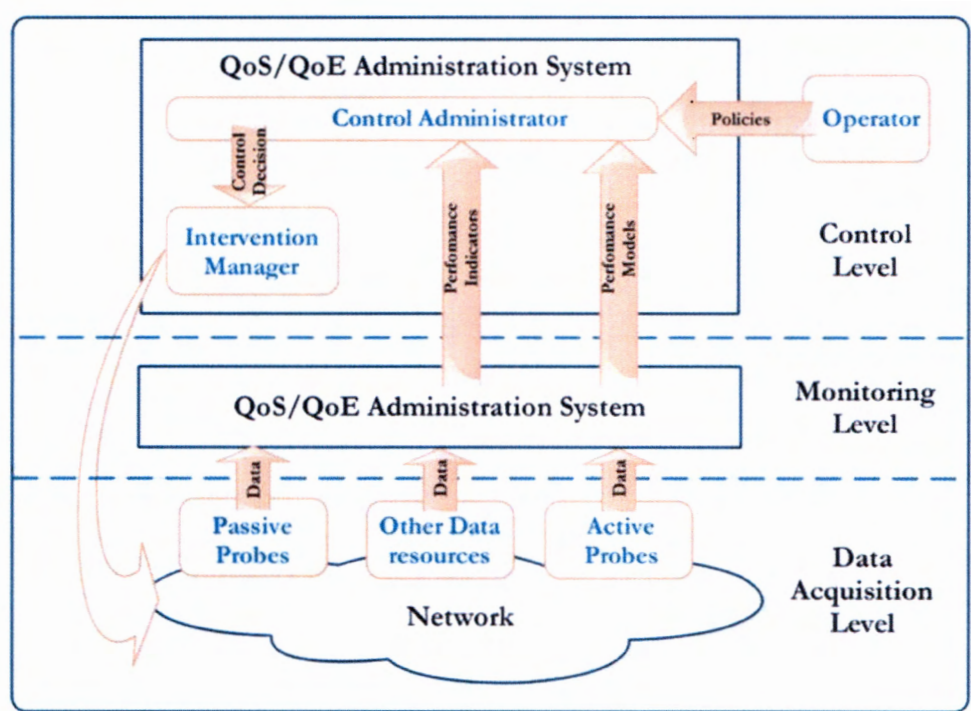


Figure 3-1 Customer experience administration mechanism (CEAM).

- **Control Manager (CM)**

The CM is responsible for determining the network condition and take the necessary decisive action that the intervention manager will perform. There are three parameters which control the decision process of the control manager. These are: policies; performance indicators (Per_Ind) and performance models (Pe_Mo). Therefore, the CM can be seen as a relation depicted in equation (11):

$$CM = Per_Ind * Pe_Mo \rightarrow Decision \quad (11)$$

a) Policies.

Policies involve the various guidelines and bounds that the network operator sets. Equation (12) presents these policies as a collection of network constraints as far as the field of operation is concerned. In the case of ISPs, policies are likely to be based on service level agreements (SerLevAg) policies (SLAP) and subscriber policies (SubsPol). Application policy (AppPol) is used more seldom but it is still important to this mechanism.

$$Policy = AppPol|SubsPol|SerLevAg \quad (12)$$

b) Performance Indicators(Per_Ind)

Performance indicators measure the current performance of the network. A common set of performance metrics are the QoS indicators (QoSIn) such as packet loss, jitter and delay. Other parameters are also discretionary, e.g, time to establish a connection, time for address allocation and average response time of some services. QoE indicators (QoEIn), such as Mean Opinion Score (MOS), MOS estimations, Peak-to-Signal Noise ratio (PSNR) and Mean Squared Error (MSE) are defined under QoE indicators (QoEIn). Since these options are not mutually exclusive, we define performance indicators to be a set of QoS, QoE and other metrics, this definition is represented in equation (13).

$$\text{Per_Ind} = \text{QoSIn}|\text{QoEIn}|\text{OtherInd} \quad (13)$$

Performance indicators express the current condition of the network, but we are also interested to know how control decisions affect the network QoS-parameters, and how a change in QoS parameters in turn affects QoE. In our mechanism, these types of frameworks are called performance models.

c) Performance Models(Per_Mod):

Performance models as shown in equation (14) hold a key role within the control manager especially where the need to optimize other parameters becomes inevitable to achieve favourable QoE. Examples of these parameters may include: costs, resource utilization and power consumption.

$$\text{Per_Mod} = \text{QoSMod}|\text{QoEMod}| \quad (14)$$

These models, along with a possible addition of usage models (Equations 15 and 16), can be used to estimate QoE (EstQoE) or QoS performance:

$$\text{EstQoE} = \text{Per_Ind} * \text{QoEMod} \rightarrow \text{QoEIn} \quad (15)$$

$$\text{EstQoS} = \text{Per_Ind} * \text{QoSMod} \rightarrow \text{QoSIn} \quad (16)$$

The control manager integrates the data provided by the three information sources (i.e., policies, performance models and performance indicators). This data is used in the control decision module to make corrective actions on the network. The actions are executed by the intervention manager, which is presented in Section 3.4. There is no restriction of what actions can or should be taken, but as the mechanism is designed for a network-level means of operation, we limit the possible actions to the following items and shown in equation (17):

$$\text{Action} = \text{ProvisionA} | \text{ShapingA} | \text{MarkingA}|$$

$$\text{AcceeCtrlA} \mid \text{HandoverA} \mid \text{AlertA} \mid. \quad (17)$$

The decisions are in respective order, dynamic network resource provisioning (ProvisionA), packet/traffic shaping (ShapingA), packet marking (MarkingA) for ensuing control, per-user or per-session access control (AcceeCtrlA), horizontal (HandoverA) and/or vertical handover and alerting (AlertA) the network operator. These control mechanisms are also discussed in Section 3.4. The last element required for the control manager is the decision, which is simply a list of actions as equation (18) shows:

$$\text{Decision} = [\text{Action}] \quad (18)$$

In simple cases, the decision can be a single action. In the prototype presented later in Chapter 4, we combine several actions in the decision process to deal with different scenarios and to mitigate quality degradations.

3.4 QoE Intervention Manager

The intervention manager is a physical device on the network which incorporates the necessary tools to actualize the decisions set by the control manager. This component may act upon all traffic traversing the point of control, traffic flows of certain users or applications, or individual packets, depending on the given task.

3.4.1 Dynamic Network Resource Provisioning

Dynamic network resource provisioning is the act of augmenting or re-distributing network resources, either adaptively during congestion or statically for certain services or users. This can be done through bandwidth reallocation or re-routing traffic.

3.4.2 Shaping

Traffic shaping enforces certain performance bounds on different types of traffic. The focus of this mechanism is to limit the throughput of certain users or services, e.g., peer-2-peer traffic. The same method can also be used to throttle down other traffic in order to give better performance to certain users or services. In other words, shaping allows the establishment of traffic classes with different priorities. Also, an important aspect in traffic shaping is that if certain traffic is restricted to enter the outgoing link, the packets of that traffic are queued instead of just being dropped, which results in an increase in delay instead of in packet losses (unless a buffer overflow occurs).

3.4.3 Access Control

If the control manager detects that the network does not have enough capacity to meet any new resource requirements, access control which is also known as admission control is one way to prevent traffic from congesting the network. Access control can be implemented on the user level by preventing users from accessing the network; a good example where this happens is in ATM and cellular networks. Another option is to allow all users to access the network but terminate new traffic flows if these flows would deteriorate the quality of existing ones. This model is seen to be more befitting in packet switched networks which are not service-specific oriented since the behaviour of users can't be predicted beforehand. Users may consume a large portion of the network capacity and use very quality strict applications or simply idle in the network. Access control can be used in conjunction with other mechanisms to gain flexibility to the administration system.

3.4.4 Handover

In a scenario where any point of a network gets congested, the current and fresh access point connections can easily be transferred to the other access interfaces. This is known as a

“handover”. Handovers are either vertical or horizontal depending on where the handover takes place, thus either amongst single type network interfaces such as two WIFI-hotspots or an array of contrasting network technologies, for instance 4G to 3G[165]. In this context the role of the QoE control manager is to decide the best network selection during a vertical handover. To achieve this goal, different vertical handover decision algorithms may be applied.

3.4.5 Marking and Alerting

Two actions which do not have a concrete impact on the network by themselves are packet marking and alerting. Alerting simply means that the system notifies the network operator under certain conditions, for example, if the network or a part of it becomes heavily encumbered. Marking is the act of modifying packets by adding a specific mark in a protocol header. The Differentiated Service Code Point (DSCP) on the network layer can be a desirable candidate to be flagged for marking. There are also some technology-specific fields, like 802.11p or 3GPP QoS classes meant for mobile traffic. Other network components outside the system can be given instructions on how to deal with different markings. Marking and alerting are not effective actions by themselves, however, the latter can still be deemed effective if they can bring about various procedures outside the system.

3.5 Chapter Review

This chapter presented the methodology that was used in this thesis. It also introduced the various concepts behind the design and modelling of the mechanism mentioned in RO4 in Chapter 1. This was done by describing a QoS/QoE administration system that mainly had three layers of operation. The description of these layers sets prepared the implementation of the mechanism prototype presented in the next chapter.

CHAPTER 4

IMPLEMENTATION

4.1 Chapter Overview

This chapter presents the implementation of the independent network access point management mechanism prototype of our solution presented in Chapter 3, this was able to monitor, manage and evaluate traffic traversing from the Internet to end users' devices. It can be placed at the edges of a 3G or 4G network before the radio channel, but because the system operates on the network layer and above, the underlying technology is not restricted in any way other than requiring it to have an IP layer.

4.2 Prototype implementation

The solution we presented in previous chapter demonstrated some features of the mechanism and also presented a strong instance of how the mechanism can be integrated into a concrete application. An outline of the mechanism design and its components are given in Figure 4-1, which follows the same layering principle in Figure 3-1. The bottom level (network path), contains the physical path from the rest of the ISP's network to the transfer media. The Control Manager and associated types described in Section 3.2.1 are located within the monitoring and control levels in Figure 4-1. Clients connect to the Internet through this path using various devices. This path is marked with solid arrows (packet traffic). Other signalling between different software components are marked with dashed arrows (control data). The goal is to monitor and manage the downlink connections of the clients using several tools and methods in the access point. To improve scalability, the access point mechanism operates in an independent nature thus it is possible to deploy and de-activate the Access Point (AP) independently. There are two components in the network path: the Linux socket buffer and scheduler. The Linux socket buffer is a double-linked list which contains the packets traversing

from the network towards clients' devices. This is the default network queuing component in Linux operating systems, both in desktop computers and embedded systems such as routers. We use a queuing discipline to implement our traffic control. Despite the fact that a queuing algorithm appears on almost all levels, the scheduler can be placed in the network path. A scheduler in this case is tasked with making the decisions on the dequeuing order of packets. It is this scheduler that physically implements traffic control. The scheduler used in this implementation is a part of the hierarchical token bucket (HTB) queuing discipline[166]. This discipline allows us to form traffic classes with different real-time constraints. This is the tool that the intervention manager uses to realize the control decisions. The use of this tool in our work will further be described later. Data acquisition and monitoring levels are the next levels within the system. All probes and sources that gather raw data reside within the data acquisition level; the various segments that are tasked with refining the raw data produced to knowledge reside on the monitoring level. A traffic classification probe, a queuing discipline information collector and QoS assessment probe are the main data sources in our implementation. First, all the incoming traffic is mirrored to the classifier probe, which sorts packets into flows identified by source and destination IP addresses, ports and transport layer protocol. The probe also collects flow-based statistical information which is forwarded to the next level where the information meets a classification algorithm. The main task of this algorithm is to classify each flow based on the acquired statistical information. The passive QoS probe, on the other hand, is placed on the exit point of the AP after the scheduler. This probe collects a few QoS-parameters for quality assessment on the next level, namely packet loss rate and packet loss burst size. This information is used by the quality assessment component to produce a quality estimation for each relevant flow in real-time. The queuing discipline information collector is the last portion of the data acquisition layer. Its main task is to extract the queuing discipline classification information for the performance model algorithms. The performance model

algorithms produce a quality estimate, which helps in the decision process to determine the best course of action.

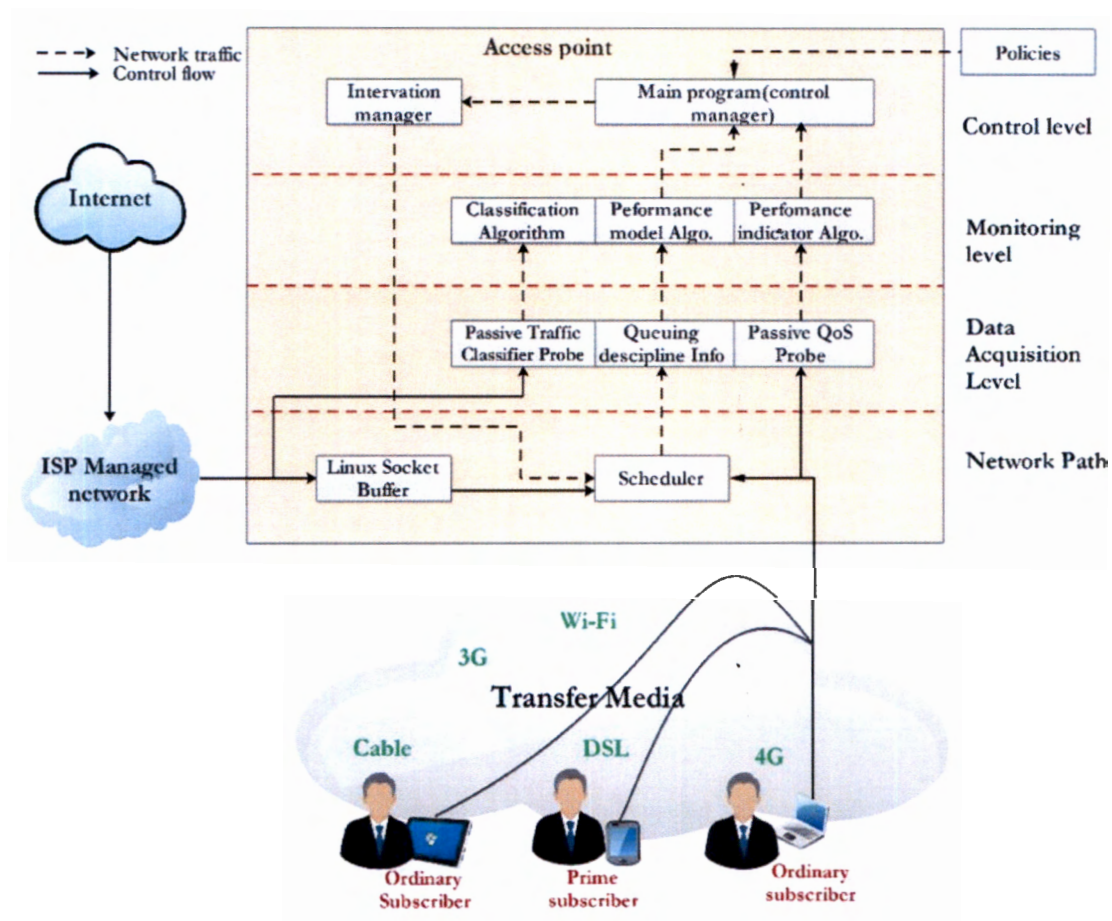


Figure 4-1: A conceptual view of the access point software.

Starting from the general definition of the control manager (CM_a) in Section 3.2.1, the management decisions ($Decision_a$) are governed by policies ($Policy_a$), performance indicators (Per_Ind_a) and performance models (Pe_Mod_a) as given by equation (19):

$$CM_a = Per_Ind_a * Pe_Mod_a * Policy_a \rightarrow Decision_a . \quad (19)$$

The information provided by the classification algorithm is also required by the control manager, since it contains the information of each traffic flow. The following subsections present the four elements in detail.

4.2.1 Policy

The policy of the designed system consists of application policies and subscriber policies. Applications are divided into three categories based on their real-time constraints: interactive, streaming and bulk. Interactive traffic has high two-way real-time constraints in bandwidth, jitter, packet loss and delay. Video conferencing, online games and VoIP calls are examples of applications that generate interactive traffic.

The second category in priority is streaming. This incorporates one-way streaming applications, which are intolerant to packet loss, low bandwidth and jitter. Applications such as IPTV and video-on-demand services fall in this category. It should be noted though, that applications in the same category may show different behaviour during quality degradations. For example, a packet loss in an RTP video stream causes visual quality degradations, but a packet loss in an HTTP-based streaming causes TCP retransmission to occur. This may result into a decrease in throughput, which may manifest itself as stalls and long buffering times on the application layer or a quality adaptation in the case of dynamic HTTP streaming systems such as DASH.

The last category contains all elastic traffic and traffic which goes beyond the interest of the network operator. In this work, focus was mainly on RTP-based video streaming. On top of application differentiation, we also implemented a user/customer differentiation strategy. We put users in two different categories based on their choice of subscription in the agreement with their ISP. Thus we had the first category as that of the prime users and the other as ordinary users' category. The prime users are those that purchase a much more expensive subscription type, this subscription provides them better quality in service mainly because of the availability of a high-end connection type. Ordinary users on the other hand are customers who rely on the traditional "best effort" service quality. The goal was to make the prime subscription more appealing to "heavy users", who would consume a large portion of ISP's base station capacity. Although we wanted to place prime users above ordinary users in service quality, less constrained applications could not supersede or pre-empt more constrained ones. Therefore, we settled on the following traffic categories (in the order of priority from the top to bottom):

- a) Prime interactive media;
- b) Ordinary interactive media;
- c) Prime streaming media;
- d) Ordinary streaming media;
- e) Prime bulk traffic;
- f) Ordinary bulk traffic/Default class;



The first four classes were QoE-managed. The last and lowest category is the ordinary bulk traffic (default class) in which all the traffic were originally sent to (with the exception of prime bulk traffic, which is always directed to the second lowest class). From the prioritized list, it can be seen that prime traffic always gets a higher priority in the same application class. However, this cannot be said for the different classes of applications. In order to implement our policy scheme, we used a sample subscriber database and a traffic classifier as the tools of

our choice. For the traffic classifier tool, we went specific with a machine learning algorithm which is known as a two-phased traffic classification tool. This tool relies on statistical features of packet traffic in the identification process, and it can be trained to classify individual applications or traffic which meets certain statistical criteria. The tool is well described in detail in [167]. However, the subscription information was managed by the AP software and we decided to leave out the subscriber database in the implementation.

4.2.2 Performance Indicator

In our experiments and simulations results we used MOS as a way to express the estimated quality quantitatively employing the average opinions as sourced from user feedback. The scale of MOS is a number between 1 and 5 as already presented in Table 2-1. In order to be able to give a verbal scale presentation, we adopted the degradation category rating (DCR)[97]. In this rating, the MOS score of 5 represented those impairments within the multimedia that were deemed “imperceptible” on the other hand, the MOS score of 1 represented the impairments that were “bothersome” in nature.

For the purposes of forming a complete data path from the network to the CM, we needed components on the data acquisition level and the monitoring level. We used the Qosmet tool [29] to perform single-point QoS measurements within the access point with a passive probe. The probe collects information at the data acquisition level, which is then transferred to the Qosmet controlling entity at the monitoring level. Although the performance indicator of this system considers only QoE, we used QoS along with application-level information to estimate the quality of experience in real-time. For this purpose, we used a machine-learning algorithm called pseudo subjective quality assessment described in [168, 169]. The algorithm basically maps a set of QoS and RTP-video stream parameters into a single MOS score. These parameters were packet loss rate, packet loss burst size, video resolution and the amount of

movement in the video. The data collected from a subjective video assessment campaign was used to train the algorithm to map these parameters into a MOS score. We therefore trained and validated the algorithm with an estimate of about 97 training samples and 21 validation samples, and the validation showed that the method indeed mapped well with the subjective data, giving rise to a 0.92 correlation coefficient.

4.2.3 Performance Model

It is the scheduler's responsibility to coordinate the traffic according to the QoE requirements. In our implementations, we concentrated on 6 different queues thus: four for each QoE-managed traffic class, one for prime bulk traffic and one for default traffic class. Since the traffic pattern was bursty in nature, we were interested not only in calculating the packet loss probability but also the mean loss burst size. This information was sufficient in our implementation in order to estimate the perceived quality of RTP-based video streams. The difference between a performance model and performance indicator therefore is that the indicator expresses the current flow-based QoE, while the performance model can be used to predict the impact of changing the priority class of a flow. Thus, with the help of the performance model, we were able to take more reliable decisions in terms of controlling traffic. The scheduling process had 6 separate M/M/1/ K_i queueing systems, where K_i denotes the storage capacity of each queue. For simplicity, we assumed that the arrival rate of each type i follows a Poisson process, denoted as λ_i and the service time of each type i is exponentially distributed with mean $1/\mu_i$ thus, the load ρ generated by type i is given by equation (20):

$$\rho_i = \lambda_i / \mu_i \quad (20)$$

Also, the loss probability p_{L_i} for each traffic type i can be expressed in equation (21)

$$p_{L_i} = \begin{cases} \frac{p_i^k(1-\rho_i)}{1-\rho_i^{(K+1)}} & \rho_i \neq 1 \\ \frac{1}{K+1} & \rho_i = 1 \end{cases} \quad (21)$$

If we assume that a burst of losses for the traffic type i is j_i , then the mean loss burst size for the traffic type m_i is given by equation (22).

$$m_i = \sum_{j=1}^{\infty} \left(\frac{\rho_i}{1+\rho_i} \right)^{j-1} \frac{1}{(1+\rho_i)} = \frac{1}{(1+\rho_i)} \quad (22)$$

While this is a simplification of the way the HTB scheduler works, it is a reasonable one for the short time scales involved in the decision making process.

4.2.4 Decision

As defined in Section 3.2., a decision is a list of actions. There are three types of control in this implementation, namely shaping, dynamic network provisioning and access control. Access control may only be performed when a new flow arrives. The other aspect is that different classes of traffic are always treated differently from the management point of view. It is at this point that provisioning and shaping are seen to operate. Top priority traffic classes do get a much bigger portion of the AP's capacity, and they are also dequeued with a higher priority. If there's no flow promotion into the higher class, the AP operates in a "best-effort" manner. The network resources are managed by hierarchical token bucket algorithm[166], which distributes the available bandwidth to the classes, while keeping utilization ratio high (that is, no bandwidth is ever left unused as long as there are packets waiting for transmission). Shaping encompasses intra and inter-class scheduling. Classes of higher priorities are always de-queued first, as long as they have bandwidth available. In addition, it is possible to schedule the classes

internally with the help of the stochastic fairness queuing discipline. The queuing mechanism is described in [170, 171].

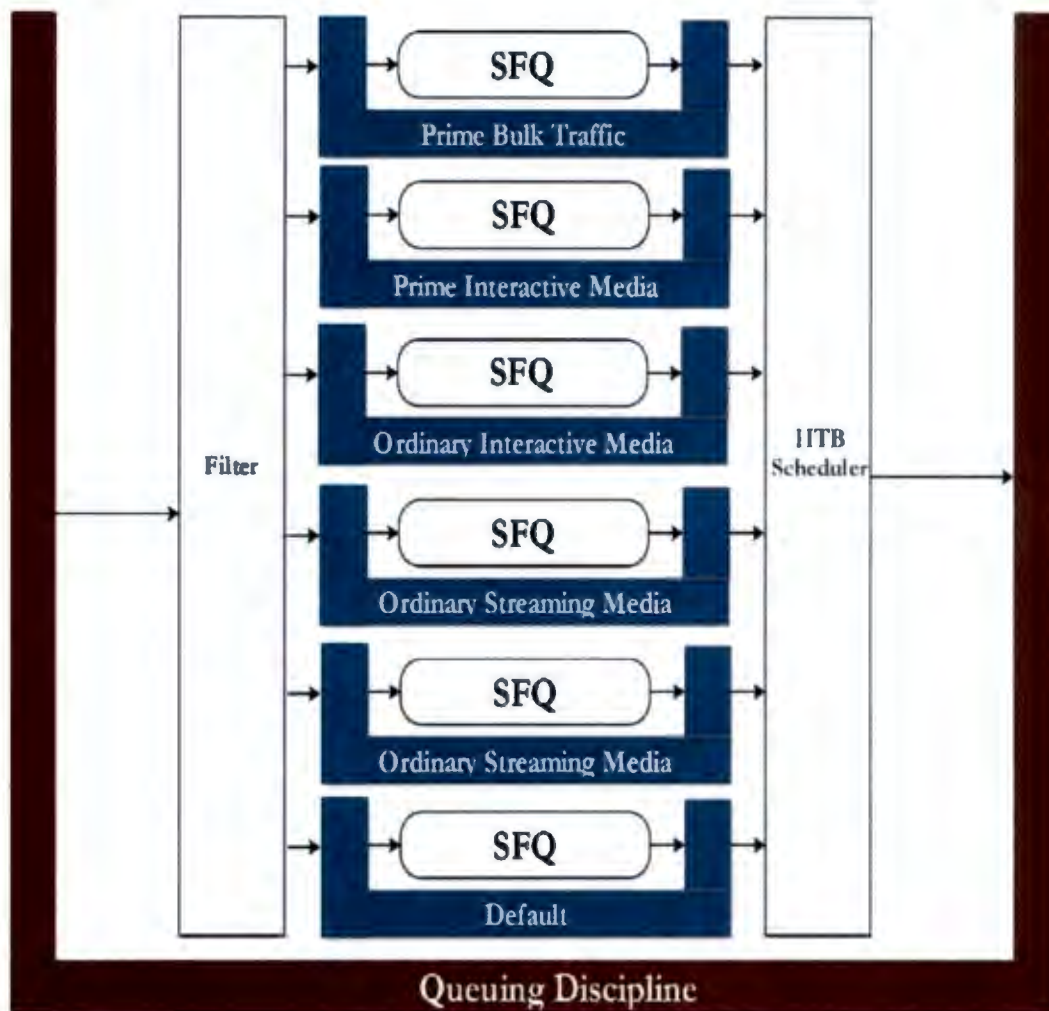


Figure 4-2: The queuing policy.

The complete queuing system is depicted in Figure 4-2. Just as a new flow attempted to get into the AP, the control manager made a decision as to whether the new flow could be supported or not, and thus if it could be permitted to enter. This decision was based on the application policy and the performance model. The admission control decision tree is depicted in Figure4-3.

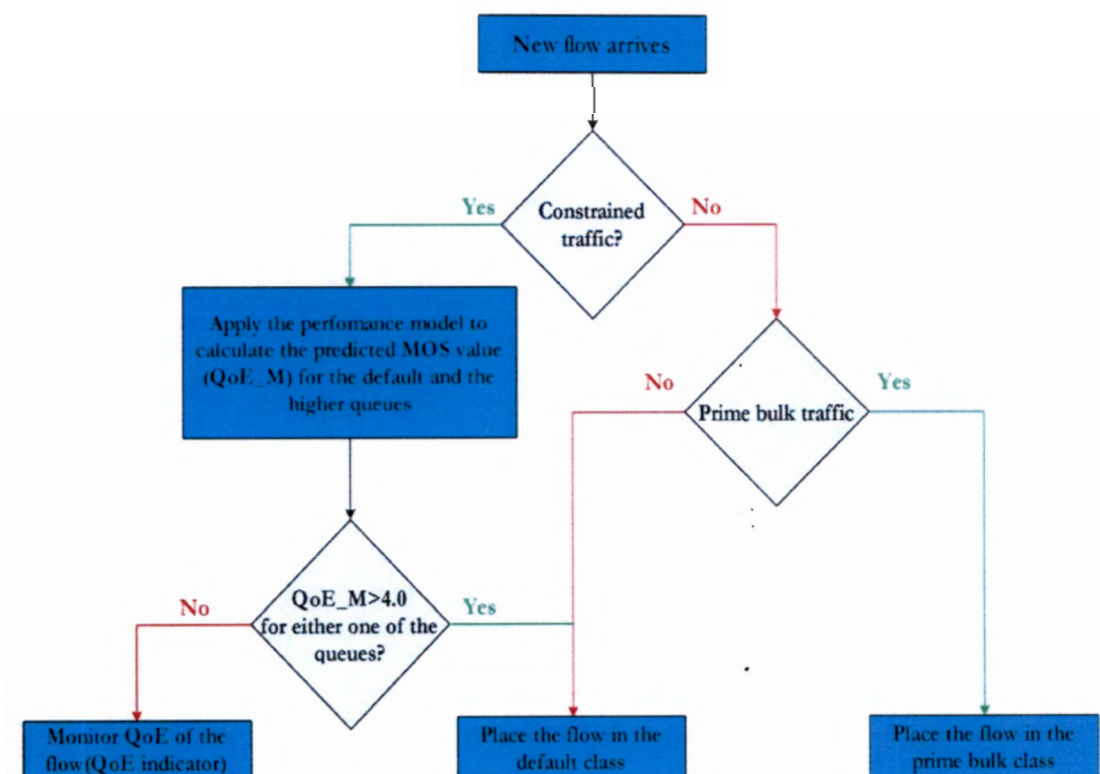


Figure 4-3 Access control decision process

Firstly, the flows which were singled out as bulk traffic were undoubtedly let into the access point. Since these flows are elastic, their resources could be re-allocated to other flows whenever needed, so there was no reason to reject them. If the bulk traffic belonged to a prime user, the flow was permanently placed in the second lowest class. On the other hand, if a flow belonging to a more constrained application class attempted to enter, the performance model was used to predict the QoE this flow would receive either in the default traffic class or the higher traffic class with respect to the application type of the flow. There were two possible queues for each constrained flow, either the corresponding higher class of the flow (e.g. an interactive application flow generated by a prime user was promoted to the highest class etc.) or to the default queue. At a point where it was decided that the quality within the queues was found to be below the threshold for admission (of which we set this to 4.0 in our system) the flow was rejected. Otherwise we came to a conclusion that the quality of the flow can be maintained in a satisfactory level in the system in at least one queue, so the flow was permitted to access.

As stated earlier on, prime bulk traffic was placed in the second lowest class, and ordinary bulk traffic and all permitted constrained traffic was placed in the default class. The flows with bulk traffic got to remain permanently in the classes though the same cannot be said for application flows such as interactive or streaming flows. Figure 4-4 depicts the decision process for constrained flows, which was processed every second for each flow.

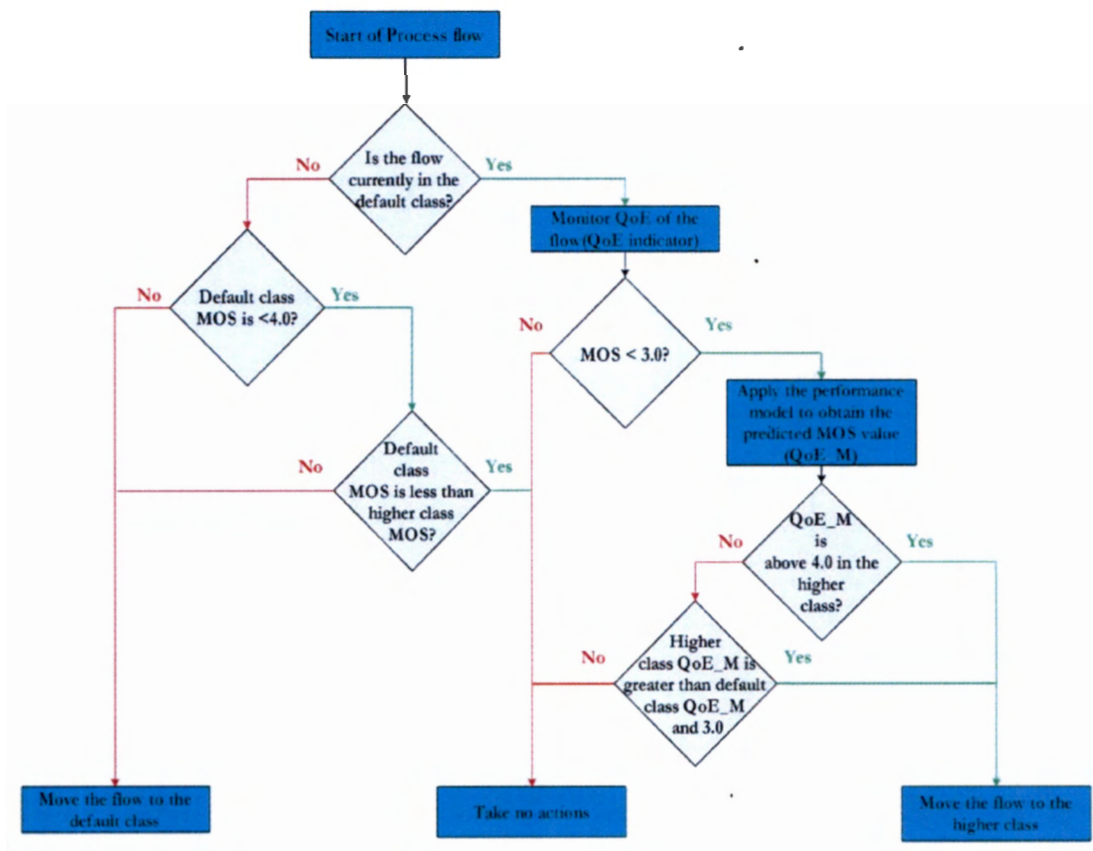


Figure 4-4 Per-flow QoE decision process.

Firstly we cross-checked the class that the constrained flow was presently in and if it was the default class, it stayed there until the estimated MOS of the flow dropped below “bad” threshold (was set at 3.0 in this demonstration). We then embarked on determining whether the flow should be moved to the higher class or keep it in the current class. For this task we used the performance model again. The estimated MOS-score was then calculated for both queues, and the flow was moved to the higher queue if the predicted MOS score of this queue was greater than the “good” threshold (which was set at 4.0). However, if this is false, then the system would compare whether the predicted MOS was still above the “bad” threshold and also above the predicted MOS of the default class. If these conditions were met, then moving the flow to the higher class was still beneficial. If all these simulations failed, then the best option was to take no action and keep the flow in the default class. The process of determining the best course of action for flows in higher classes was somewhat similar. However, to be able to circumvent over-prioritization that could have been brought about by a lot of flows being sent to the higher classes, we had to determine when the conditions in the default class were viable for flow de-promotion. First, the system checked if the predicted MOS of the default class was above the good threshold. If so, the flow could be moved to the default class right away. In the other case, moving the flow back to the default class was still a good option if the predicted MOS of the default class was above the MOS of the higher class. With this decision tree, we could limit the amount of flows in the higher classes so that the quality of these flows remained good. Otherwise flows could overload the higher queues during congestion periods.

4.3 Test bed Evaluation setup

Figure 4-5 shows the test bed architecture used to evaluate our mechanism after being implemented.

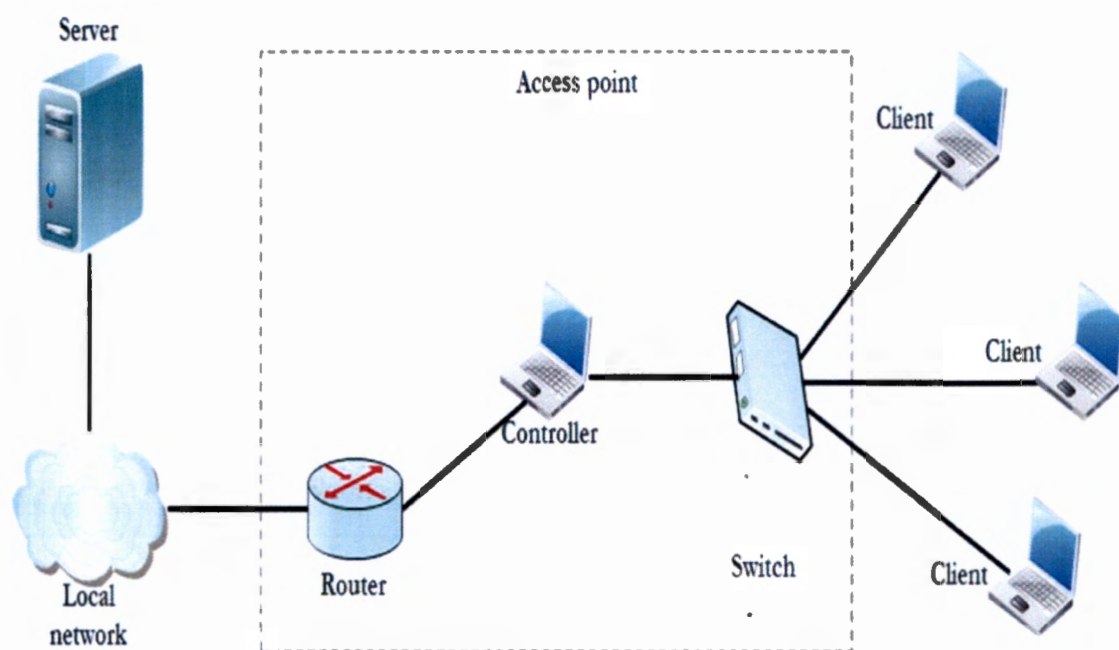


Figure 4-5: Evaluation test bed setup

Three client computers were used in this set up. The access point mechanism was implemented on a laptop (shown in the centre of the Figure 4-5). An external computer was used as a D-ITG and VLC media player server, and was used to transmit both bulk traffic and video streams. NAT on the subnet was implemented by the router and the switch. This was to allow more connections of multiple devices to the AP. This software can also be deployed on a router that is Linux based.

For all the simulations, we used the same reference video, which can be obtained freely from [36]. The main video featured a 1280x546 resolution video stream of 24 frames per second that was MP4-encoded. To emulate TCP bulk traffic, we used the Distributed Internet Traffic Generator (D-ITG) [37], which was set to generate 2048 packets per second, each having a payload of 256 bytes. For emulating packet losses, we used Netem, which is integrated into Linux iproute2 toolset [38] we also used the queuing disciplines required to implement and realize the required behaviour. We then ran several simulations in order to study the implemented mechanism. It was always going to be a challenge to complete such holistic simulation for this mechanism and the same can therefore be said for extracting the desired information from the same mechanism. It was for these reasons that we decided to choose several simpler scenarios to validate the different aspects of the mechanism's functionality: we tested the differentiation between traffic classes to observe how inter-class prioritization impacts the perceived quality. Then we ran an intra-class simulation where ordinary and prime users have to share common resources. We then performed a more complex simulation where the AP had to resolve the distribution of resources to a mix of different applications and users with different classes. We also tested the admission control algorithm to make sure it didn't allow the AP to become over-burdened. Finally, an important thing was now to measure how fast the system could react to a quality degradation, this was because of the reason that the reaction time is a critical component in user satisfaction. Each simulation except the

performance simulation was run in pairs; one without the administration mechanism (i.e. a simple “best effort” FIFO-system) and the other with the administration mechanism. These paired runs were repeated 10 to 40 times depending on the simulation to acquire statistically significant values showing that the system operates as demonstrated in the figures. The standard deviation of MOS scores for each simulation was calculated to achieve consistency and reliability of the results. The deviations were calculated for each video stream as follows. In a single simulation run, samples were taken every second; each sample was rounded to the nearest second and MOS scores were windowed with 4 samples; For each second where there are at least 2 samples a standard deviation was calculated, and all samples were finally averaged to get a mean standard deviation.



4.4 Chapter review

Using a portion of the QoS/QoE administration system described in chapter 3, this chapter presented in detail the prototype implementation and evaluation set up of the mechanism that addressed the main research objective (RO4) in Chapter 1. The major purpose of the evaluation set up was to assess the performance of the deployed mechanism. The next chapter presents the analysis and discussion of the results generated from the evaluation set up.

CHAPTER 5

RESULTS ANALYSIS AND DISCUSSION

5.1 Chapter Overview

In this chapter, we present the results generated from our various simulations from the evaluation set up discussed in Chapter 4. It is with these results that the ways in which the designed mechanism impacts on QoE are discussed. The chapter firstly, presents two simple simulation cases where traffic and client differentiation are demonstrated. Then the performance simulation which measured the reaction time of the system follows. The chapter ends by presenting last two simulations that show how the admission control prevented traffic from over-burdening the AP, and how the system was able to balance the load of different traffic types in a more complex scenario. For all our simulations, we use NetEm traffic shaper. NetEm is an enhancement of the Linux traffic control facilities that allow to add delay, packet loss, duplication and more other characteristics to packets outgoing from a selected network interface. NetEm is built using the existing Quality of Service (QoS) and Differentiated Services (diffserv) facilities in the Linux kernel.

5.2 Simulation 1: AP performance during congestion without the mechanism

The first simulation results showed how the Access Point (AP) performed with and without the administration system during congestion. As shown in Figures 5-1 and 5-2, the simulation was run using an RTP-video stream against background traffic. In figure 5-1, the RTP-stream was initiated at the beginning of the measurement, and background bulk traffic at 60s marked by the vertical dashed line. The administration system was then de-activated in run no.1 and then activated in run no.2.

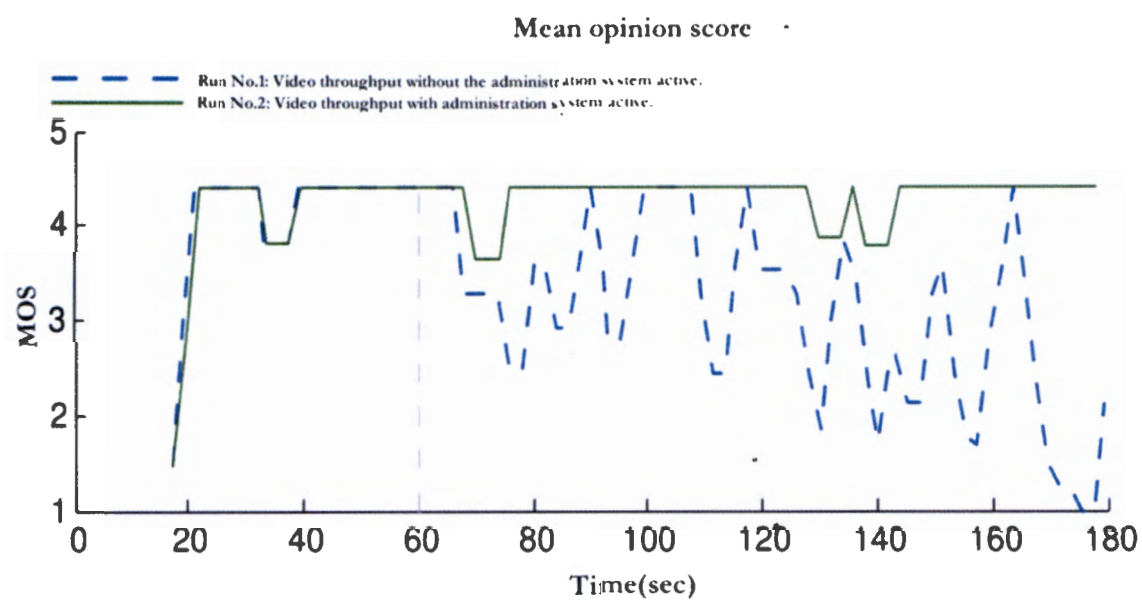


Figure 5-1 RTP video MOS score in run no.1 and 2

5.2.1 Discussion

Without the administration system in run no.1, we started to observe a decline in the quality of the RTP stream once the background traffic was initiated. Although the background traffic operated on TCP, which features congestion control mechanisms, the process remained quite slow and thus packet loss and increased delay started to have effects on the RTP- stream.

In run no.2, the decline in quality was picked up by the administration system, and the RTP-stream was at this point escalated to a higher class. Now these two traffic types were given different bandwidth and dequeuing priorities, which resulted in an improved QoE of the RTP-stream. After repeating this simulation 15 times, we observed that there was an increment in RTP-video MOS score of approximately 0.9 points on average (0.05 standard deviation), and the variance of MOS reduced from 1.92 to 0.21, indicating that the quality did not only improve but also remained more stable during the simulation.

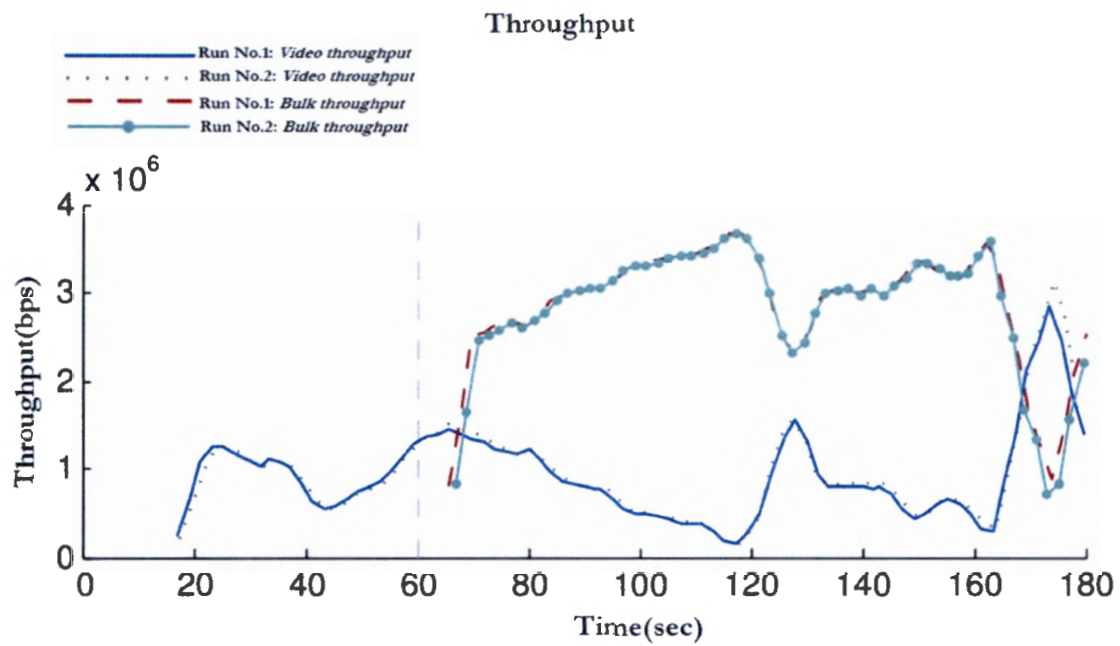


Figure 5-2: Bulk transfer and the throughput of the RTP vide in the two runs

From Figure 5-2, we observed and noted that, while the video stream quality increased tremendously, there was hardly any difference in bulk throughput between these two simulation runs. This was most likely due to the fact that in both cases the bulk transfer suffered from packet losses due to rate adaptation, but in the latter case the reaction was faster because of prioritization. Therefore, it can be noted, especially between 170 and 180s that, the rate of the bulk transfer slowed down faster when the administration system was on, but the difference was marginal since the RTP video only required a slightly larger bandwidth to operate with good quality.

5.3 Simulation 2: User-centric differentiation other than application-centric differentiation.

Another simulation was also executed which demonstrated the differentiation between user classes instead of applications. In this simulation, two RTP-streams were initiated simultaneously, one belonging to a prime user and the other to an ordinary user.

The results of this simulation are illustrated in Figures 5-3 to 5-6. For this simulation, the access point bandwidth was set to 2048 kbps – enough to support only one RTP stream, and the simulation was repeated 15 times, resulting in a standard deviation less than 0.21 for both video streams when the administration system is on.

5.3.1 Discussion

The variable bit rate of the video stream caused two congestion events to occur, one at 25–45s and another at 60–90s when the administration system was on.

Without the administration system, both video signals exhibited poor quality in during these congestion periods. Upon activating the administration system in run no.2, we began to observe a drop in quality and the streams were escalated to higher classes. Despite the fact that the two streams used the same application, it was during the second event of congestion that the prime user gained the preference. Further 15 repetitions of this simulation showed that the quality of

the prime flow increased on average from 2.41 to 3.81, similarly reducing the MOS variance from 2.83 to 1.68.

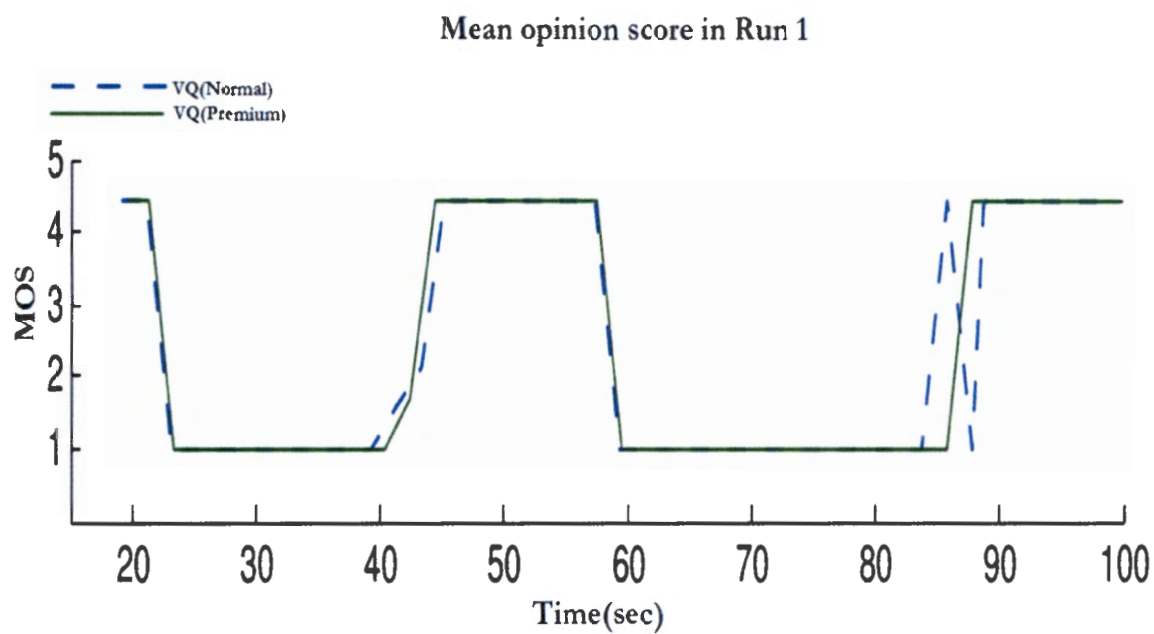


Figure 5-3: RTP videos' MOS score in run no.1

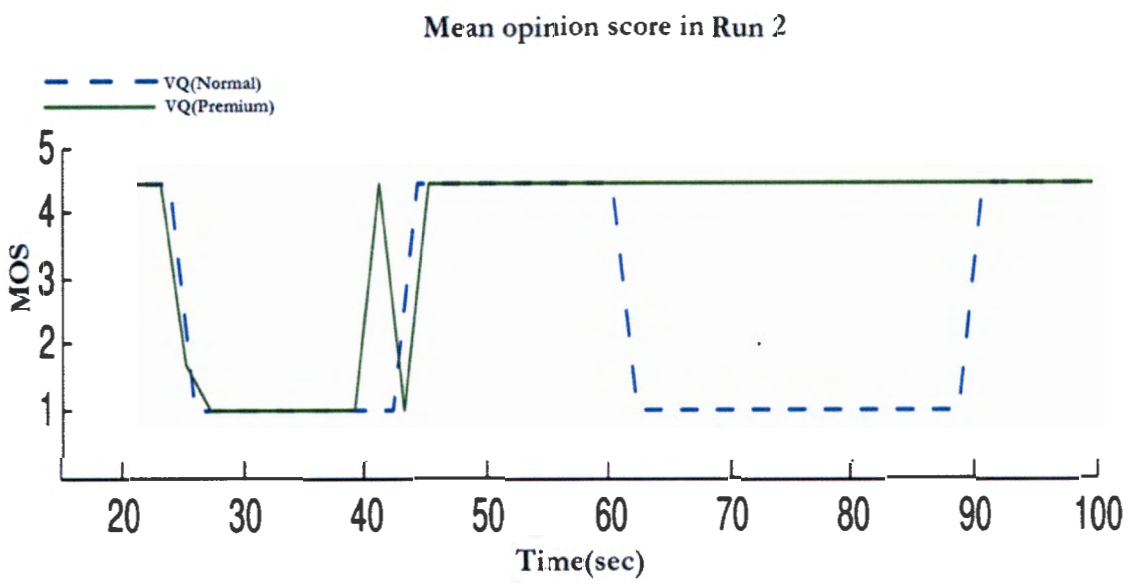


Figure 5-4 RTP videos' MOS score in run no.2

There were no statistically significant changes in the quality of the ordinary users' flow, which can also be noted from Figure 5-3 and 5-4.

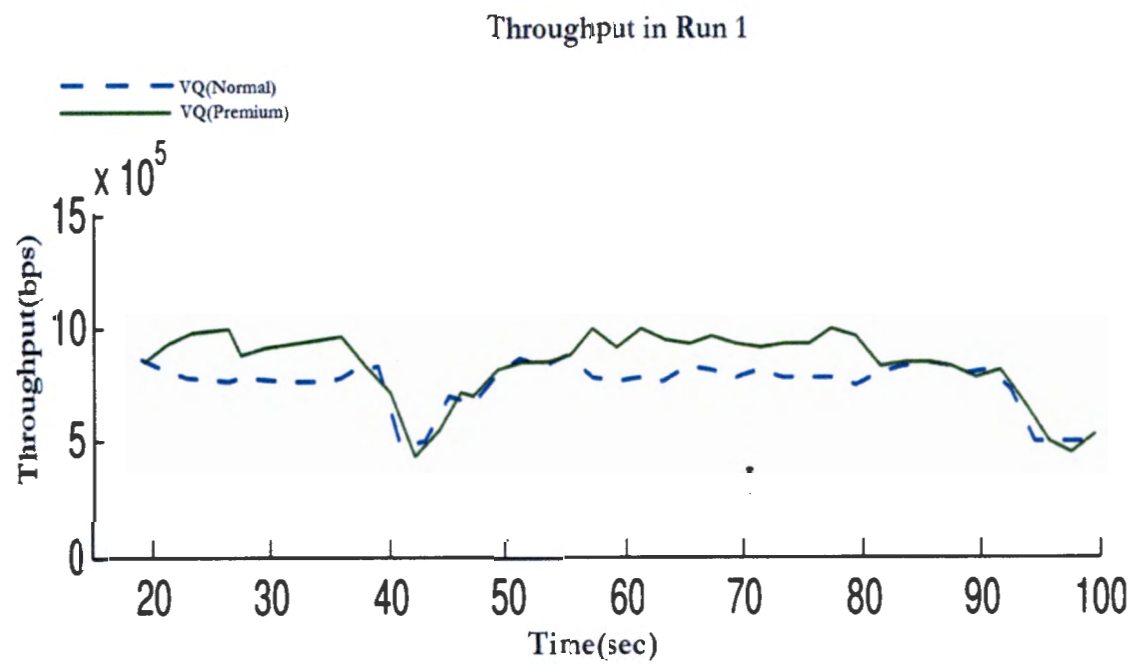


Figure 5-5: RTP videos' throughput in run no.1

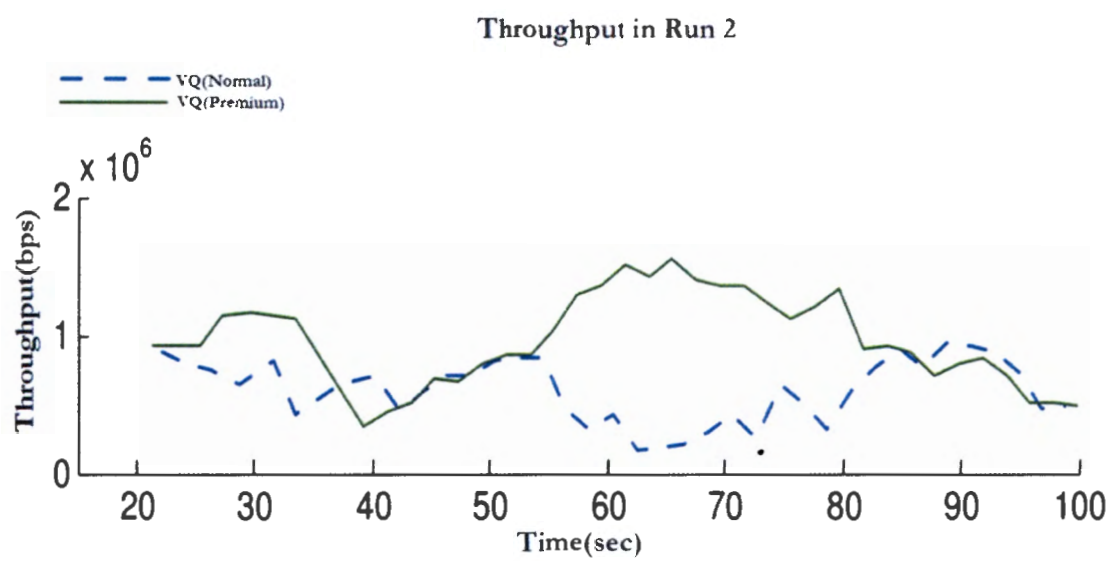


Figure 5-6: RTP videos' throughput run no.2

In Figures 5-5 and 5-6, we made an observation of the variation in the throughput at various points in time. This can be explained by the simple fact that once the quality drops to 1, the throughput of the flow can be distributed to other flows, as the quality is already intolerable and cannot become worse if more bandwidth is distributed elsewhere. This also further explains the drop in the throughput of VQ (normal) and a sudden rise in the throughput of VQ (Premium) as shown in Fig 5-6.

5.4 Simulation 3: Performance evaluation of the administration system.

The next simulation was for performance analysis: this simulation was executed to study how fast the administration system could react to a video stream quality drop. The simulation was performed using an RTP-video stream and assigning a static packet loss rate to the default bulk traffic class with Netem [APPENDIX A]. This simulation was repeated 40 times.

5.4.1 Discussion

In Figure 5-6 we observed that the system picked up the deteriorations in quality in 2.8s (with 1.44s of standard deviation) on average and it had the ability to recover the quality to an acceptable level (i.e., above 4.0) in 4.0s on average (with 1.78s of standard deviation). But during these simulations we also realized that the reaction time was much larger in a congestion event than in this simulation as seen in Figure5-5.

We further examined the issue and found out that when the default class consumes all the bandwidth of the AP, the flow which was escalated to a higher class did not instantly receive the bandwidth the class is supposed to gain. This resulted into heavy packet losses to the promoted flow for approximately 5s, which negatively affected the results somewhat acquired in congestion simulations. This issue was attributed either to Hierarchical Token Bucket (HTB) or Linux iproute2 implementation. We were, however, still able to demonstrate the benefits of the mechanism. Although the AP administration mechanism improved the perceived quality of selected applications, the system was bound to fail nevertheless if there was simply not enough capacity to support all the flows. If we got to know earlier on that the new arriving flow did not

receive satisfactory quality as a result of congestion, we would immediately reject the entry of that flow.

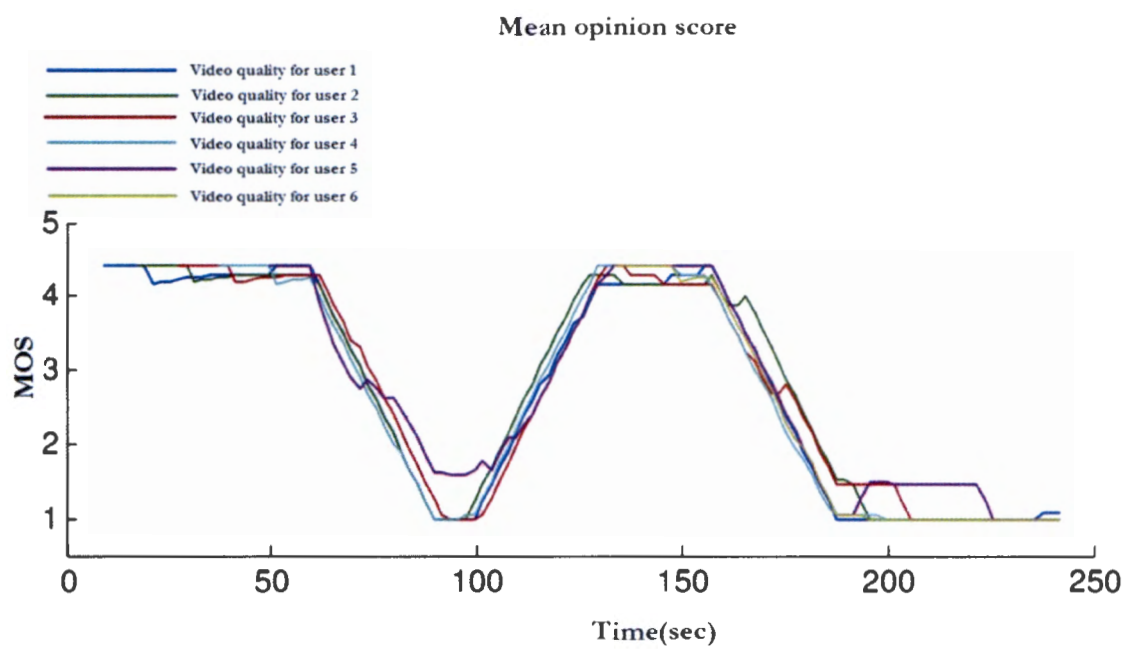


Figure 5-7 MOS scores for the videos when admission control is inactive

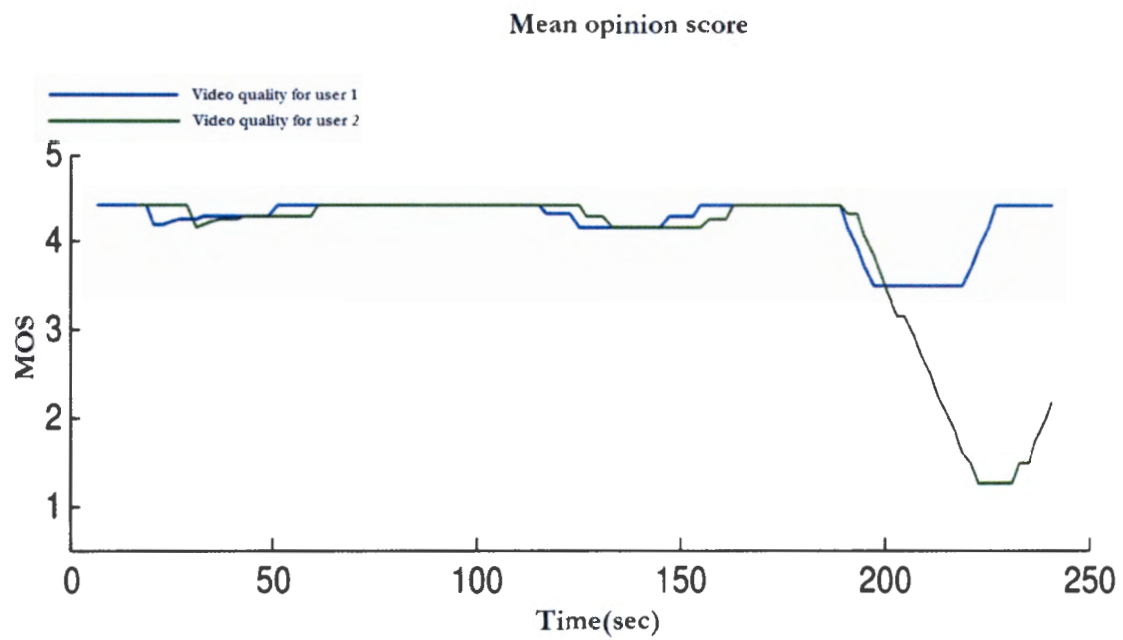


Figure 5-8 MOS scores for the videos when admission control is active

We demonstrated how the admission control works in Figure 5-7 and 5-8. We repeated this simulation 10 times to ensure consistency of the results. In 8 runs, the admission control permitted 2 flows to enter the AP, and in 2 runs 3 flows were permitted. When the administration system was off, the average MOS score of all video streams was 2.29 with a standard deviation of 0.38, while the values were 4.17 MOS score and 0.14 standard deviation when the administration system, and therefore admission control was on.

Figure 5-7 depicts the situation where multiple RTP flows attempted to enter the access point in 10s intervals. With absence of admission control, the flows started to experience poor quality immediately, especially during peak bit rates. Figure 5-8 shows a similar situation with the presence of admission control. Only two flows were permitted to enter the AP, as the admission control algorithm calculated that there was enough capacity to maintain a satisfactory quality. Nonetheless, the bit rates of the videos increased significantly at 190s and the administration system had to step in and differentiate the flows. In order to use performance models in quality prediction, we needed to know how much capacity the new arriving flow would consume. This was difficult to predict beforehand, but there were several ways to make a good guess. One possibility was to perform Deep Packet Inspection (DPI) on packets sent during the handshake procedure to see what kind of data the flow would contain. For an instance, where the data wasn't available we could make various assumptions using the other information available such as the URL field or the history of the client. We predicted the flow throughput based on the average throughput of flows already in the AP, assuming that a new RTP flow will consume approximately the same amount of throughput as all the other flows on average.

5.5 Simulation 4: Effects of traffic and client differentiation.



Finally, we introduced a simulation where both traffic and client differentiation were present.

In this simulation, we had two prime and two ordinary users, each downloading an RTP video.

Additionally, there was a TCP bulk transfer running in the background. This simulation was also run 10 times, resulting in an average standard deviation less than 0.27 in MOS scale.

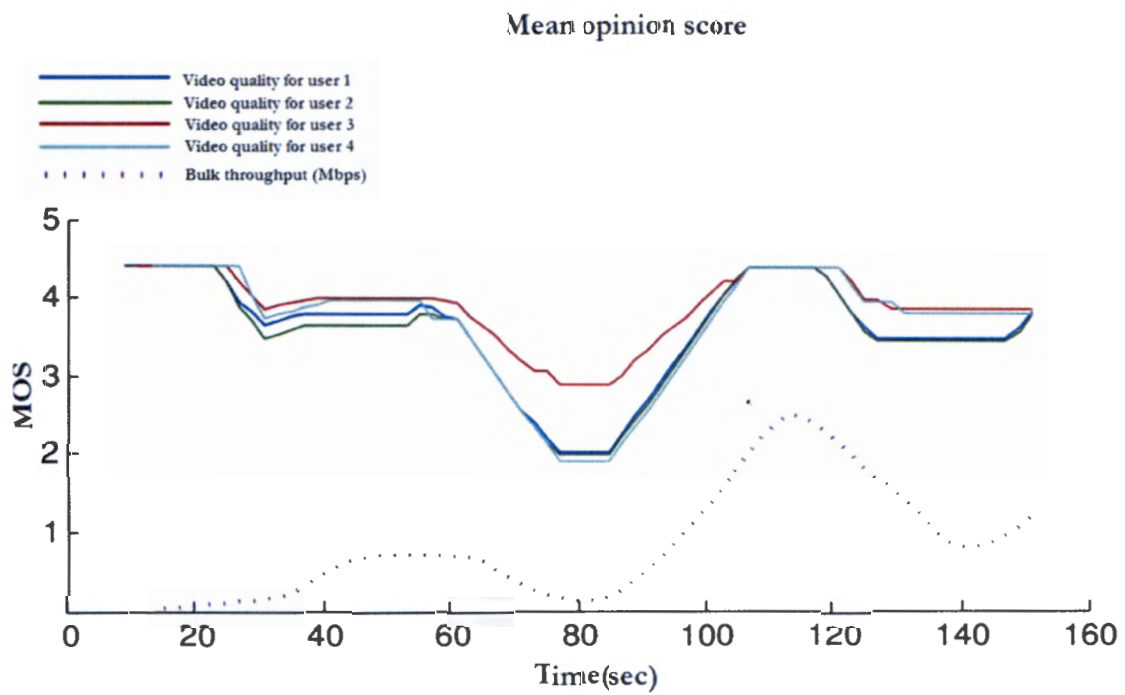


Figure 5-9 MOS scores for the videos when the administration system is inactive

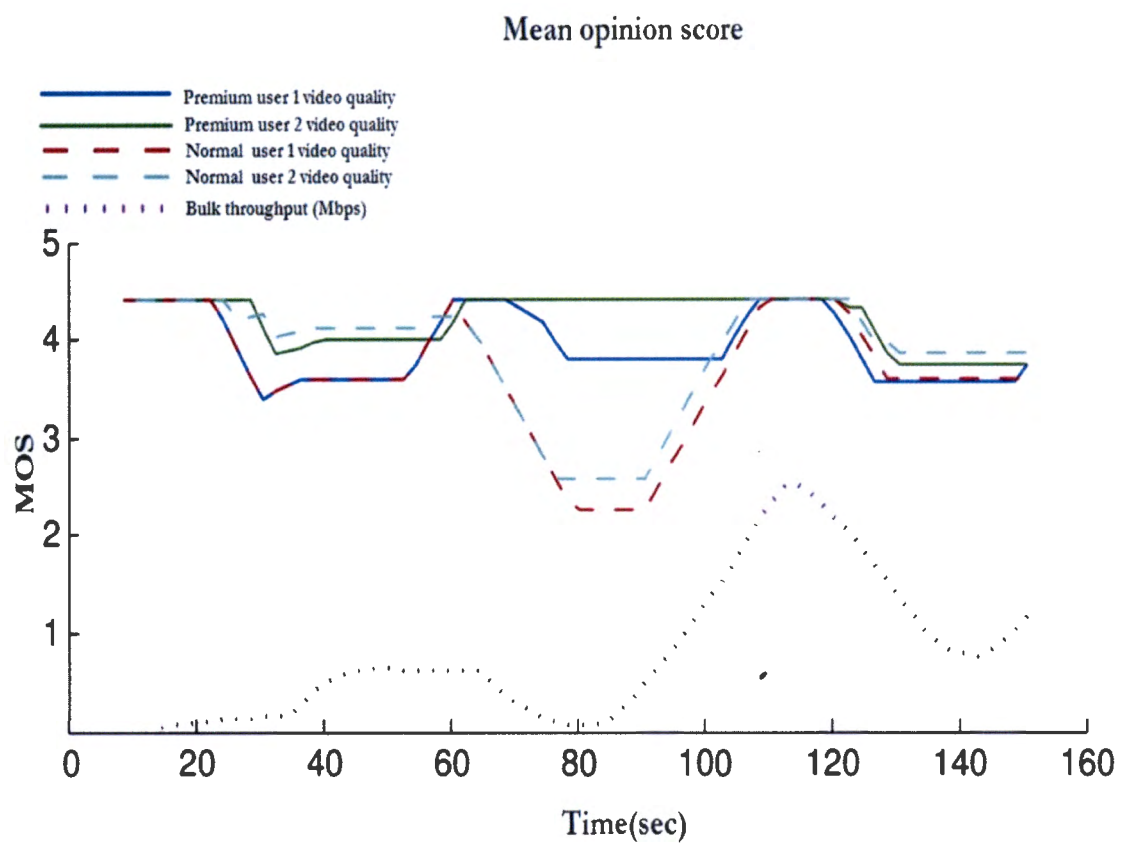


Figure 5-10: MOS scores for the videos when the administration system is active

5.5.1 Discussion

The benefit of having the administration system is shown in Figures 5-9 and 5-10. Figure 5-9 shows the basic scenario without any traffic management. Several flows entered the AP and used approximately an even amount of resources. The TCP bulk transfer adapted to the bandwidth requirements of these flows, which could be seen as a correlation between bulk throughput and video quality (i.e., videos entered a small bit rate phase, which improved the perceived quality and allowed more capacity to the bulk transfer). It should still be noted that the quality of the video streams during high bit rates was unacceptable. The administration system was enabled in Figure 5-10, and an improvement could be seen for streams belonging to prime users. At approximately 30s, the administration system performed a promotion of prime flows to a higher class (but as stated earlier in Chapter 4, the flow promotion during congestion periods still caused more packet losses, hence the slightly worse quality compared to ordinary users). At approximately 70s, the system ran a check using performance models and noticed that it could not fully support the prime flows if both of them resided in the higher class. Therefore, one of these flows was demoted back to the default class. After, one of the flows received the best possible quality during congestion, while the other prime user received satisfactory video quality.

For the ordinary users, there was hardly any difference between having and not having the administration system. From these results we concluded that, performing prioritization was not the only benefit but also the intelligent optimization of the different class loads whenever it became feasible.

5.6 Chapter Review

In this chapter, simulation information in evaluation setup in Chapter 5 was given and results of these simulations presented in detail with the help of graphs. These results were analysed and discussed in the sections of the chapter that followed. The performance of the implemented mechanism and functionalities of traffic differentiation, client differentiation, and admission control and quality optimization were all presented in this chapter. The next chapter presents the conclusions based on the results and observations in this chapter.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Chapter Overview

In this section, we sum up the presentations in all the earlier chapters. We also repeat our research contributions section 1.7 of chapter 1. This chapter also includes a section that is dedicated to highlighting the possibilities of taking this work to the next level in the future. The chapter then closes with a final remark as far as the impact of this research is concerned.

6.2 Research Summary

By the year 2017, video traffic is estimated to have risen to slightly above 90% more than any other global network traffic[172]. As already highlighted in section 1.2 of chapter 1, different users continue to appreciate and use video streaming applications. Various researchers have presented works that concluded that a small drop in the quality of video can easily bring about a recognizable abandonment of a service by users. Much as the mobile devices continue to take the lead in the generation of video traffic, less research has been done to design and implement video streaming infrastructures that have the ability to address the main challenges faced by the streaming applications and their users within the Next Generation Mobile Networks. The ultimate goal of our research was to formulate, design and implement an independent user-centered QoE and QoS monitoring and evaluation mechanism for video streaming in Next Generation Mobile Networks. Concentration was mostly given to the management aspect of the mechanism. In our simulations, we showed how data could be collected from a network and then be refined into knowledge of quality perceived by users and how to take corrective measures when it became necessary to do so. We also showcased this mechanism by introducing and demonstrating network access point management software, built around the

proposed architecture of the mechanism. We then executed several simulations where RTP video streams were subjected to quality-driven network controls.

According to the results we generated, the mechanism performed as expected and each of the simulations resulted in improved quality for the relevant users. It should be noted that, customer differentiation presented the most expected results, which was an important aspect to ISPs. The simulations also showed that the perceived quality of the users could be improved without sacrificing the quality of streams belonging to ordinary users.

Currently, the demonstrated AP administration mechanism supports quality-driven control for RTP streams only. This is due to the availability of a suitable model for RTP video quality and could trivially be extended by adding other models (e.g., for HTTP-based video streaming). Although more applications need to be supported, the need is to support only applications which are most relevant to the ISP and customers. These include, real-time applications such as VoIP, Video and Online games. It is in these applications where quality control is most relevant. HTTP video streaming would be a good app as it has recently become very popular among OTT content providers.

Another limitation in this work was the throughput estimation of new flows. This is as explained in the admission control simulation. Regardless of how the estimation is done, with this kind of administration system it is more important not to under-utilize the access point. The management software can deal with large amounts of traffic, but it is more beneficial for the user to know at the very beginning if the stream will be supported or not. This also enables the system to offer more bandwidth to non-constrained traffic most especially in cases where the AP is free of high priority traffic that may constantly feature “unacceptable” quality. This is why admission control is essential although not seemingly necessary from the point of quality management point of view. When congestion occurs, distributing the resources QoE-

wise becomes an optimization problem. Users might have differing priorities but the same can also be said for applications. When the perceived quality of one application drops below an acceptable level, the application in question may be deemed worthless, thus rendering all the resources allocated to that application useless since they are wasted. Traditional approaches of conflating application quality with the bandwidth allocated to it no longer apply; different multimedia applications require not only capacity but also other guarantees in terms of delay, jitter and packet losses. When these factors are combined with human perception, optimization becomes even more difficult. A momentary good QoE does not mean much to users if their application is operational only occasionally, as seen in the simulations performed. The quality also needs to remain stable. We introduced a mechanism which is not only able to make a good decision for a given time instant, but also to predict the outcome of different decisions and pick the most optimal one. The system is not only able to improve the quality of selected streams, but also to be conservative with the available resources and identify when they are next needed.

6.3 Research contributions

This work managed to address all the research objectives as outlined in Chapter 1 as indicated in the various chapter reviews at the end of each chapter. Chapter 2 was able to extensively address RQ1, RQ2, and RQ3 and also achieve RO1, RO2, and RO3 respectively. Chapter 3 then presented the answers to RQ4 hence achieving RO4. It is in Chapter 5 where the realization of RO4 was seen. It is making this achievement that made it possible for us to make the following research contributions at large:

- a) **Monitoring network QoS and QoE.** The mechanism had the ability to gather the necessary network data and use this data to monitor the network or a user application and be able to take measures where it was deemed necessary.

b) Network traffic management. The mechanism addressed the main challenges in network traffic management, prioritization, and differentiation and resource optimization based mainly on QoE evaluation.

c) Network convergence. The mechanism also supported the convergence in networks with a variety of infrastructure models of access, transport, and control service solutions to cater for services in a multimedia environment. Over all, the mechanism made it possible for network operators to pay more attention to their users as far as QoE is concerned.

d) User expectations and network provider's revenue.

The implemented mechanism had the ability to allow the network operators to effectively address the user expectations in terms of QoE. Meanwhile it also introduced subscriber differentiation as a means of increasing revenue and simplifying resource allocation.

e) Network performance.

Based on the results generated from the various simulations, the desired network performance was achieved and the system performed as envisaged. Below are the various observations that were made before coming to the above conclusion:

- There was a reduction in the compression of data for the effective delivery of lossless data packets;
- There was a reduction in the loss of data packets over the network;
- The mechanism recorded less latencies in data transfer which improved user privacy;
- The mechanism had the ability to administer security and interoperability of data. this was achieved through performing access control on the new traffic flows;



- The same mechanism had the ability to use application types and subscriber classification to select the appropriate priority of the flows;
- The mechanism was able to reduce resolution shortfalls which in turn led to an increase in the user data rate in the process of achieving the desired network conditions and ultimately estimated the current QoE of the user.

6.4 Future Work

Our work mainly concentrated on evaluating user-centric video streaming QoS and QoE in NGMNs. Apart from finding ways of incorporating or factoring in the research limitations identified in Chapter 1 under Section 1.6, more studies can be carried out to include more traffic types and also try to explore the possible impact of QoS/QoE implementations as far as the latest architectures of NGMNs are concerned. Along that line of thought, we are certain that it is of importance to look at complicating these network scenarios further (much nearer to real-life deployments) and test the degree of impact keeping the same parameters as implemented in this work.

In terms of the ever evolving importance of QoE and QoS in NGMNs, it is possible to investigate the merits of QoS/QoE if the various stakeholders, for example, network operators, ISPs, content providers and regulators are also factored into the evaluations especially from the marketing point of view. This is mainly because QoS/QoE is considered a competitive factor as far as telecommunications services are concerned.

Investigating the links between mobile device energy consumption and QoE/QoS could yet be another opportunity to pursue in terms of research and development. Saving energy in any set up of life is always seen as an achievement, therefore, the effects of online time interruptions, the link between the latter and QoE/QoS seek to indicate that it is realistic enough to save

energy in NGMNs infrastructures. It is modelling such relationships that would make the next steps of this work very interesting.

Lastly, we have a vision of modelling resource based QoE/QoS that factors in HCI aspects to improve user perception and experience of consumption of a service in NGMNs.

6.5 Closing Remark

QoE involves bringing together user experience, expectations, emotions and perceptions onto an application and the performance of a network (expressed by QoS parameters). Quantifiable links between QoS and QoE are of paramount importance for the implementation of QoE control management systems with quantifiable QoS parameters. Therefore, the mechanism presented here in our work can be used by a service provider to manage and monitor the QoE of a user and be in a position to effectively respond to complaints of video degradation during a streaming session.

REFERENCES

- [1] X. M. Wu and H. S. Wu, "Based on IPv6 Network and 4G Network System Security Analysis and Research," in *Applied Mechanics and Materials*, 2014, pp. 773-776.
- [2] A. Kuwadekar and K. Al-Begain, "User Centric Quality of Experience Testing for Video on Demand over IMS," in *Computational Intelligence, Communication Systems and Networks, 2009. CICSYN'09. First International Conference on*, 2009, pp. 497-504.
- [3] B. Lewcio, B. Belmudez, T. Enghardt, and S. Möller, "On the way to high-quality video calls in future mobile networks," in *Quality of Multimedia Experience (QoMEX), 2011 Third International Workshop on*, 2011, pp. 43-48.
- [4] S. Panchadcharam, Q. Ni, and G. Taylor, "Evaluation of QoS Parameters for Real-time Rich Media Applications Using a Test-bed Implementation," in *Computer and Information Technology (CIT), 2010 IEEE 10th International Conference on*, 2010, pp. 343-348.
- [5] C. V. N. Index, "The zettabyte era—trends and analysis," *Cisco white paper*, 2013.
- [6] O. Issa, F. Speranza, and T. H. Falk, "Quality-of-experience perception for video streaming services: Preliminary subjective and objective results," in *Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*, 2012, pp. 1-9.
- [7] I. Rec, "P. 10/G. 100 Amendment 1: New Appendix 1-Definition of Quality of Experience (QoE)," *International Telecommunication Union (Jan. 2007)*, 2007.
- [8] Y. Huang, J. Xiao, J. W.-K. Hong, A. Mehaoua, and R. Boutaba, "Event-based estimation of user experience for network video streaming," in *Network Operations and Management Symposium (APNOMS), 2012 14th Asia-Pacific*, 2012, pp. 1-8.
- [9] N. Staelens, S. Moens, W. Van den Broeck, I. Marien, B. Vermeulen, P. Lambert, *et al.*, "Assessing quality of experience of IPTV and video on demand services in real-life environments," *Broadcasting, IEEE Transactions on*, vol. 56, pp. 458-466, 2010.
- [10] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G evolution: HSPA and LTE for mobile broadband*: Academic press, 2010.
- [11] T. Halonen, J. Romero, and J. Melero, *GSM, GPRS and EDGE performance: evolution towards 3G/UMTS*: John Wiley and Sons, 2004.
- [12] B. T. Scheme, "LTE: the evolution of mobile broadband," *IEEE Communications magazine*, vol. 45, 2009.
- [13] G. m. S. Association, "GSA confirms LTE in mainstream, 145 networks in commercial service," ed: On line]. URL Address: http://www.gsacom.com/news/gsa_369.php, accessed 6/8/2014
- [14] S. Sesia, I. Toufik, and M. Baker, *LTE: the UMTS long term evolution*: Wiley Online Library, 2009.

- [15] I. F. Akyildiz, J. Xie, and S. Mohanty, "A survey of mobility management in next-generation all-IP-based wireless systems," *Wireless Communications, IEEE*, vol. 11, pp. 16-28, 2004.
- [16] C. E. Perkins, S. R. Alpert, and B. Woolf, *Mobile IP; Design Principles and Practices*: Addison-Wesley Longman Publishing Co., Inc., 1997.
- [17] R. Bolla, R. Bruschi, F. Davoli, and A. Ranieri, "Energy-aware performance optimization for next-generation green network equipment," in *Proceedings of the 2nd ACM SIGCOMM workshop on Programmable routers for extensible services of tomorrow*, 2009, pp. 49-54.
- [18] B. Li, M. Hamdi, D. Iang, X.-R. Cao, and Y. T. Hou, "QoS enabled voice support in the next generation Internet: issues, existing approaches and challenges," *Communications Magazine, IEEE*, vol. 38, pp. 54-61, 2000.
- [19] C.-F. Wu, L.-T. Lee, H.-Y. Chang, and D.-F. Tao, "A novel call admission control policy using mobility prediction and throttle mechanism for supporting QoS in wireless cellular networks," *Journal of Control Science and Engineering*, vol. 2011, p. 21, 2011.
- [20] M. Ismail and W. Zhuang, "A distributed multi-service resource allocation algorithm in heterogeneous wireless access medium," *Selected Areas in Communications, IEEE Journal on*, vol. 30, pp. 425-432, 2012.
- [21] S. Kajioka, N. Wakamiya, and M. Murata, "Autonomous and adaptive resource allocation among multiple nodes and multiple applications in heterogeneous wireless networks," *Journal of Computer and System Sciences*, vol. 78, pp. 1673-1685, 2012.
- [22] K. Akkarajitsakul, E. Hossain, and D. Niyato, "Distributed resource allocation in wireless networks under uncertainty and application of Bayesian game," *Communications Magazine, IEEE*, vol. 49, pp. 120-127, 2011.
- [23] Q. Song, Y. Cui, and L. Li, "An Adaptive Bandwidth Management Scheme for Heterogeneous Wireless Networks," in *Recent Advances in Computer Science and Information Engineering*, ed: Springer, 2012, pp. 431-437.
- [24] S. Gupta, "Wireless communication," *IJSER, USA*, vol. 3, pp. 1-8, 2012.
- [25] H. Zang and T. E. Dawson, "Telecom fraud detection using social pattern," ed: Google Patents, 2013.
- [26] Y. Park and T. Park, "A survey of security threats on 4G networks," in *Globecom Workshops, 2007 IEEE*, 2007, pp. 1-6.
- [27] T. Ahmed, G. Buridant, and A. Mehaoua, "Encapsulation and marking of MPEG-4 video over IP differentiated services," in *Computers and Communications, 2001. Proceedings. Sixth IEEE Symposium on*, 2001, pp. 346-352.
- [28] T. Stockhammer, "Dynamic adaptive streaming over HTTP--: standards and design principles," in *Proceedings of the second annual ACM conference on Multimedia systems*, 2011, pp. 133-144.

- [29] J. Erman, A. Gerber, K. Ramadrihnan, S. Sen, and O. Spatscheck, "Over the top video: the gorilla in cellular networks," in *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, 2011, pp. 127-136.
- [30] E. S. Westberg, "Understanding Viewer Abandonment Trends in Short-Form Online Video Content," *Broadcasting, IEEE Transactions on*, vol. 47, pp. 168-172, 2012.
- [31] A. Khan, L. Sun, E. Jammeh, and E. Ifeachor, "Quality of experience-driven adaptation scheme for video applications over wireless networks," *IET communications*, vol. 4, pp. 1337-1347, 2010.
- [32] I. E. Richardson, *The H. 264 advanced video compression standard*: John Wiley and Sons, 2011.
- [33] J.-R. Ohm and G. J. Sullivan, "High efficiency video coding: the next frontier in video compression [Standards in a Nutshell]," *Signal Processing Magazine, IEEE*, vol. 30, pp. 152-158, 2013.
- [34] A. D. Gupte, B. Amrutur, M. M. Mehendale, A. V. Rao, and M. Budagavi, "Memory bandwidth and power reduction using lossy reference frame compression in video encoding," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 21, pp. 225-230, 2011.
- [35] S. Issa and O. O. Khalifa, "Performance analysis of Dirac video codec with H. 264/AVC," in *Computer and Communication Engineering (ICCCE), 2010 International Conference on*, 2010, pp. 1-6.
- [36] S. Subramanya, "Video containers: a system for the on-demand storage, delivery, and management of television programs," in *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on*, 2000, pp. 1245-1249.
- [37] C. V. N. I. Cisco, "Global mobile data traffic forecast update, 2013–2018," *white paper*, 2014.
- [38] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-advanced for mobile broadband*: Academic press, 2013.
- [39] K. D. Singh, A. Ksentini, and B. Marienval, "Quality of experience measurement tool for SVC video coding," in *Communications (ICC), 2011 IEEE International Conference on*, 2011, pp. 1-5.
- [40] S. R. Haddock, J. N. Chueh, S. T. Merchant, A. H. Smith, and M. Yip, "Policy based quality of service," ed: Google Patents, 2000.
- [41] I. Ketykó, K. De Moor, W. Joseph, L. Martens, and L. De Marez, "Performing QoE-measurements in an actual 3G network," in *Broadband Multimedia Systems and Broadcasting (BMSB), 2010 IEEE International Symposium on*, 2010, pp. 1-6.
- [42] X. Yu, H. Chen, W. Zhao, and L. Xie, "No-Reference QoE Prediction Model for Video Streaming Service in 3G Networks," in *Wireless Communications, Networking and Mobile Computing (WiCOM), 2012 8th International Conference on*, 2012, pp. 1-4.

- [43] T. N. Minhas, "Network impact on quality of experience of mobile video," *Blekinge Institute of Technology*, 2012.
- [44] A. Alvarez, S. Cabrero, X. G. Pañeda, R. Garcia, D. Melendi, and R. Orea, "A flexible QoE framework for video streaming services," in *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*, 2011, pp. 1226-1230.
- [45] A. Khan, L. Sun, E. Ifeachor, J. O. Fajardo, and F. Liberal, "Video quality prediction model for H. 264 video over UMTS networks and their application in mobile video streaming," in *Communications (ICC), 2010 IEEE International Conference on*, 2010, pp. 1-5.
- [46] C.-R. Lan, H.-H. Lu, C.-W. Yi, and C.-C. Tseng, "A P2P HD live video streaming system," in *Multimedia Technology (ICMT), 2011 International Conference on*, 2011, pp. 475-478.
- [47] C. Zheng and Z. Ming, "An efficient video similarity search strategy for video-on-demand systems," in *Broadband Network and Multimedia Technology, 2009. IC-BNMT'09. 2nd IEEE International Conference on*, 2009, pp. 174-178.
- [48] A. C. Begen, T. Akgul, and M. Baugher, "Watching video over the web: Part 1: Streaming protocols," *Internet Computing, IEEE*, vol. 15, pp. 54-63, 2011.
- [49] D. Wu, Y. T. Hou, W. Zhu, Y.-Q. Zhang, and J. M. Peha, "Streaming video over the Internet: approaches and directions," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 11, pp. 282-300, 2001.
- [50] Z. Wang, H. R. Sheikh, and A. C. Bovik, "Objective video quality assessment," *The handbook of video databases: design and applications*, pp. 1041-1078, 2003.
- [51] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, "Study of subjective and objective quality assessment of video," *Image Processing, IEEE transactions on*, vol. 19, pp. 1427-1441, 2010.
- [52] T. Stockhammer, P. Fröjdh, I. Sodagar, and S. Rhyu, "Information technology—MPEG systems technologies—Part 6: Dynamic adaptive streaming over HTTP (DASH)," *ISO/IEC, MPEG Draft International Standard*, 2011.
- [53] S. Winkler and P. Mohandas, "The evolution of video quality measurement: from PSNR to hybrid metrics," *Broadcasting, IEEE Transactions on*, vol. 54, pp. 660-668, 2008.
- [54] S. Winkler, "Video quality and beyond," in *Proc. European Signal Processing Conference*, 2007, pp. 3-7.
- [55] K. Kilkki, "Quality of Experience in Communications Ecosystem," *J. UCS*, vol. 14, pp. 615-624, 2008.
- [56] K. Knightson, N. Morita, and T. Towle, "NGN architecture: generic principles, functional architecture, and implementation," *IEEE Communications Magazine*, vol. 43, pp. 49-56, 2005.
- [57] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam, "Objective video quality assessment methods: A classification, review, and performance comparison," *Broadcasting, IEEE Transactions on*, vol. 57, pp. 165-182, 2011.

- [58] S. Jelassi, G. Rubino, H. Melvin, H. Youssef, and G. Pujolle. "Quality of experience of VoIP service: a survey of assessment approaches and open issues." *Communications Surveys and Tutorials, IEEE*, vol. 14, pp. 491-513, 2012.
- [59] W. Lin and C.-C. J. Kuo. "Perceptual visual quality metrics: A survey." *Journal of Visual Communication and Image Representation*, vol. 22, pp. 297-312, 2011.
- [60] R. Serral-Gracià, E. Cerqueira, M. Curado, M. Yannuzzi, E. Monteiro, and X. Masip-Bruin. "An overview of quality of experience measurement challenges for video applications in IP networks," in *Wired/Wireless Internet Communications*, ed: Springer, 2010, pp. 252-263.
- [61] E. Kafetzakis, H. Koumaras, M. A. Kourtis, and V. Koumaras. "QoE4CLOUD: A QoE-driven multidimensional framework for cloud environments," in *Telecommunications and Multimedia (TEMU), 2012 International Conference on*, 2012, pp. 77-82.
- [62] T. Hobfeld, R. Schatz, M. Varela, and C. Timmerer. "Challenges of QoE management for cloud applications," *Communications Magazine, IEEE*, vol. 50, pp. 28-36, 2012.
- [63] G. Gómez, J. Lorca, R. García, and Q. Pérez, "Towards a qoe-driven resource control in lte and lte-a networks," *Journal of Computer Networks and Communications*, vol. 2013, 2013.
- [64] J.-O. Fajardo, F. Liberal, I.-H. Mkwawa, L. Sun, and H. Koumaras, "QoE-driven dynamic management proposals for 3G VoIP services," *Computer Communications*, vol. 33, pp. 1707-1724, 2010.
- [65] M. Mu, E. Cerqueira, F. Boavida, and A. Mauthe, "Quality of experience management framework for real-time multimedia applications." *International Journal of Internet Protocol Technology*, vol. 4, pp. 54-64, 2009.
- [66] F. Matos, A. Matos, P. Simoes, and E. Monteiro, "QoS adaptation in inter-domain services," in *Integrated Network Management (IM), 2011 IFIP/IEEE International Symposium on*, 2011, pp. 257-264.
- [67] A. Vakili and J.-C. Grégoire, "QoE management for video conferencing applications," *Computer Networks*, vol. 57, pp. 1726-1738, 5/8/ 2013.
- [68] H.-J. Kim, K.-H. Lee, and J. Zhang, "In-service feedback QoE framework," in *Communication Theory, Reliability, and Quality of Service (CTRQ), 2010 Third International Conference on*, 2010, pp. 135-138.
- [69] J. Lloret, M. Garcia, M. Atenas, and A. Canovas. "A QoE management system to improve the IPTV network," *International Journal of Communication Systems*, vol. 24, pp. 118-138, 2011.
- [70] K. Piamrat, A. Ksentini, C. Viho, and J.-M. Bonnin, "Qoe-aware admission control for multimedia applications in ieee 802.11 wireless networks," in *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*, 2008, pp. 1-5.
- [71] E. Gallo, M. Siller, and J. Woods. "An ontology for the quality of experience framework," in *Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on*, 2007, pp. 1540-1544.

- [72] A. Bouch, M. A. Sasse, and H. DeMeer, "Of packets and people: a user-centered approach to quality of service," in *Quality of Service, 2000. IWQOS. 2000 Eighth International Workshop on*, 2000, pp. 189-197.
- [73] K. Kilkki, *Differentiated services for the Internet*: Macmillan Publishing Co., Inc., 1999.
- [74] N. Muhammad, D. Chiavelli, D. Soldani, and M. Li, "Introduction," *QoS and QoE Management in UMTS Cellular Systems*, pp. 1-8, 2006.
- [75] I. Rec, "E. 800, Terms and definitions related to quality of service and network performance including dependability," *International Telecommunication Union*, vol. 22, 1994.
- [76] I. Recommendation, "E. 800: Terms and definitions related to quality of service and network performance including dependability," *ITU-T August 1994*, 1994.
- [77] E. N. Aspects, "General Aspects of Quality of Service (QoS) and Network Performance (NP)," Technical Report ETR003 1994.
- [78] J. Gozdecki, A. Jajszczyk, and R. Stankiewicz, "Quality of service terminology in IP networks," *Communications Magazine, IEEE*, vol. 41, pp. 153-159, 2003.
- [79] T. Itu, "G. 1010: End-user multimedia QoS categories," Technical report, ITU2001.
- [80] G. Recommendation, "1010 End-user multimedia QoS categories," *ITU-T, November*, 2001.
- [81] N. Bayer, M. C. De Castro, P. Dely, A. Kassler, Y. Koucheryavy, P. Mitoraj, *et al.*, "VoIP service performance optimization in pre-IEEE 802.11 s Wireless Mesh Networks," in *Circuits and Systems for Communications, 2008. ICCSC 2008. 4th IEEE International Conference on*, 2008, pp. 75-79.
- [82] T. ITU, "Terms and definitions related to quality of service and network performance including dependability," *Recommendation E*, vol. 800, 1994.
- [83] I. Rec, "P. 10 (2007) Vocabulary for performance and Quality of Service," *International Telecommunication Union, Geneva*, vol. 22, 2007.
- [84] S. Iekin, K. Wac, M. Fiedler, L. Janowski, J.-H. Hong, and A. K. Dey, "Factors influencing quality of experience of commonly used mobile applications," *Communications Magazine, IEEE*, vol. 50, pp. 48-56, 2012.
- [85] K. De Moor, W. Joseph, I. Ketykó, E. Tanghe, T. Deryckere, L. Martens, *et al.*, "Linking users' subjective QoE evaluation to signal strength in an IEEE 802.11 b/g wireless LAN environment," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, p. 6, 2010.
- [86] V. Roto, E. Law, A. Vermeeren, and J. Hoonhout, "User experience white paper," *Bringing clarity to the concept of user experience*, 2011.
- [87] T. M. O'Neil, "Quality of experience and quality of service for IP video conferencing," *Polycom Video Communications, Milpitas, CA, USA, White paper*, 2002.
- [88] P. Odence, "Assuring QoE on next generation networks," *Annual Review of Communications*, 2004.

- [89] M. Siller and J. Woods, "QoS arbitration for improving the QoE in multimedia transmission," in *IEE conference publication*, 2003, pp. 238-241.
- [90] D. Soldani, "Means and methods for collecting and analyzing QoE measurements in wireless networks," in *Proceedings of the 2006 International Symposium on on World of Wireless, Mobile and Multimedia Networks*, 2006, pp. 531-535.
- [91] A. Van Ewijk, J. De Vriendt, and L. Finizola, "Quality of Service for IMS on Fixed Networks. Business Models and Drivers for Next-Generation IMS Services," ed: International Engineering Consortium, USA, 2007.
- [92] K. De Moor, I. Ketyko, W. Joseph, T. Deryckere, L. De Marez, L. Martens, *et al.*, "Proposed framework for evaluating quality of experience in a mobile, testbed-oriented living lab setting," *Mobile Networks and Applications*, vol. 15, pp. 378-391, 2010.
- [93] S. Baraković, J. Baraković, and H. Bajrić, "QoE dimensions and QoE measurement of NGN services," in *Proceedings of the 18th Telecommunications Forum, TELFOR 2010*, 2010.
- [94] P. Le Callet, S. Möller, and A. Perkis, "Qualinet white paper on definitions of quality of experience," *European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)*, 2012.
- [95] D. Durkee, "Why cloud computing will never be free," *Queue*, vol. 8, p. 20, 2010.
- [96] H. Rifai, S. Mohammed, and A. Mellouk, "A brief synthesis of QoS-QoE methodologies," in *Programming and Systems (ISPS), 2011 10th International Symposium on*, 2011, pp. 32-38.
- [97] R. ITU-T, "800: Methods for subjective determination of transmission quality," *International Telecommunication Union-Telecommunication Standardisation Sector (ITU-T)*, 1996.
- [98] P. Ballon, J. Pierson, and S. Delaere, "Fostering innovation in networked communications: Test and experimentation," *Designing for networked communications: strategies and development*, p. 137, 2007.
- [99] L. De Marez and G. Verleye, "Innovation diffusion: The need for more accurate consumer insight. Illustration of the PSAP scale as a segmentation instrument," *Journal of Targeting, Measurement and Analysis for Marketing*, vol. 13, pp. 32-49, 2004.
- [100] A. Følstad, "Towards a living lab for the development of online community services," *The Electronic Journal for Virtual Organizations and Networks*, vol. 10, 2008.
- [101] S. S. Intille, E. M. Tapia, J. Rondoni, J. Beaudin, C. Kukla, S. Agarwal, *et al.*, "Tools for studying behavior and technology in natural settings," in *UbiComp 2003: Ubiquitous Computing*, 2003, pp. 157-174.
- [102] M. P. De Leon, M. Eriksson, S. Balasubramaniam, and W. Donnelly, "Creating a distributed mobile networking testbed environment-through the living labs approach," in *Testbeds and Research Infrastructures for the Development of Networks and Communities*, 2006. *TRIDENTCOM 2006. 2nd International Conference on*, 2006, pp. 5 pp.-139.

- [103] L. De Marez and K. De Moor, "The challenge of user-and QoE-centric research and product development in today's ICT-environment," *Observatorio (OBS*)*, vol. 1, 2007.
- [104] B. Fehnert and A. Kosagowsky, "Measuring user experience: complementing qualitative and quantitative assessment," in *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, 2008, pp. 383-386.
- [105] A. Perkis, S. Munkeby, and O. I. Hillestad, "A model for measuring quality of experience." in *Signal Processing Symposium, 2006. NORSIG 2006. Proceedings of the 7th Nordic*, 2006, pp. 198-201.
- [106] A. Schmidt, M. Strohbach, K. Van Laerhoven, A. Friday, and H.-W. Gellersen, "Context acquisition based on load sensing," in *UbiComp 2002: Ubiquitous Computing*, ed: Springer, 2002, pp. 333-350.
- [107] V.-P. Niitamo, S. Kulkki, M. Eriksson, and K. A. Hribernik, "State-of-the-art and good practice in the field of living labs," in *Proceedings of the 12th International Conference on Concurrent Enterprising: Innovative Products and Services through Collaborative Networks. Italy: Milan*, 2006, pp. 26-28.
- [108] T. Farley, "Mobile telephone history," *Teletronikk*, vol. 101, p. 22, 2005.
- [109] A. Goldsmith, "5G and beyond: what lies ahead for wireless system design," *PPT Presentation at BUPT*, 2014.
- [110] V. K. Bhargava and A. Leon-Garcia, "Green cellular networks: A survey, some research issues and challenges," in *Communications (QBSC), 2012 26th Biennial Symposium on*, 2012, pp. 1-2.
- [111] I. F. Akyildiz, D. M. Gutierrez-Estevez, R. Balakrishnan, and E. Chavarria-Reyes, "LTE-Advanced and the evolution to Beyond 4G (B4G) systems," *Physical Communication*, vol. 10, pp. 31-60, 2014.
- [112] C. V. N. Index, "Global mobile data traffic forecast update, 2010-2015," *White Paper, February*, 2011.
- [113] L. Ericsson, "More than 50 billion connected devices," *White Paper*, 2011.
- [114] D. Ehsan, M. Fick, M. S. Aminian, and Y. Dong, "The 5th generation of mobile networks and wireless systems," 2014.
- [115] S. Patel, M. Chauhan, and K. Kapadiya, "5G: future mobile technology-vision 2020," *International Journal of Computer Applications*, vol. 54, 2012.
- [116] V. SA. (Feb 25, 2015, 1 July 2015). *Vodacom Baby crying network buffering*.
- [117] B. Mobile, "Mobile Analytics Report," ed: June, 2011.
- [118] L. Pocock, "A survey of the telephone transmission-rating problem," *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, vol. 95, pp. 253-265, 1948.

- [119] M. Siller and J. Woods, "Improving quality of experience for multimedia services by QoS arbitration on a QoE framework," in *Proc. of the 13th Packed Video Workshop 2003*, 2003.
- [120] C. Assigned, "9.2.9 E-RAB Level QoS Parameters," *LTE; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 Application Protocol (X2AP)(3GPP TS 36.423 version 11.6.0 Release 11)*, 2013.
- [121] P. Brooks and B. Hestnes, "User measures of quality of experience: why being objective and quantitative is important," *Network, IEEE*, vol. 24, pp. 8-13, 2010.
- [122] B. Bauer and A. S. Patrick, "A human factors extension to the seven-layer OSI reference model," *Retrieved January*, vol. 6, p. 2004, 2004.
- [123] D. Lopez, F. Gonzalez, L. Bellido, and A. Alonso, "Adaptive multimedia streaming over IP based on customer oriented metrics," in *Computer Networks, 2006 International Symposium on*, 2006, pp. 185-191.
- [124] R. Callon and E. C. Rosen, "Multiprotocol label switching architecture," 2001.
- [125] R. Braden, D. Clark, and S. Shenker, "Integrated services in the internet architecture: an overview," 2070-1721, 1994.
- [126] M. Carlson, E. Davies, U. Nortel, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services Status of this Memo This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited," 1998.
- [127] T. Wang, A. Pervez, and H. Zou, "VQM-based QoS/QoE mapping for streaming video," in *Broadband Network and Multimedia Technology (IC-BNMT), 2010 3rd IEEE International Conference on*, 2010, pp. 807-812.
- [128] I. Rec, "P. 800.1," "Mean Opinion Score (MOS) Terminology "," *International Telecommunication Union, Geneva*, 2006.
- [129] O. Nemethova, M. Ries, E. Siffel, and M. Rupp, "Quality assessment for H. 264 coded low-rate and low-resolution video sequences," in *Communications, Internet, and Information Technology*, 2004, pp. 508-512.
- [130] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, pp. 600-612, 2004.
- [131] I. Rec, "P. 862.1: Mapping function for transforming P. 862 raw result scores to MOS-LQO," *International Telecommunication Union, Geneva*, vol. 24, 2003.
- [132] D. Hands, O. V. Barriac, and F. Telecom, "Standardization activities in the ITU for a QoE assessment of IPTV," *IEEE Communications Magazine*, vol. 46, pp. 78-84, 2008.
- [133] J. Korhonen, N. Burini, J. You, and E. Nadernejad, "How to evaluate objective video quality metrics reliably," in *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*, 2012, pp. 57-62.

- [134] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *Broadcasting, IEEE Transactions on*, vol. 50, pp. 312-322, 2004.
- [135] S. Wulf and U. Zolzer, "Full-reference video quality assessment on high-definition video content," in *Signal Processing and Communication Systems (ICSPCS), 2012 6th International Conference on*, 2012, pp. 1-10.
- [136] F. Agboma and A. Liotta, "QoE-aware QoS management," in *Proceedings of the 6th international conference on advances in mobile computing and multimedia*, 2008, pp. 111-116.
- [137] G. Fechner, "Elements of psychophysics," 1965.
- [138] W. R. Klecka, *Discriminant analysis*: Sage, 1980.
- [139] M. Elkotob, D. Grandlund, K. Andersson, and C. Ahlund, "Multimedia QoE optimized management using prediction and statistical learning," in *Local Computer Networks (LCN), 2010 IEEE 35th Conference on*, 2010, pp. 324-327.
- [140] V. Menkovski, A. Oredope, A. Liotta, and A. C. Sánchez, "Predicting quality of experience in multimedia streaming," in *Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia*, 2009, pp. 52-59.
- [141] J. R. Quinlan, "Induction of decision trees," *Machine learning*, vol. 1, pp. 81-106, 1986.
- [142] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, pp. 199-222, 2004.
- [143] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Ijcai*, 1995, pp. 1137-1145.
- [144] V. A. Machado, N. Carlos, R. S. O. Silva, A. M. Melo, M. Silva, C. R. Francês, *et al.*, "A new proposal to provide estimation of QoS and QoE over WiMAX networks," ed: LatinCom, 2011.
- [145] M. S. Mushtaq, B. Augustin, and A. Mellouk, "Empirical study based on machine learning approach to assess the QoS/QoE correlation," in *Networks and Optical Communications (NOC), 2012 17th European Conference on*, 2012, pp. 1-7.
- [146] H. Du, C. Guo, Y. Liu, and Y. Liu, "Research on relationship between QoE and QoS based on BP Neural Network," in *Network Infrastructure and Digital Content, 2009. IC-NIDC 2009. IEEE International Conference on*, 2009, pp. 312-315.
- [147] P. Frank and J. Incera, "A neural network based test bed for evaluating the quality of video streams in IP networks," in *Electronics, Robotics and Automotive Mechanics Conference, 2006*, 2006, pp. 178-183.
- [148] P. Calyam, P. Chandrasekaran, G. Trueb, N. Howes, R. Ramnath, D. Yu, *et al.*, "Multi-resolution multimedia QoE models for IPTV applications," *International Journal of Digital Multimedia Broadcasting*, vol. 2012, 2012.
- [149] A. Khan, L. Sun, E. Ifeachor, J.-O. Fajardo, F. Liberal, and H. Koumaras, "Video quality prediction models based on video content dynamics for H. 264 video over UMTS networks," *International Journal of Digital Multimedia Broadcasting*, vol. 2010, 2010.

- [150] T. Hoßfeld, M. Seufert, M. Hirth, T. Zinner, P. Tran-Gia, and R. Schatz, "Quantification of YouTube QoE via crowdsourcing," in *Multimedia (ISM), 2011 IEEE International Symposium on*, 2011, pp. 494-499.
- [151] M. Hirth, T. Hoßfeld, and P. Tran-Gia, "Anatomy of a crowdsourcing platform-using the example of microworkers. com," in *Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2011 Fifth International Conference on*, 2011, pp. 322-329.
- [152] T. Hobfeld, R. Schatz, and S. Egger, "SOS: The MOS is not enough!," in *Quality of Multimedia Experience (QoMEX), 2011 Third International Workshop on*, 2011, pp. 131-136.
- [153] A. Watson and M. A. Sasse, "Measuring perceived quality of speech and video in multimedia conferencing applications," in *Proceedings of the sixth ACM international conference on Multimedia*, 1998, pp. 55-60.
- [154] B. Han, X. Zhang, Y. Qi, Y. Gao, and D. Yang, "QoE model based optimization for streaming media service considering equipment and environment factors," *Wireless Personal Communications*, vol. 66, pp. 595-612, 2012.
- [155] I. Khan and N. Crespi, "Quantitative and qualitative assessment of qoe for multimedia services in wireless environment," in *Proceedings of the 4th Workshop on Mobile Video*, 2012, pp. 7-12.
- [156] L. Skorin-Kapov and M. Varela, "A multi-dimensional view of QoE: the ARCU model," in *MIPRO, 2012 Proceedings of the 35th International Convention*, 2012, pp. 662-666.
- [157] M. Fiedler, T. Hossfeld, and P. Tran-Gia, "A generic quantitative relationship between quality of experience and quality of service," *Network, IEEE*, vol. 24, pp. 36-41, 2010.
- [158] H. L. Kim and S. G. Choi, "A study on a QoS/QoE correlation model for QoE evaluation on IPTV service," in *Advanced Communication Technology (ICACT), 2010 The 12th International Conference on*, 2010, pp. 1377-1382.
- [159] H. Koumaras, C.-H. Lin, C.-K. Shieh, and A. Kourtis, "A framework for end-to-end video quality prediction of MPEG video," *Journal of Visual Communication and Image Representation*, vol. 21, pp. 139-154, 2010.
- [160] R. K. Mok, E. W. Chan, and R. K. Chang, "Measuring the quality of experience of HTTP video streaming," in *Integrated Network Management (IM), 2011 IFIP/IEEE International Symposium on*, 2011, pp. 485-492.
- [161] J. W. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage publications, 2013.
- [162] J. Hu, S. Choudhury, and J. D. Gibson, "Video capacity of WLANs with a multiuser perceptual quality constraint," *Multimedia, IEEE Transactions on*, vol. 10, pp. 1465-1478, 2008.
- [163] J. Collis and R. Hussey, "Business Research (3: e uppl.)," *Hampshire: Palgrave Macmillian*, 2009.

- [164] G. Gibbs, *Qualitative data analysis: Explorations with NVivo (Understanding social research)*: Buckingham: Open University Press, 2002.
- [165] A. Sgora and D. D. Vergados, "Handoff prioritization and decision schemes in wireless cellular networks: a survey," *Communications Surveys and Tutorials, IEEE*, vol. 11, pp. 57-77, 2009.
- [166] M. Devera, "HTB Homepage," URL: <http://luxik.cdi.cz/~devik/qos/htb>, 2003.
- [167] M. Hirvonen and J.-P. Laulajainen, "Two-phased network traffic classification method for quality of service management," in *Consumer Electronics, 2009. ISCE'09. IEEE 13th International Symposium on*, 2009, pp. 962-966.
- [168] S. Mohamed, "Automatic evaluation of real-time multimedia quality: a neural network approach," *Rennes: University of Rennes I*, 2003.
- [169] M. Varela, "Pseudo-subjective quality assessment of multimedia streams and its applications in control," *Université de Rennes*, vol. 1, 2005.
- [170] M. A. Brown, "Linux traffic control howto," URL: <http://tldp.org/HOWTO/Traffic-Control-HOWTO/index.html> Accessed July, vol. 8, 2008.
- [171] B. Hubert, "Linux advanced routing and traffic control HOWTO," *setembro de*, 2002.
- [172] N. K. Vlachos, "Global Internet Traffic Volume Prediction: Towards the Discovery of a New Formula," *American Journal Of Advanced Computing*, vol. 2, 2015.

APPENDIX A - EMULATING WIRELESS/MOBILE NETWORK DELAYS

This is the simplest example, it just adds a fixed amount of delay to all packets going out of the local Ethernet.

```
# tc qdisc add dev eth0 root netem delay 100ms
```

Now a simple ping simulation to host on the local network should show an increase of 100 milliseconds. The delay is limited by the clock resolution of the kernel (HZ). On most 2.4 systems, the system clock runs at 100hz which allows delays in increments of 10ms. On 2.6, the value is a configuration parameter from 1000 to 100 hz.

Later examples just change parameters without reloading the qdisc

Real wide area networks show variability so it is possible to add random variation.

```
# tc qdisc change dev eth0 root netem delay 100ms 10ms
```

This causes the added delay to be $100\text{ms} \pm 10\text{ms}$. Network delay variation isn't purely random, so to emulate that there is a correlation value as well.

```
# tc qdisc change dev eth0 root netem delay 100ms 10ms 25%
```

This causes the added delay to be $100\text{ms} \pm 10\text{ms}$ with the next random element depending 25% on the last one. This isn't true statistical correlation, but an approximation.

Delay distribution

Typically, the delay in a network is not uniform. It is more common to use a something like a ordinary distribution to describe the variation in delay. The netem discipline can take a table to specify a non-uniform distribution.

```
# tc qdisc change dev eth0 root netem delay 100ms 20ms distribution ordinary
```

The actual tables (ordinary, pareto, paretoordinary) are generated as part of the iproute2 compilation and placed in /usr/lib/tc: so it is possible with some effort to make your own distribution based on experimental data.

Packet loss

Random packet loss is specified in the 'tc' command in percent. The smallest possible non-zero value is:

```
232 = 0.0000000232%
```

```
# tc qdisc change dev eth0 root netem loss 0.1%
```

This causes 1/10th of a percent (i.e 1 out of 1000) packets to be randomly dropped.

An optional correlation may also be added. This causes the random number generator to be less random and can be used to emulate packet burst losses.

```
# tc qdisc change dev eth0 root netem loss 0.3% 25%
```

This will cause 0.3% of packets to be lost, and each successive probability depends by a quarter on the last one.

$$\text{Probn} = .25 * \text{Probn-1} + .75 * \text{Random}$$

Caveats

When loss is used locally (not on a bridge or router), the loss is reported to the upper level protocols. This may cause TCP to resend and behave as if there was no loss. When testing protocol response to loss it is best to use a netem on a bridge or router

Packet duplication

Packet duplication is specified the same way as packet loss.

```
# tc qdisc change dev eth0 root netem duplicate 1%
```

Packet corruption

Random noise can be emulated (in 2.6.16 or later) with the corrupt option. This introduces a single bit error at a random offset in the packet.

```
# tc qdisc change dev eth0 root netem corrupt 0.1%
```

Packet re-ordering

There are two different ways to specify reordering. The first method **gap** uses a fixed sequence and reorders every *Nth* packet. A simple usage of this is:

```
# tc qdisc change dev eth0 root netem gap 5 delay 10ms
```

This causes every 5th (10th, 15th, ...) packet to go to be sent immediately and every other packet to be delayed by 10ms. This is predictable and useful for base protocol testing like reassembly.

The second form **reorder** of re-ordering is more like real life. It causes a certain percentage of the packets to get mis-ordered.

```
# tc qdisc change dev eth0 root netem delay 10ms reorder 25% 50%
```

In this example, 25% of packets (with a correlation of 50%) will get sent immediately, others will be delayed by 10ms.

Newer versions of netem will also re-order packets if the random delay values are out of order. The following will cause some reordering:

```
# tc qdisc change dev eth0 root netem delay 100ms 75ms
```

If the first packet gets a random delay of 100ms (100ms base - 0ms jitter) and the second packet is sent 1ms later and gets a delay of 50ms (100ms base - 50ms jitter); the second packet will be sent first. This is because the queue discipline *tfifo* inside netem, keeps packets in order by time to send.

Caveats

- Mixing forms of reordering may lead to unexpected results
- Any method of reordering to work, some delay is necessary.
- If the delay is less than the inter-packet arrival time then no reordering will be seen.

Rate control

There is no rate control built-in to the netem discipline, instead use one of the other disciplines that does do rate control. In this example, we use Token Bucket Filter (TBF) to limit output.

```
# tc qdisc add dev eth0 root handle 1:0 netem delay 100ms

# tc qdisc add dev eth0 parent 1:1 handle 10: tbf rate 256kbit buffer 1600 limit 3000

# tc -s qdisc ls dev eth0

qdisc netem 1: limit 1000 delay 100.0ms

Sent 0 bytes 0 pkts (dropped 0, overlimits 0 )

qdisc tbf 10: rate 256Kbit burst 1599b lat 26.6ms

Sent 0 bytes 0 pkts (dropped 0, overlimits 0 )
```

Check on the options for buffer and limit as you might find you need bigger defaults than these (they are in bytes)

APPENDIX B – ABSTRACTS OF PUBLICATIONS

PUBLICATIONS

Topic: Performance Evaluation of OSPF and EIGRP Routing Protocols for Video Streaming over Next Generation Networks.

By: Francis L Lugayizi, Naison Gasela, and Esiefarienrhe Bukohwo Michael

Published in:

Advances in Computer Science: International Journal

Volume 4, Issue 5, September 2015

Abstract: The challenge in performance of Next Generation Networks (NGNs) is a growing concern. Bandwidth limitation for multimedia applications in NGNs such as voice and video telephony along with the increasing number of applications on the Internet, service classification and efficient resource management have all become quite challenging tasks. In order to improve the service quality in video streaming especially, there has to be efficient resource and traffic management, and using routing protocols is one way in which this can be done. We designed three network models that are configured with OSPF, EIGRP and one with both OSPF and EIGRP routing protocols, and then used the QoS parameters of throughput, packet loss, convergence time, mean latency and end-to-end delay as our metrics to evaluate the performance of OSPF and EIGRP. Our main source of network traffic was a typical video conferencing application. The results obtained during our experiments indicate that combining both EIGRP and OSPF is more reliable in providing Quality of Service than OSPF routing protocol when the main traffic used in the network is video, but when dealing with a standalone real time application network, EIGRP is better than OSPF.

Topic: Evaluation of Video Transmission Quality of Service over Next Generation Networks.

By: Obeten. O. Ekabua and Francis L Lugayizi

Published in:

Proceedings of the World Congress in Computer Science, Computer Engineering, and Applied Computing (WORLDCOMP'12) July 16-19, 2012, Las Vegas Nevada, USA.

Abstract: The challenge in service quality for the performance of Next Generation Networks (NGNs) is a growing concern. The bandwidth limitation for multimedia applications in NGNs such as voice and video telephony along with the increasing number of applications on the Internet, service classification and efficient resource management have all become quite challenging tasks. In order to handle the different types of applications in the network to improve the service quality in video and voice transmission, there has to be efficient resource and traffic management, and using routing protocols is one way in which this can be done. We designed three network models that are configured with OSPF, EIGRP and one with both OSPF and EIGRP routing protocols, and then used the QoS parameters of their throughput, packet loss, convergence time, packet delay variation and end-to-end delay as our performance evaluation metrics. Our main source of network traffic was a typical video conferencing application. The results obtained during the simulation indicates that combining both EIGRP and OSPF is more reliable in providing Quality of Service than OSPF routing protocol when the main traffic used in the network is video, but when dealing with a standalone real time application network, EIGRP is better than OSPF.

Topic: A Comparative Evaluation of QoS Routing in VANETs.

By: Francis L Lugayizi, Naison Gasela and Esiefarienrhe Bukohwo Michael

Publication Under review in:

IETE Journal of Research ISSN - 0377-2063 (Print), 0974-780X (Online)

Abstract: *The field of Vehicular Ad hoc Networks (VANETs) has gained great interest among network researchers in the past few years. Research in to VANETs has been aided to a great extent by the vast amount of research done in MANETs, and because VANETs are simply a particular type of MANET; most of the core principles found in MANETs are applied in VANETs, this has made the transition of research a great deal easier. One of the major areas of great interest in VANET research is routing, and particularly, quality of service routing. The challenge is finding ways to counter the high speed nature of VANETs and the constantly changing topologies, and then determining which routing protocols are best suited for a particular transmission type; and which ones provide more consistent and stable routing performances. Great efforts have and are still being made to improve the routing capabilities of the protocols in existence, continued upgrades and re-evaluations to QoS mechanisms in routing protocols are an integral practice in the field of VANETs. Our research work presents a comparative QoS evaluation of two routing protocols, AODV and DSR. We carried out our research by means of simulations, performed using the NS2 simulation tool, we generated realistic traffic environment models using the VanetmobiSim tool. We compared the two protocol's QoS routing performances based on the throughput and end-to-end delay metrics and a set of defined parameters. In various scenarios, either protocol out performs the other; however, an aggregated evaluation does give DSR the upper hand. We concluded that overall, DSR was the preferable protocol as it showed more stable and consistent results than AODV.*

Topic: QoS Performance analysis of Bit Rate Video Streaming in Next Generation Networks Using TCP and UDP.

By: Francis L Lugayizi, Esiefarienrhe Bukohwo Michael and Naison Gasela

Publication Under review in:

Advances in Computer Science: International Journal

Abstract: *The tremendous growth of users streaming videos on the internet has increased leading to the need of improved video quality and reception. Quality of Service (QoS) evaluation of real time services such as video streaming over next generation networks. Using routing and communication protocols is one way this improvement can be implemented. This paper presents the behavior of bit rate video streaming and monitoring the performance of video streaming QoS in relationship with the main communication protocols namely; Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). Our tool of choice in developing a test bed was the OPNET modeler. A network topology was created and duplicated 3 times in order to configure each with the different protocols and the other with both protocols. The results generated from a video conferencing application's end-to-end delay, packet delay variation and traffic dropped indicated that when a given network is to combine both protocols in video streaming, there is a recognized degree of performance in the QoS of video streaming applications.*

Topic: Performance Analysis of Quality of Service in Video Streaming Applications over Vehicular Adhoc Networks

By: Francis L Lugayizi, Thulani Phakathi, Esiefarienrhe Bukohwo Michael and Naison Gasela

Publication Under review in:

IETE Journal of Research ISSN - 0377-2063 (Print), 0974-780X (Online)

Abstract: As technology in Vehicle manufacture advances, there is a problem that seems to have never been eliminated in the past years. The problem identified is that of road accidents and traffic jams. Road accidents and traffic jams have escalated to the point that researchers, service providers and other stakeholders have joined hands in effort to find a viable solution. They sought out ways in which they started using a new technology called Vehicular Adhoc Networks (VANETs). VANETs are part of Mobile Adhoc Networks (MANETs) with the slight exception that VANETs uses vehicular nodes and moving at higher speeds. This research is based on the theory that even though it is difficult to ensure Quality of Service in VANETs because of the variating topology infrastructure but it is possible to achieve the best of it by the implementation of a routing protocol in vehicular nodes. OPNET 14.5 Modeler was the simulation tool used to design and implement scenarios to generate results. The generated results were analysed and used to compare the influence of having few or more nodes on video transmission based on throughput, delay and routing traffic received. We can conclude that proactive routing protocol, OLSR, has better service is rendered when there are few nodes in the network and AODV shows better service when there are too many nodes in a network and thus guaranteeing the user a better video quality when streaming online in High Quality Video Conferencing application.
