DOCTORAT EN CO-ACCREDITATION
TELECOM SUDPARIS ET L'UNIVERSITE EVRY VAL D'ESSONNE

Spécialité : Informatique et Réseaux

Ecole doctorale : Sciences et Ingénierie

Présentée par
Jeevan POKHREL
Pour obtenir le grade de
DOCTEUR DE TELECOM SUDPARIS

Analyse intelligente de la qualité d'expérience (QoE) dans les réseaux de diffusion de contenu Web et Multimédia

Soutenue le Décembre 2014
Devant le jury composé de :

Directeur de thèse
Ana Rosa Cavalli, Professeur, Télécom SudParis, France

Rapporteurs
Abdelhamid Mellouk, Professeur, Université Paris-Est Créteil, France
Patrick Sénac, Professeur, Supérieur de l'Aéronautique et de l'Espace (ISAE-SUPAERO), France

Examinateurs
Bachar Wehbi , Docteur, Montimage, France
Stéphane Maag, Professeur, Télécom SudParis, France
Edgardo Montes de Oca, CEO, Montimage, France
Javier Baliosian, Associate Professor, Universidad de la Republica, Uruguay

Thèse n° 2014TELE0027

Acknowledgments

I would like to thank all those people who made my thesis possible. I am indebted to them for their support and cooperation; however, given the limits of this document, I am able to mention only few of them.

First of all, I would like to express my deepest sense of **Gratitude** to my supervisor Professor Ana Cavalli at Institut Mines Télécom, who offered her continuous advice and encouragement throughout the course of this thesis. I would also like to thank Edgardo Montes de Oca for having a belief on me and giving me an opportunity to purse my PhD at Montimage. I am also indebted to my co-supervisor Doctor Bachar Wehbi at Montimage for his advice and guidance in preparing this thesis. He has not only guided but also encouraged me to excel in my research work. His trust in my abilities made me work even harder and better.

During my PhD, I collaborated with various experts and it gives me an immense pleasure to thank all those wonderful people who made this thesis possible. I would like to thank my co-authors Indira Paudel (Institut Mines Telecom), Dr. Anderson Morais (Institut Mines Telecom), Dr. Badii Jouaber (Institut Mines Telecom), Dr. Felipe Lalanne (NIC Chile Research Labs, Inria, Chile), Dr. Wissam Mallouli (Montimage), Eric Allilaire (Vierling Communication SAS Plaisir, France), Muhammad Shahid (Blekinge Institute of Technology, Sweden), Jacob Søgaard (Technical University of Denmark (DTU), Denmark), Professor Kjell Brunnström (Acreo AB, Sweden), Kun Wang (Royal Institute of Technology and Acreo Swedish ICT, Sweden), Samira Tavakoli (Universidad Politécnica de Madrid, Spain), Dr. Natalia Kushik (Tomsk State University, Russia) and Professor Nina Yevtushenko (Tomsk State University, Russia) for their valuable inputs on our collaborative research and paper writing work.

I would also like to thank secretary of LOR Department Brigitte Laurent; she was always very kind and generous in helping me to solve tedious administrative works. Undoubtedly, all the members of the LOR department and Montimage deserve a special thanks and appreciation for an exceptionally friendly working environment.

With profound love and respect, I thank my father, Uddhav Pokhrel and mother, Sabitra Pokhrel, my sisters, Sita Dahal and Sunita Dhakal and brother Suman Pokhrel who always boosted me with encouragement to achieve the new milestones in my life.

Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of this thesis work.

Jeevan Pokhrel

Abstract

Today user experience is becoming a reliable indicator for service providers and telecommunication operators to convey overall end to end system functioning (client, terminal, network, services infrastructure, media encoding, etc.). Moreover, to compete for a prominent market share, different network operators and service providers should retain and increase the customers' subscription. To fulfil these requirements they require an efficient Quality of Experience (QoE) monitoring and estimation. QoE is a subjective metric which deals with user perception and can vary due to the user expectation and context. Moreover, subjective QoE evaluation is expensive and time consuming since it requires human participation. Therefore, there is a need for a tool that can objectively measure the QoE with reasonable accuracy in real-time.

In the context of service and network providers, to fulfill user requirements, they need to provide sufficient QoS to relevant services/applications. However, QoS parameters do not necessarily reflect the user's satisfaction and feelings towards a particular service/application. Therefore it is necessary to compute the mapping or correlation between QoS and QoE. This mapping helps them understand the behavior of the overall network on user experience (QoE) and efficiently manage the network resources.

In this thesis work, we utilize this correlation between QoS and QoE to objectively estimate the QoE. For this, we use subjective methodology to create a dataset which represents the correlation between objective QoS parameters and subjective QoE. Following this, different intelligent machine learning algorithms are trained with this subjective dataset to objectively predict the QoE. The machine learning algorithms we used in our work are fuzzy expert system, fuzzy rough hybrid expert system and random neural network.

As a first contribution, we analyzed the impact of network conditions on Video on Demand (VoD) services. We also proposed an objective QoE estimation tool that uses fuzzy expert system to estimate QoE from network layer QoS parameters. As a second contribution, we analyzed the impact of MAC layer QoS parameters on VoD services over IEEE 802.11n wireless networks. We also proposed an objective QoE estimation tool that uses random neural network to estimate QoE from the MAC layer perspective. As our third contribution, we analyzed the effect of different adaption scenarios on QoE of adaptive bit rate streaming. We also developed a web based subjective test platform that can be easily integrated in a crowdsourcing platform for performing subjective tests. As our fourth contribution, we analyzed the impact of different web QoS parameters on web service QoE. We also proposed a novel machine learning algorithm (i.e., fuzzy rough hybrid expert system) for objectively estimating web service QoE.

Keywords: Quality of Experience, Quality of Service, Video, Over the top, wireless network, video on demand, web services

List of publications

Books Chapter

- **Jeevan Pokhrel**, Natalia Kushik, Bachar Wehbi, Nina Yevtushenko and Ana R. Cavalli, "Multimedia quality-of-experiences", Book chapter in Handbook of Research on Networked Multimedia Communication Systems, IGI-global, 2014 (Under review)

Journals

- Natalia Kushik, **Jeevan Pokhrel**, Nina Yevtushenko, Ana R. Cavalli, and Wissam Mallouli, "QoE prediction for Multimedia services: Comparing fuzzy and logic network approaches", International Journal of Organizational and Collective Intelligence (IJOCI), 2014 (Under publication)

Conferences

- **Jeevan Pokhrel**, Bachar Wehbi, Anderson Morais, Ana R. Cavalli, and Eric Allilaire, "Estimation of QoE of Video Traffic using a Fuzzy Expert System", 10th Annual IEEE CCNC- Multimedia Networking & Services & Applications, Las Vegas, USA, January 11-14 2013. (Published)

- Indira Paudel, **Jeevan Pokhrel**, Bachar Wehbi, Ana R. Cavalli and Badii Jouaber, "Estimation of video QoE from MAC parameters in wireless network: A Random Neural Network approach", ISCIT, 2014, IEEE, Korea (Accepted)

- **Jeevan Pokhrel**, Felipe Lalanne, Ana R. Cavalli and Wissam Mallouli, "QoE estimation for web service selection using a Fuzzy-Rough hybrid expert system", The 28th IEEE International Conference on Advanced Information Networking and Applications (AINA-2014) Victoria, Canada, May 13-16, 2014. (Published)

- **Jeevan Pokhrel**, Indira Paudel, Bachar Wehbi, Ana R. Cavalli and Badii Jouaber, "Performance Evaluation of Video transmission over 802.11n wireless network: A MAC layer perspective", to appear in The 5th International Conference on Smart Communication in Network Technologies (SaCoNeT - IWQM 2014),IEEE, June 18-20 2014, Vilanova I la Geltrù, Spain (Published)

- Natalia Kushik, Nina Yevtushenko, Ana R. Cavalli, Wissam Mallouli and **Jeevan Pokhrel**, "Evaluating web service QoE by learning logic networks", WEBIST 2014 - 10th International Conference on Web Information Systems and Technologies, 3 - 5 Apr 2014, Barcelona, Spain.(Published)

- Muhammad shahid, Jacob Søgaard, **Jeevan Pokhrel**, Kjell Brunnström, Kun Wang, Samira Tavakol, "Crowdsourcing based subjective quality assessment of adaptive video streaming", to appear in IEEE QoMEX 2014, Singapore (Under publication)

Table of Contents

List of figures

List of tables

List of abbreviations

QoE	Quality of Experience
QoS	Quality of Service
MAC	Medium Access Control
VoD	Video on Demand
ITU	International Telecommunication Union
IPTV	Internet Protocol Television
MOS	Mean Opinion Score
KPI	Key Performance Indicator
VQM	Video Quality Monitoring
RNN	Random Neural Network
RST	Rough Set Theory
OSI	Open systems Interconnection Model
XML	Extensible Markup Language
REST	Representational state transfer
SOAP	Simple Object Access protocol
WSDL	Web Services Description Language
UDDI	Universal Description, Discovery and Integration
ACR	Absolute Category rating
PC	Personal Computer
SLA	Service Level Aggreement
HTTP	Hypertext Transfer Protocol (HTTP)
RTP	Real Time Protocol
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
RTCP	Real Time Control Protocol
HLS	HTTP Live Streaming
MSS	Microscoft's Silverlight Smooth Streaming
HDS	HTTP Dynamic Streaming
DASH	Dynamic Adaptive Streaming over HTTP

PLR	Packet Loss Ratio
HD	High Defination
SD	Standard Defination
QCIF	Quater Common Intermediate Format
CBR	Constant Bit Rate
VBR	Variable Bit Rate
DSCQS	Double Stimulus Continuous Quality Scale
DSIS	Double Stimulus Impairment Scale
SSCQS	Single Stimulus Continuous Quality Scale
SDSCE	Simultaneous Double Stimulus Continuous Evaluation
SAMVIQ	Subjective Assessment Methodology for Video Quality
FR	Full Reference
NR	No Reference
RF	Reduce Reference
TS	Transport Stream
PSNR	Peak Signal-to-Noise Ratio
SSIM	Structural Similarity
AQoS	Application Quality of Service
NQoS	Network Quality of Service
SVM	Support Vector Machine
DT	Decision Tree
ANN	Artificial Neural Network
BP	Back Propagation
ANFIS	Adaptive Neural Fuzzy Inference System
DMOS	Degraded Mean Opinion Score
TFIFO	Time-ordered First In First Out
PFIFO	Packet-ordered First In First Out
MBL	Mean Burst Loss
NB	Naives Bayes
4-NN	4-Nearest Neighbour
RF	Random Forest

RBF	Radial Basis Function
VQEG	Video Quality Expert Group
FRTV	Full Reference Television
RRNR-TV	Reduce Reference No reference Television
AVHD	AudioVisual Television
HDR	High Dynamic Range
MOAVI	Monitoring of Audio Visual
QART	Quality Recognition Tasks
RICE	Real Time Interactive Communication Evaluation
IIF	IPTV Interoperability Forum
IGMP	Internet Group Management Protocol
PEVQ	Perpectual Evaluation of Video Quality
MDI	Media Delivery Index
DF	Delay Factor
MLR	Media Loss Rate
GoP	Group of Picture
PDF	Probability Density Function
DPI	Deep Packet Inspection
MMT	Montimage Monitoring Tool
DSLAM	Digital Subscriber Line Access Multiplexer
CO	Central Office
CPE	Customer Permises Equipement
VOIP	Voice over IP
MIMO	Multiple Input Multiple Output
OFDM	Orthogonal Frequency Divison Multiplexing
BA	Block Acknowledgement
SDM	Spatial Division Multiplexing
MPDU	Message Protocol Data Unit
MSDU	MAC service Data Unit
BER	Bit Error Rate
TXOP	Transmit Opportunity

NS	Network Simulator
SIFS	Short Interframe Space
BT	Bradley-Terry
WWW	World Wide Web

Chapter 1

Introduction

1.1 Introduction

In this Internet revolution era, technology became an essential part of our daily life. Most of us possess technological gadgets that help accessing different Internet services like voice, video, emails. Providing Internet access to these different services involves three different players: network operators, service providers, and content providers. This multi-party ecosystem survives as these different players fulfill customers' demands and Quality of Experience (QoE) requirements. In return, the customers pay for the service they get.

According to the ITU-T Focus Group on IPTV [1], Quality of Experience (QoE) refers to "*the overall acceptability of an application or service, as perceived subjectively by the end- user*". Similarly, Qualinet [2] defines QoE as "*the degree of delight or annoyance of the user of an application or service. It results from the fulfilment 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*". Therefore, QoE is a subjective measure and can vary according to the user expectation and context. Moreover, it is an overall end to end system effect (client, terminal, network, services infrastructure, media encoding, etc.) and depends on a number of factors that cannot simply be measured. It requires tests with actual users in a controlled environment to properly estimate the QoE; which is costly and time consuming.

In the Internet, multimedia services and applications have become an important source of income for the network operators and service providers. To compete for a prominent market share, different network operators and service providers should retain and increase the customers' subscription. For this, they should fulfil user's multimedia QoE requirements. To fulfil these requirements they require an efficient QoE monitoring and estimation tool. However, QoE is a subjective metric and can vary due to the user expectation and context. Moreover, subjective QoE evaluation is expensive and time consuming since it requires human participation. Therefore, there is a need of a tool that can objectively measure the QoE with reasonable accuracy.

Traditional approaches for measuring the user satisfaction rely on objective QoS parameters collected from the network. In this case, the QoS parameters are monitored and controlled in order to provide a satisfactory level of service quality. Different QoS parameters like bandwidth, delay, packet loss, etc. are essential metrics for determining the service quality from a technical point of view. However, QoS parameters do not necessarily reflect the users' satisfaction and feelings towards a particular service.

In this thesis work, to overcome these shortcomings, we proposed different hybrid approaches that combine both subjective and objective metrics to objectively predict QoE. These approaches utilize subjective methodology to create a dataset which represents the correlation between objective QoS and subjective QoE. Following this, different intelligent machine learning algorithms are trained with this subjective dataset to objectively predict the QoE. We applied our approaches to estimate the QoE of multimedia and web services. Moreover, we also created a novel web based video subjective test platform that can be easily integrated into a crowdsourcing platform.

1.2 Motivation

The motivation for this thesis work revolves around searching answers for two basic questions.

Why multimedia services?

The popularity of multimedia services over the Internet is immense. It is supported by the exponential growth in consumption and adoption of Internet multimedia services. Figure 1, Figure 2 and Figure 3, show the forecast of video consumption and adoption from 2013 to 2018 [3] [4]. This growth has created a threat on network resources scarcity as well as opportunities for network operators and service providers for revenue generation. As a consequence, in the recent years, multimedia traffic has diverted the attention of the research as well as industrial communities.

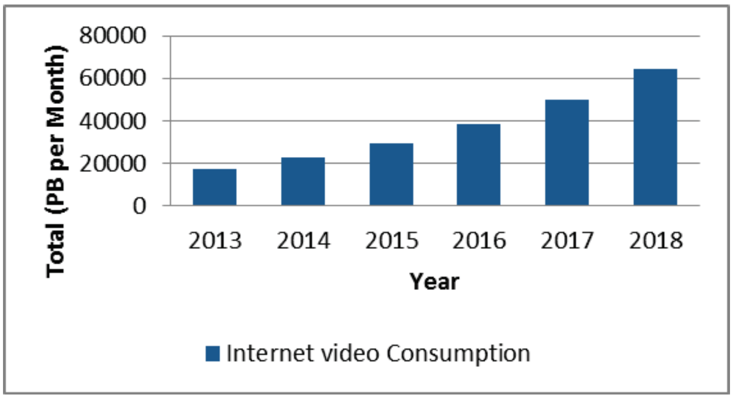

Figure 1: Consumer Internet video consumption [3]

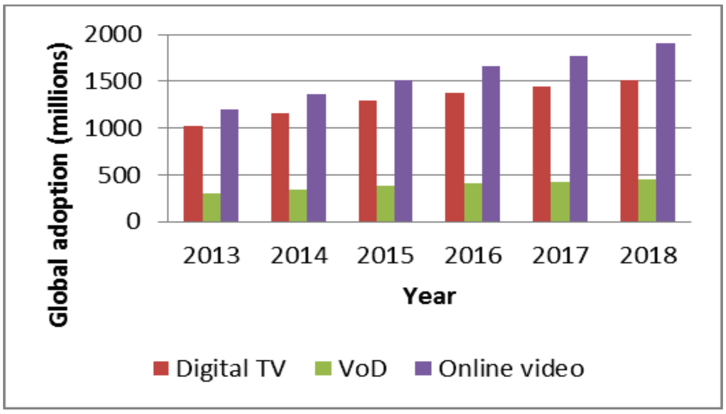

Figure 2: Residence video service adoption [4]

Online video: Video downloaded from or streamed over the Internet.

Digital TV: Services such as digital cable TV, Internet Protocol Television (IPTV), digital satellite TV(DTH), and digital terrestrial TV (DTT)

VoD: On-demand video programming that is streamed or downloaded through a TV set-top box, using a next-generation TV service.

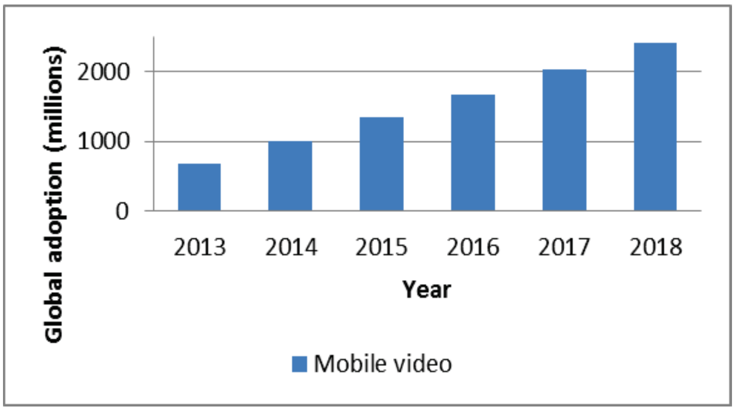

Figure 3: Consumer mobile video service adoption [4]

Mobile video: On-demand video content downloaded or streamed to the mobile handset.

Why multimedia QoE monitoring and estimation?

Multimedia QoE monitoring and estimation is a multi-disciplinary approach which involves user psychology, engineering science, economics, etc. The QoE depends on different elements (i.e., content, network, application, business model, etc.) that directly or indirectly affect the user's perception towards the multimedia service. Moreover, the diversity in these elements makes the QoE estimation rather complex and unpredictable. This motivates us to address some prominent challenges related to QoE monitoring and estimation as described below.

QoE is a subjective metric: QoE is a subjective metric as it reflects the perception of the users to a particular service. User attitude and expectation towards multimedia services play a vital role in determining the QoE. Moreover, the QoE can depend on different user profiles like age, sex, interest, skills, frame of mind, experience, etc. Also, different environmental conditions can impact how users perceive the multimedia service content. Multimedia QoE can vary according to when, where, and with whom the service is used.

Subjective vs. objective evaluation: There exist two methodologies for evaluating multimedia QoE i.e., subjective and objective. The subjective evaluation requires human participation and controlled environment to perform the test. However, this methodology cannot be applied in real-time and requires more time and cost (due to human participation). On the other hand, the objective evaluation of multimedia QoE is based on objectively measured network/media parameters. However, it may include complex mathematics and algorithms and may require objective metrics that can be hard to extract and also may not highly correlate with subjective QoE (no human participation).

Network QoS impact on multimedia QoE: The ability to identify the perceived degree of multimedia impairment due to network perturbation is a key point in the multimedia QoE estimation. Moreover, the effect of network perturbations on multimedia traffic can range from distortion-less to intolerable distortion. Also the selection of network related key performance indicators (KPIs) or QoS parameters that can impact multimedia QoE is difficult.

Need for objective tool: There is a pressing need for objective QoE estimation tools capable of processing high speed network links in real time to extract QoS parameters and correlate them with user perceived QoE. However, the relationship between QoS and QoE is fuzzy and non-linear. To address this issue there is a large number of intelligent algorithms proposed in the literature however, there is still room for innovative mechanisms to efficiently correlate QoE from QoS in real time.

1.3 Contribution of the thesis

This thesis work starts with investigating different QoE estimation techniques from the state of the art. At first, we focused on identifying the key performance indicators (KPIs) that can impact the users perceived QoE. These KPIs are basically QoS parameters relevant to different OSI layers i.e., network layer, Medium access control (MAC) layer and application layer. Then, we utilize the QoS – QoE relationship of different services to objectively estimate the QoE in real-time. For this, a subjective test is performed to correlate selected QoS parameters to users perceived QoE for a particular service. This correlation is then used by different intelligent machine learning algorithms to objectively predict QoE in real-time. In this work, we proposed novel subjective test platforms as well as different intelligent algorithms. The most relevant contributions of this thesis are listed below:

- *Objective QoE evaluation framework- based on QoS and QoE relationship*

An objective QoE evaluation framework that correlates the relationship between QoS parameters and subjective QoE to objectively estimate the QoE in real-time is presented here. For this purpose, we designed a subjective test platform for conducting subjective tests with end user participants in order to build a learning set for correlating objective QoS metrics with the subjective QoE provided by the participants. This correlation was then used by different intelligent algorithms to objectively estimate the QoE. This framework includes two novel subjective test platforms i.e., laboratory test platform and crowdsourcing platform. Moreover to correlate QoS with QoE, we used different intelligent algorithms out of which we present two novel intelligent algorithms that are fuzzy rough hybrid system [5] and logic networks [6].

- *Objective QoE estimation of VOD services based on network QoS metrics*

A methodology and a system based on fuzzy expert system to estimate the impact of network conditions (QoS) on the QoE of video traffic (VoD/IPTV) [7] is presented here. At first, we conducted subjective tests to correlate network QoS metrics with participants' perceived QoE of video traffic. Second, we proposed a No Reference method based on fuzzy expert system to estimate the network impact on the video QoE. The membership functions of the proposed fuzzy system are derived from normalized probability distributions correlating the QoS metrics with QoE. We proposed a simple methodology to build the fuzzy inference rules. We evaluated our system in three different sets of experiments. The estimated video quality showed high correlation with the subjective QoE obtained from the participants in a controlled test. We integrated our system as part of a monitoring tool in an industrial IPTV test bed and compared its output with standard Video Quality Monitoring (VQM). Additionally, we performed benchmarking analysis where we compared the result obtained from our system with Indra's probe [162] for QoE estimation in the real IPTV test bed. The evaluation results show that the proposed video quality estimation method based on fuzzy expert system can effectively measure the network impact on the QoE.

- *Objective QoE estimation of VOD services based on MAC-level QoS metrics*

The impact of different MAC-level parameters on video QoE over IEEE 802.11n wireless networks [8] is investigated here. Moreover, we present a methodology and a system based on Random Neural Network (RNN) to analyze the impact of different MAC-level parameters on video QoE over IEEE 802.11n wireless networks [9]. At first, subjective tests are performed to correlate MAC-level parameters with user's perceived video QoE. Secondly, we propose a RNN technique to estimate the impact of these parameters on video QoE. The experimental results show that the proposed method can effectively measure the impact of MAC-level parameters on video QoE. The high correlation between subjective QoE and estimated QoE validates the designed system. The study shows that careful parameterization of MAC-level parameters can improve the QoE for video streaming applications when used over a wireless network.

- *Subjective QoE estimation of OTT video services based on crowdsourcing*

We present insights into a study that investigates perceptual preferences of various adaptive video streaming scenarios through crowdsourcing based subjective quality assessment [10] here. For this, we created a novel web based video subjective test platform that can be easily integrated into the crowdsourcing platform. With the help of this platform we performed a subjective test of OTT video services to analyze whether users prefer bit rate switching or buffering events in adverse network conditions. Our study suggests that in a network environment with fluctuations in the bandwidth, a medium or low video bitrate which can be kept constant is the best approach. Moreover, if there are only a few drops in bandwidth, one can choose a medium or high bitrate with a single or few buffering events.

- *Objective QoE estimation for web service selection*

Apart from multimedia services, we also worked on estimating QoE for web service selection. Here, we present a novel method based on a fuzzy-rough hybrid expert system for estimating QoE of web services for web service selection [5]. We also analyzed how different QoS parameters impact the QoE of web services. For this, we conducted subjective tests in a controlled environment with real users to correlate QoS parameters to subjective QoE. Based on this subjective test, we derive membership functions and inference rules for the fuzzy system. Membership functions are derived using a probabilistic approach [11] [12] and inference rules are generated using Rough Set Theory (RST) [13]. We evaluated our system in a simulated environment in MATLAB. The simulation results show that the estimated web quality from our system has a high correlation with the subjective QoE obtained from the participants in controlled tests.

1.4 Overview of the work

The remainder of the thesis is divided into different chapters. In chapter 2, we present a QoE estimation framework. The proposed framework can be used by different Internet services to objectively estimate the QoE. This chapter also proposes different novel subjective platforms as well as intelligent machine learning algorithms for correlating QoS-QoE. In Chapter 3, we present a deeper insight on multimedia services, components of multimedia QoE ecosystem, different methodologies and techniques for estimating multimedia QoE etc. In Chapter 4, we analyze the impact of different network QoS parameters on users perceived video QoE for VoD services. We proposed a methodology based on a fuzzy expert system to objectively estimate the video QoE. To validate our methodology, we integrated our system as part of a monitoring tool in an industrial IPTV test bed and compared its output with standard Video Quality Monitoring (VQM). In Chapter 5, we analyze the impact of different MAC-level QoS parameters on the video QoE for VOD services. We utilize the correlation between QoS parameters and QoE to train a random neural network which is then used for objectively estimating QoE. In Chapter 6, we performed a crowdsourcing experiment to study the perceptual preferences of various adaptive video streaming scenarios. We performed a crowdsourcing based subjective test of OTT video services to analyze whether users prefer bit rate switching or buffering events in adverse network conditions. We found that users are very sensitive to buffering events however, if the bitrate is high enough and the frequency of the buffering events is low enough, this seems to be a viable alternative to decreasing the bitrate of the video or having a constant low bitrate. In Chapter 7, we proposed a novel algorithm based on fuzzy rough hybrid expert system to estimate the web QoE from objective web QoS parameters. The estimated web QoE can be used to select the most performing service among different web services. Finally Chapter 8 concludes the thesis work by highlighting the main contributions and results, and discussing future perspective and research directions.

Chapter 2

QoE evaluation framework

2.1 Introduction

In this chapter, we describe an overall QoE evaluation framework that is used for QoE estimation of different application/services in this thesis work. The proposed QoE evaluation framework is shown in Figure 4. This framework exploits the relationship between QoS parameters and QoE to objectively estimate the QoE. In order to achieve this relationship we perform a subjective test in which participants provide QoE scores to different service quality. These different service qualities are achieved by varying different QoS parameters. The QoS parameters can be associated to different OSI (Open Systems Interconnection model) layers. After that intelligent machine learning algorithms learns the relationship between QoS parameters and QoE, this learned intelligent system can be used as an objectively tool for QoE estimation.

2.2 Components of the QoE evaluation framework

In the following section, we describe each components of the QoE evaluation framework and highlight our contributions.

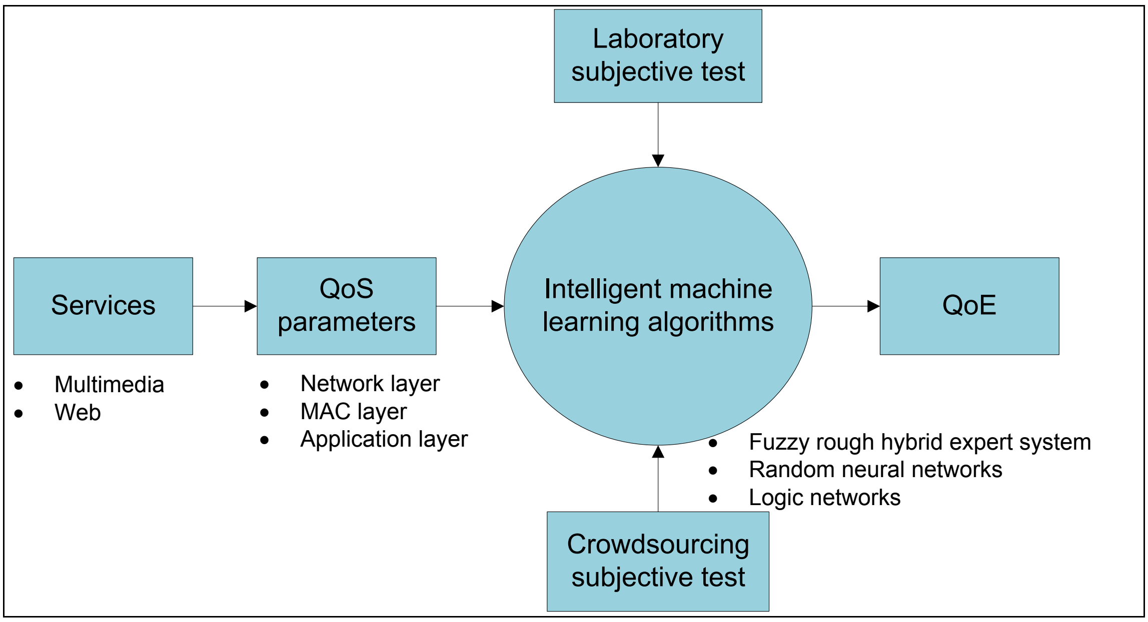

Figure 4: QoE estimation framework

2.2.1 Applications/services

Applications/services help users to perform certain tasks. We can find abundant of Internet applications that helps us to perform different activities in our daily life such as multimedia, web, email, games, etc. The popularity of these applications is growing rapidly and, moreover, we can find a number of applications having similar features and performing similar tasks. This offers a number of options to the customers and introduces a higher demand for quality. Applications and services that fulfill the customers experience requirements are the ones which are favorite to the customers. Therefore, estimating the QoE is a must for application/service providers. In this thesis work, we are focused on estimating QoE of multimedia and web services.

Multimedia applications/services: Applications that communicate any combination of audio, image, video, or data traffic are known as multimedia applications. In this modern era, multimedia applications are the medium of information, communication and entertainment for the customers. Some of the most used multimedia applications are audio/video conferencing, television, video-on-demand, telephony, etc. Moreover, multimedia traffic has become a principal source of traffic in today's Internet. However, these multimedia applications are delay sensitive, bandwidth intense, and loss tolerant [15], which makes them very sensitive to varying network conditions. Hence, it is of utmost importance for network and service providers to monitor and estimate user experience to retain the higher service quality.

Web applications/services: Web Services (WSs) are self-contained software systems that can be published, advertised, located, and invoked through the web, usually relying in standardized XML technologies (REST, SOAP, WSDL, and UDDI [15]) for description and publication, and on Internet Protocols for invocation [16]. Web applications/services are very common in the Internet. Some of the popular web services are weather forecast, currency conversion, search engine, maps, etc. With the proliferation of web services on the Internet, it has become important for service providers to select the best web services for their clients in accordance to their functional and non-functional requirements. Users gets intolerant if the content is not served in expected time and easily switch to other options if their needs are not fulfilled [17]. About 90% of the people do not want to complain for the low service quality. They just leave the service and move to another ones [18]. So service providers and operators should not wait for user feedback for improving the service quality, instead they should continuously monitor QoE and improve it as required. They should provide users with services that can offer high QoE values.

2.2.2 Quality of Service (QoS) parameters

The term QoS refers to the probability of the telecommunication network meeting a given traffic contract [19]. QoS is actually the ability of a network element (e.g., an application, host, or router) to have some level of assurance that its traffic and service requirements would be satisfied [20]. QoS is defined in terms of QoS parameters. QoS parameters are generally used to represent network layer quality however; it can correspond to different OSI layers. Moreover, the state of QoS parameters at different OSI layers can only depict the overall service quality. Therefore, every service or application should meet QoS at different OSI layers to deliver higher service quality. In this thesis, we considered QoS parameters from Medium Access Control (MAC) layer, Network layer, and Application layer. Detail of these QoS parameters will be discussed later in the different chapters.

2.2.3 Subjective test

The subjective test involves human participants rating the test sequences in predefined environmental conditions. There are different methodologies for performing the subjective test depending on the service or application to be tested. Subjective tests are mostly performed in laboratory environments however, crowdsourcing based subjective test environment is gaining popularity as well. In our thesis work, we perform subjective test to correlate QoS parameters to the QoE score provided by the users. For this purpose, we create test sequences by injecting perturbation (i.e., by varying different QoS parameters) in the original sequence. After that, these test sequences are shown to different participants in random other, who rate this sequences providing one of the MOS [21] scores. The MOS scores have five categories i.e., imperceptible (score 5), perceptible but not annoying (score 4), slightly annoying (score 3), annoying (score 2), and very annoying (score 1). Absolute Category Rating (ACR) is used for rating the quality [49]. In our work, we used the laboratory environment to perform subjective test for web and VoD services and the crowdsourcing environment for OTT services.

Laboratory subjective test environment: The laboratory experiment provides a controlled environment for performing subjective tests for evaluating the multimedia quality. Different parameters associated with the test like noise-level, distance between screen and users, screen size, etc. can be easily controlled according to the requirements. However, lab-based experiments have limitations such as 1) high cost in terms of time and labor, 2) limited participants' diversity [22]. A laboratory experiment takes weeks for preparing tests, recruiting users, and scheduling time slots for supervising the experiments. Also users need to be physically present in the laboratory to perform the test [22]. Generally lab tests are performed in university or research laboratory so the participants for the test are either students or researchers. This limits the diversity of participants: most of them are boys which are young and well-educated, etc.

Crowdsourcing subjective test environment: In crowdsourcing environment, subjective experiments can be performed from distance and there is little control over the participant's environment. Application software that is integrated in a web platform (crowdsourcing platform) is used to perform the test. This methodology mainly involves collecting subjective assessment of quality through ubiquitous streaming via the Internet. Different screen test and cheat detection mechanisms are employed in a crowdsourcing experiment [23] [24] to make the test reliable and resilient. Crowdsourcing based subjective experiments have gained attention to replace needs of lab-based tests and these experiments offer promising correlation with the former [25]. This allows an investigator to get opinions from a vast variety of subjects; in a time-flexible, test-data size scalable, and swift manner [10]. Examples of crowdsourcing application software available for multimedia subjective tests are QualityCrowd [26], PC-Video-Test-Interface [10], etc.

2.2.4 Intelligent machine learning algorithms

Intelligent machine learning algorithms automatically learn from the past observations to make accurate predictions in the future. They are mostly used in classification problems. Some of the most popular machines learning algorithms are decision tree, neural networks, fuzzy expert system, etc. In our work, the QoE estimation system uses machine learning algorithms to learn the correlation between QoS parameters and subjective QoE. The learning process involves training the algorithm with subjective data set. Once the system has learnt, it can predict the QoE based on any combination of input QoS parameters.

Advantages of using machine learning algorithms include [27]:

- They are much more accurate than the human crafted rules since they are data driven.
- They do not require expert or programmer for making decision.
- They can be performed automatically and in real time.
- They are flexible and cheap.

Disadvantages of using machine learning algorithms include:

- They required lot of labelled data.

- Impossible to get perfect accuracy.

2.2.5 Quality of Experience (QoE)

In this thesis work, the output QoE scores are represented in terms of MOS scores [21] as presents in Section 2.2.3. Therefore, the output of the proposed QoE estimation framework is a MOS score that reflects the mean opinion score of different customers. This MOS score is as close as possible to those voted by the human users.

2.3 Conclusion

In this chapter, we presented a QoE evaluation framework for estimating QoE of different application/service. This framework learns the correlation between QoS parameters and subjective QoE with machine learning algorithms to objectively estimate QoE. The QoS parameters are associated with different layers of the OSI model. The framework utilizes laboratory based subjective test or crowdsourcing based subjective test for correlating QoS parameters to subjective QoE. Furthermore, it incorporates different intelligent machine learning algorithms like fuzzy rough hybrid expert system, random neural networks, and logic networks to learn this correlation for objectively estimating QoE.

Chapter 3

State of art

3.1 Introduction

The multimedia era started in the middle of the last century when the television appeared and people got used to watch TV shows and movies at home. Moreover, with the VCR, DVD and Blue-ray disks and players, viewers could use the TV set to watch recorded material. In parallel, in the late 90s internet video service started. This was a time when video Internet technologies were merely innovation and novelty. Due to the technological limitations with dial-up connections and slow modems, it was hard to achieve good transfer rate and superior quality. In addition, the limited graphics processing power at that time prevented wide adoption of high quality videos. However, with the proliferation of broadband internet and the tremendous increase in processing power, multimedia technologies boomed [28]. The combination of high speed internet and sophisticated powerful devices introduced a completely new way for multimedia consumption through Internet. It was the boom of "Over The Top" and "Video on Demand" services.

Moreover, the advancement in different concurrent digital multimedia technologies and the proliferation of smart mobile terminals with their application ecosystem have exponentially increased the popularity of Internet multimedia services, which are a key player in current ICT business. It is expected that video traffic will reach 66% of the global mobile traffic by the year 2015 with one million minutes of video content crossing the Internet every second [29]. At the same time, video consumers are getting more demanding about the quality of multimedia content. Therefore, in order to satisfy users' demands with acceptable viewing quality, it is utmost necessary to monitor their satisfaction.

The traditional approach of measuring the user satisfaction relies on QoS parameters collected from the network. In this case, the QoS parameters are monitored and controlled in order to provide a satisfactory level of service quality. Different QoS parameters like bandwidth, delay, packet loss, etc. are essential metrics for determining the service quality from a technical point of view. However, QoS parameters do not necessarily reflect the user's satisfaction and feelings towards a particular service. In order to accurately address the human perception of the service quality, a new concept of measuring the Quality of experience (QoE) is involved.

The QoE refers to “the degree of delight or annoyance of the user of an application or service. It results from the fulfilment 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” [2]. Accordingly, the QoE is a subjective metric and can vary due to the user expectation and context. Moreover, the QoE is a reliable indicator for service providers and telecommunication operators to convey overall end to end system functioning (client, terminal, network, services infrastructure, media encoding, etc.). Furthermore, it is a multi-disciplinary approach when determining QoE, which involves user psychology, engineering science, economics, etc. The QoE depends on different elements (i.e., content, network, application, etc.) that directly or indirectly affect the user’s perception towards the multimedia service. These elements should perform to their best to provide high user experience. However, the diversity in these elements makes the QoE estimation rather complex and unpredictable.

It has been shown that multimedia customers are willing to pay for better quality of experience [30]. The success of paid VoD services (not to mention that of Netflix) is just a simple proof. However, customers get intolerant if the multimedia service is not satisfactory and they easily shift to other options if their needs are not fulfilled. Therefore, the user satisfaction is utmost important for retaining customers and has become the main differentiator for the success of network operators and service providers. Correspondingly, network operators and service providers should accurately estimate and monitor the multimedia QoE in order to track the performance and quality of a particular service.

In the following sections, we will look at the different components of the multimedia QoE ecosystem and their roles in the multimedia QoE. We will present different multimedia transmission components and highlight some key performance indicators relevant to the multimedia QoE. We will also explain different multimedia QoE estimation methods and techniques along with the QoS/QoE learning algorithms. Finally, we will discuss some of the future challenges and issues related to multimedia QoE.

3.2 Multimedia QoE ecosystem

Quality of Experience (QoE) is an important indicator for network operators and service providers to help them assessing the user acceptability towards a particular service or a particular application. As the paradigm is shifting towards user-centric evaluation of service or application performance, the real time estimation of QoE is becoming a necessity for network operators and service providers in order to attract and bind users to their service.

Generally, QoS parameters are used for evaluating the quality of multimedia transmissions which do not necessarily reflect the user satisfaction with the service. The QoS parameters reflect network and service level performance; however, they do not address the user's reaction to the service or application. On the other hand, QoE is a multi-disciplinary metric, which is subjectively evaluated and can vary according to the user expectation and context. Moreover, it is an overall end to end system effect (client, terminal, network, service infrastructure, media encoding, etc.) that depends on a number of factors and cannot be simply measured.

The multimedia QoE ecosystem incorporates all the possible components that directly or indirectly affect the user's perception towards the multimedia service. Figure 5 shows different components that play an important role in determining the user's perceived QoE. The diversity of these components makes the QoE estimation rather complex and unpredictable.

Figure 5: Multimedia Quality of Experience ecosystem

3.3 Role of each component in multimedia QoE ecosystem

Each component of a multimedia QoE ecosystem determines the effectiveness of the service. A component's behavior depends on the performance of other components and their behavior can affect the overall multimedia QoE. The multimedia service provider should monitor, evaluate and adapt these components according to different service requirements in order to offer a better experience. The success of a multimedia service depends on how well all these components are behaving in the ecosystem. In the following section, the role of different components of the multimedia QoE ecosystem is described.

3.3.1 Content

If the original multimedia quality is unsatisfactory, then the QoE at the destination is also unsatisfactory. In order to have a better multimedia QoE, the multimedia content should be of high quality. Some of the characteristics of multimedia contents and their description are listed in Table 1.

Table 1: Multimedia content characteristics and their description

Content Characteristics	Description
Bit rate	Bit rate in terms of video transmission refers to the minimum rate at which video bits are transferred from a source to a destination. The higher is the multimedia bit rate, the better is the multimedia quality.
Frame rate	Multimedia frame rate refers to a number of multimedia frames presented per second. The higher is the frame rate, the smoother the video appears and hence, the better is the video QoE.
Resolution	Video resolution refers to the number of pixels in both directions (width and height) of a video frame. A higher frame resolution yields to a better video quality.
2D/3D	Video types i.e., 2D/3D, refers to the visual dimensions of a video content. These content types have different service and network requirements.

Therefore, different content characteristics represent the diversity of multimedia content and their properties influence the multimedia QoE. However, content characteristics like higher bit rate, frame rate and resolution increase the complexity of the video encoding and require higher network bandwidth for getting better QoE. Moreover, different multimedia devices and applications should support different content characteristics for a smooth playback.

3.3.2 Network

Network represents the segments from the content servers, the core and distribution segments and the access network with a combination of copper, fiber and wireless links. In order to experience better multimedia QoE, transmission conditions at the network end should be reliable. The network condition is represented by the quality of service (QoS) parameters.

Each multimedia service has its own QoS requirements. Some of the network QoS parameters and their descriptions are presented in Table 2.

Table 2: Network QoS parameters and their description

QoS parameters	**Description**
Packet loss rate	Packet loss rate is the ratio of the total number of packets lost in transmission compared to the total number of packets sent. The higher is the packet loss rate, the lower is the multimedia QoE.
Burst loss	If a group of consecutive packets are lost then it is defined as a burst packet loss. A higher burst loss results in a lower multimedia QoE.
Jitter	Jitter is the variation in the packet inter-arrival delay. Higher jitter results in a lower multimedia QoE.
Bandwidth	Bandwidth is the amount of information that can flow in a network during a specific period of time. To some extent, the higher is the network bandwidth, the higher is the multimedia QoE.
Packet error rate	Packet error rate is the ratio of the total packets received with errors to the total number of transmitted packets. The higher is the packet error rate, the lower is the multimedia QoE.

The variation in time of the network conditions will directly impact the network QoS parameters and therefore, the multimedia QoE.

3.3.3 Users

Each individual user has his/her own perception towards a multimedia quality. User attitude and expectation towards multimedia services play a vital role in determining the quality of the service. Multimedia QoE can depend on different user profiles like age, sex, interest, skills, frame of mind, experience, etc. For example, if a person is young and sportive, he/she might be inclined to sport content and his/her perception towards the QoE of sport content can be different than for a movie content. Thus, it is necessary to categorize users and to analyze their needs.

3.3.4 Environment

Different environmental conditions can impact how users perceive the multimedia service content. Multimedia QoE can vary according to when, where, and with whom the service is used. For example, multimedia QoE will be different when the multimedia service is used while commuting, in a cafe, at home, or in a bar. These places offer different levels of external noise that can directly or indirectly impact how a user perceives the service. Moreover, time of day (morning, afternoon, evening, etc.) also has an impact on the user multimedia experience. It was observed that for the same multimedia content, a higher level of user experience was obtained in the afternoon and in the mid-day [31]. Also, the results obtained from lab tests (having controlled environment) and crowdsourcing experiment for multimedia QoE are different [10].

3.3.5 Device

Multimedia content can be consumed through different devices like televisions, personal computers, laptops, tablets, mobile phones, etc. These devices offer different level of screen size resolution as well as different sound quality to the users. Therefore, the user experience can vary when different devices are utilized. For example, the user experience of watching a video content on a personal computer (PC) at a fixed distance is very different from that when watching it on a handheld mobile device where you have a possibility to move [32]. Subjective tests for multimedia QoE conducted by [32] on PC and mobile handset revealed that, for the same content, almost all users preferred the PC over the mobile handset.

3.3.6 Application

Multimedia content can be consumed as a video-on-demand (VOD) service, over the top (OTT) service, broadcast service, etc. These services provide different multimedia applications for usability and interactivity. Multimedia applications should be self-intuitive and easy to use for getting a higher QoE. In the study performed by Bernhaupt et al. [33], it was found that the usability rating is higher when the service is easier to use. Moreover, every application has its own video buffering scheme, encoding, and decoding, which can affect the playback of multimedia content and the overall multimedia QoE.

Each component of the ecosystem plays an important role in determining how users perceive the multimedia quality. If any element of the ecosystem is imbalance, the overall multimedia ecosystem is disturbed with lower QoE. This implies that concerned parties involved in the multimedia content distribution should be very careful in monitoring, evaluating and adapting these corresponding components for getting a higher QoE.

3.4 Multimedia transmission and Key performance indicators

Nowadays, multimedia traffic is the major source of traffic over the Internet. It is used for different purposes like entertainment, education, advertisement, etc. Internet technology boosted the growth of multimedia by providing a transmission platform from where each individual can stream the content to their devices. However, Internet is a best effort service and is shared by different other data greedy applications which means that multimedia traffic are vulnerable to various network perturbations. To overcome this time varying nature of Internet's network quality, different operators provide dedicated service (in the form of SLA agreement) to the managed multimedia services like IPTV and VOD. However, most of the multimedia services are over the top (OTT) and need to share the same bandwidth with competing applications. In the following section, we study different aspects of multimedia transmission in Internet.

3.4.1 Multimedia transmission in Internet

Figure 6 illustrates different components of multimedia transmission system in the Internet. The raw analog multimedia contents are huge in size and can consume a large amount of network resources. To compensate this, analog multimedia contents are digitalized and compressed by video and audio compression algorithms in the source encoder, and stored in the multimedia server. Upon the client's request, the multimedia content is retrieved from the server and the channel encoder adapts the multimedia stream to the network QoS requirements by adding redundant information for error recognition and correction. After that, the multimedia stream is transmitted (streamed) to the client terminal over Internet. Multimedia stream can suffer from different types of perturbation like packet loss and delay due to congestion. At the client terminal, a channel decoder performs the error detection and correction and transmits the digital data to the source decoder. The digital data received from the channel decoder is transformed into continuous waveform in the source decoder, which can be viewed or listened to using different players at the application layer. To compensate the network jitter different multimedia players implement a playout buffer. The playout buffer stores some multimedia content in order to have a smooth playback.

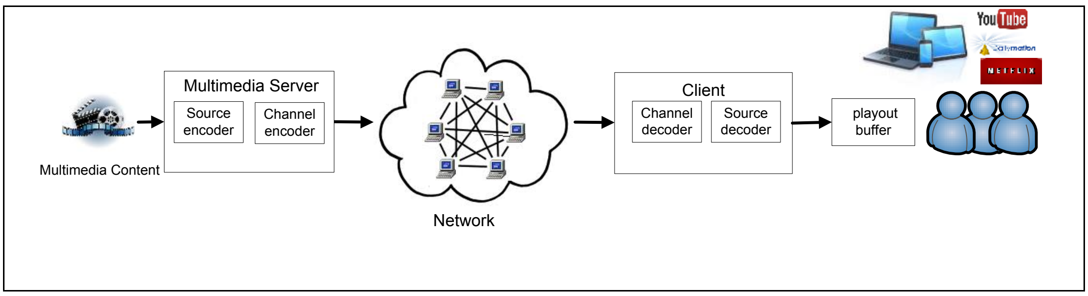

Figure 6: Multimedia transmission

Different streaming protocols for data transmission are used over the Internet. Streaming protocols control the data transfer between the multimedia server and the clients. The most popular streaming protocols used for multimedia Internet transmission are HTTP over TCP and RTP over UDP. In the following, we explain the characteristics of these protocols.

3.4.1.1 RTP/UDP based streaming

The Real Time Transport Protocol (RTP) [34] is one of the popular and most used streaming protocols for multimedia transmission. Most of the IPTV and VOD services use RTP/UDP based multimedia steaming. RTP is transmitted over UDP and is basically used for real-time transfers (TCP is unsuitable for real-time transfers due to high delay [35]). RTP works in conjunction with the Real Time Control Protocol (RTCP), which operates at the session layer. The main function of RTCP is to provide a feedback on the quality of the data distribution [36].

However, when using UDP, data packets often have difficulties getting through firewalls and network address translators [37]. UDP is an unreliable protocol and the multimedia streams can suffer packet loss which might cause distortion of the multimedia content.

3.4.1.2 HTTP/TCP based streaming

Nowadays, most of the OTT video streaming applications use HTTP/TCP as a transport mechanism for multimedia streaming. This is because HTTP/TCP can easily pass through firewalls and routers. It also does not require special proxies or caches. In addition, HTTP media streaming is easier and cheaper to deploy because web streaming can use generic HTTP solutions and does not require specialized servers at each network node [37].

However, the use of TCP for streaming has some shortcomings. Due to the TCP congestion avoidance mechanism, it produces transmission rate which is saw-tooth shaped and has high variation. Also, due to TCP reliability, retransmission of lost packets occurs, which introduces additional transmission delays. To cope up with this short term bandwidth variation of TCP, streaming services deploy playout buffer [38]. The use of TCP guarantees the reception of all packets to ensure no video impairment, however, due to severe network conditions multimedia transmission can freeze for a long time. Also to cope up with different network conditions different adaptive HTTP based video deliver schemes have been developed [39].

In adaptive HTTP based video delivery, video contents are encoded in different data rates and stored as small fragments (a few seconds each) on the server. It has an ability to switch between different data rates based upon network conditions and other variables [39]. This will eliminate the unwanted freezing events at the client terminal however, reducing the multimedia quality, i.e., the resolution. Different implementations of adaptive HTTP based video delivery are Apple's HTTP Live Streaming (HLS), Microsoft's Silverlight Smooth Streaming (MSS), Adobe's HTTP Dynamic Streaming (HDS) and Dynamic Adaptive Streaming over HTTP (MPEG-DASH) [39].

3.4.2 Impacts of different transmission components on multimedia QoE

Different components are involved when a multimedia content is being transmitted from the server to a client. Each of these components has its own effect on the multimedia QoE. Out of these different components source encoding, network and playout buffer have a significant effect on multimedia QoE. In the following section some of these components and their effect on multimedia QoE are explained.

3.4.2.1 Source coding

A source encoder converts the uncompressed raw media signals into bit streams. However, the size and the bitrate of this bit stream are huge. Therefore, different compression techniques are used in source coding to make the source data light weighted with the lowest possible bit rate in order not to exceed the available network bandwidth. The compression is performed in such a way that the same data can be easily reproduced at the decoder. However, data compression adds visible distortion and artifacts to the multimedia quality since a lot of redundant information is lost. Moreover, the highly compressed multimedia streams are easily susceptible to network impairments. Examples of video compression standards are MPEG-2, H.264, H.265, VC-2, etc.

3.4.2.2 Network

Today's Internet services rely on unreliable and best effort network. The network bandwidth available between the source and destination is unknown in advance and can change over time. Therefore, there is no guarantee for the bandwidth, the packet loss, the burst loss, the jitter for having a good multimedia quality. Moreover, multimedia streams are most sensitive to the network perturbations and the effect of network perturbations can range from distortion-less to intolerable distortion. The impact of network perturbation on multimedia content streaming can result in a frame loss, freezing, pixelization, etc.

3.4.2.3 Playout buffer

A multimedia service requires a multimedia player at the client terminal. Most of the players employ a playout buffer to compensate the jitter in the network. If the playout buffer gets empty the playback of the multimedia stream gets interrupted. The size or length of the playout buffer can determine the start-up time or freeze events of the multimedia stream. If the length of the playout buffer is high, then the start-up time will be high; that means users need to wait longer time before playback. On the other hand, if the length of the playout buffer is small, the buffer gets empty fast and there will be risk of freezing events. Therefore, it is required to design the playout buffer smartly. Till now there is no systematic guideline for buffering strategy and dimensioning of player buffer size, however; in the following section we outline video buffering strategies as discussed in [40].

- Standard buffering

In the standard buffering strategy, the player maintains one buffer threshold level (buffer size). To start playback, the player continuously downloads the content up to this threshold and when the threshold is achieved the playback starts immediately. In parallel, the player downloads the video content in order to maintain this threshold. However, if the download bandwidth drops below the video data rate, the data available in the buffer decreases continuously. This will empty the buffer and playback stalls until the buffer fills up again to the particular threshold. The main limitation of the standard buffering scheme is that the player needs to have a larger buffer threshold level to compensate longer bandwidth drops, and this leads to a longer prebuffering time which can affect the user QoE.

- Dual-threshold buffering strategy

To avoid the limitations of the standard buffering scheme, the dual-threshold buffering strategy maintains two buffer threshold levels. The lower threshold level is chosen to start the playback fast and the higher threshold level is chosen to compensate longer bandwidth drops. In this scheme, when the buffer level reaches the lower threshold it starts playback and increases its buffer threshold to the higher threshold level. Therefore, in good network conditions, the buffer grows continuously to achieve the higher buffer threshold level, which can be used in the future to compensate sudden bandwidth drops. However, if the buffer is empty due to adverse network conditions, then a lower value of buffer threshold is triggered for a fast start.

3.4.3 Indicators (KPIs) for multimedia Quality

Key performance indicators (KPIs) are parameters that indicate the overall success or quality of a particular service. In case of multimedia services, KPIs indicate the overall quality of the multimedia service. The measurement of KPI parameters at different transmission points can help to identify and to locate the problem. In general, multimedia KPIs are classified into two groups namely: network level and application level KPIs.

3.4.3.1 Network level KPIs

Network level KPIs are the QoS parameters associated with the network layer that indicate the quality of a multimedia service. Different packet level information (i.e., IP level information) is used to measure network level KPIs. Some of the important network level KPIs are described below.

Packet loss ratio (PLR)

Packet Loss Ratio is the ratio of the total number of packets lost in a transmission compared to the total number of packets sent. The higher packet loss ratio helps to conclude that the network has problems. Multimedia traffic is highly susceptible to the packet loss ratio.

PLR is defined as follows

$$\text{PLR}=N_{loss}/N_{send},$$

where N_{loss} is the total number of packets lost and N_{send} is the total number of packet sent.

Average Delay

Average delay refers to an average time needed for a packet to reach from source to destination. It can be measured in *ms* (milliseconds) or *µs* (microseconds). Larger delay results in increase of start-up time in multimedia playback.

End to End Delay = Packet receive time – Packet send time

Average delay = Σ (Packet receive time – Packet send time)/ total packets received

Burst loss

If a group of consecutive packets is lost then it is defined as burst packet loss. The larger is the burst loss, the greater is the multimedia quality degradation. A single burst loss can impact a single video frame or can propagate to a number of video frames.

Jitter

Jitter is the change in the packet inter-arrival delay. Jitter shows how much the latency changes from packet to packet. A low jitter indicates a good and uninterruptable connection. A high jitter is a sign that there is congestion in the network. If we consider the RTP protocol specified in RFC3550 [41], the jitter is calculated as:

If *Si* is the RTP timestamp from packet i, and *Ri* is the time of arrival in RTP timestamp units for packet *i*, then for two successive packets *i* and *i-1*, inter-arrival delay D may be expressed as:

$$D_i = (R_i - R_{i-1}) - (S_i - S_{i-1}) = (R_i – S_i) - (R_{i-1} – S_{i-1})$$

The inter-arrival jitter is calculated continuously according to the formula:

$$J_i = (15 * J_{i-1} + |D_i|)/16$$

Throughput

Throughput refers to the number of bits received during a time unit. It is measured in bits per second (bps). The higher is the multimedia content streaming throughput; the better is the multimedia service quality. If the throughput is less than the required multimedia bit rate, the multimedia quality degrades. For example, if throughput is less than desired multimedia bit rate, then the congestion occurs and packets are lost.

Number of duplicate packets

If the same packet is received more than once it is considered as duplicate packet. When duplicated packets appear this means that there are some configuration error in the network or some devices are defective. If the number of duplicate packets increases, the multimedia quality decreases.

Number of reordered packets

A packet is considered as reordered if the sequence number is smaller than the sequence number of the packet previously received. If the number of reordered packets increases, the multimedia quality decreases.

3.4.3.2 Application level KPIs

Application level KPIs are the performance parameters that are directly associated to the application layer and the presentation of a multimedia content. Some of the important application level KPIs are described below.

Resolution

Resolution is taken into account when video content is considered. Video resolution refers to the number of pixels in both directions (width and height) of a video frame. It is represented in width × height format. Different video formats such as HD (1280× 720), SD (720× 480), QCIF (176× 144), etc. have their own video resolutions. The higher is the video resolution; the better is the video quality. It can be noticed that, the low resolution videos are preferred for handheld devices and the higher resolution videos are preferred for bigger display screens.

Frame rate/sample rate

Multimedia frame rate refers to the number of video frames presented per second. The higher is the frame rate, the smoother the video appears i.e., the better is the video QoE. Video frame rate is measured in frame per second (*fps*). In case of audio, a sample rate is used, which is measured as the number of audio samples carried per second (Hz).

Encoding rate

Multimedia encoding rate refers to the data rate at which the multimedia files are encoded. Multimedia files are generally encoded in Constant Bit rate (CBR) or in Variable Bit rate (VBR). The higher is the bit rate, the higher is the image quality, however, a higher bit rate adds more complexity to the system and requires a larger network bandwidth. This parameter is measured in bits per second.

Freezing time

The duration of time when the multimedia playback stops is called freezing time. This phenomenon occurs when the application's playout buffer is empty. For smooth playback, application's playout buffer should be always full so that video frames/ audio samples can be always available for the playback. The higher is the freezing time, the lower is the multimedia QoE. This parameter is measured in seconds.

Freezing frequency

The number of times the multimedia playback stops or freezes per unit time is referred as freezing frequency. If the freezing frequency increases during the multimedia playback, then the multimedia QoE decreases.

Blurriness and blockiness

Blurriness and Blockiness are applicable to video content. If there are insufficient bits available to represent some details of the image, blurriness and blockiness occur. This phenomenon is basically caused due to the compression techniques.

3.5 Estimation of Multimedia Quality of Experience

Multimedia services and applications have grown exponentially in the recent years. Different multimedia services like IPTV, VOD, OTT video, video conferencing, etc. are very common in a massive Internet market. To gain a prominent market share, different vendors are competing with each other; such competition includes cable television, Internet service providers, traditional and emerging telephony carriers, etc. On the other hand, multimedia customers are expecting to access a high quality service on any device they are using (PC, smartphone, tablet...). A user gets intolerant if the multimedia service is not satisfactory and he/she easily shifts to other options if his/her needs are not fulfilled. This requires the service provider to deliver the best multimedia quality to the customers under heterogeneous network conditions as well as for different terminals' capabilities. In these challenging scenarios, it is critical to guarantee an appropriate QoE for the end users. This requires efficient QoE estimation and monitoring techniques to track the performance of the multimedia service in terms of user’s perception.

The QoE estimation methods can be implicitly categorised into subjective and objective methods. Subjective methods consist of many participants viewing sample multimedia and rating its quality according to a predefined quality scale depending on their personal perception. On the other hand, objective methods are used to measure the QoE based on objectively measured network/media parameters. Different limitations of subjective and objective methods are listed below.

Limitations of subjective methods:

- Testing environment requires strict attention.
- Real-time implementation is difficult.
- Process is hard for automation.
- Subjective estimation is costly and time consuming.

Limitations of objective methods:

- These methods are hard to correlate with human perception.
- Objective methods may require high calculation power /time.

- When using objective methods, it is hard to integrate all quality affecting parameters of the model.

As listed above, both methodologies have their own shortcoming; to overcome them, various hybrid approaches that combine both subjective and objective methodologies are proposed. For example, in different approaches such as those presented in [42] [6] [5] [7], the authors use a dataset which represents the relationship between objective QoS parameter values and the subjective QoE. Following this, different intelligent systems are trained with this subjective dataset to objectively evaluate/predict the QoE. Several researchers have shown that such hybrid approaches can objectively reflect the subjective mean opinion score of users with reasonable accuracy.

In the following section, we describe some QoE estimation techniques that can be used when estimating the QoE for a video content.

3.5.1 Video QoE estimation

Video QoE estimation refers to the quality estimation of a video signal originated from different sources (movies, sports, cartoons, etc.). Similar to the speech QoE estimation, the video QoE estimation can be classified into two categories which are subjective and objective estimation. Different methodologies for subjective QoE estimation of video signals are described below.

3.5.1.1 Subjective QoE estimation

Subjective video quality estimation methods require an appropriate test environment. General environment conditions for a subjective laboratory test and a test at home are defined in [46]. Moreover, each user should be screened for (corrected-to-) normal visual acuity on the Snellen or Landolt chart, and for normal color vision using specially selected charts (Ishihara, for instance) [46]. The participation of at least 15 participants is considered as statistically reasonable for this kind of subjective tests [46].

Double Stimulus Continuous Quality Scale (DSCQS)

In the DSCQS method, participants are required to give scores to multiple video sequence pairs. These pairs consist of original and test sequences of duration of around 10 seconds each. The pairs are shown in alternating fashion. These sequences are shown twice and in a random order. Participants are unaware which sequence is an original and which one is a test sequence. They rate the quality on a scale of bad to excellent. This scale has an equivalent scale from 0 to 100. The difference between these two scales is used to remove the uncertainties caused by the material content and/or viewer's experience [46].

Double Stimulus Impairment Scale (DSIS)

The DSIS method [46] differs from DSCQS by showing multiple video sequence pairs only once and the original sequence is always shown before the test sequence. If longer test sequences (for example, over 10 seconds) are shown when using the DSIS or DSCQS method, the time between the original sequence and the test sequence can be increased. Furthermore, it can be hard to rate the sequences accurately and a psychological "recency" effect can be noticed [47].

Single Stimulus Continuous Quality Scale (SSCQE)

In the SSCQE method, a video sequence of around 5 minutes is shown to the participants. Each participant needs to evaluate the video quality instantaneously by continuously adjusting the slider in each 1-2 seconds. The DSCQS scale (from Excellent to bad) is used in this case [46]. The reference video sequence is not provided to participants. One may notice that in this case it is difficult to compare scores for different test sequences as well as to provide the overall quality rating for a particular test sequence. Moreover, the scores provided by end-users can be also affected by a "recency or memory" effect [47].

Simultaneous Double Stimulus Continuous Evaluation (SDSCE)

In the SDSCE method, two video sequences, the original and the test, are simultaneously shown to participants. Participants will check the difference between these two sequences and rate the quality by moving the slider of a handset-voting device using a 0-100 scale. If the difference between the two sequences is null then the slider should be at the top i.e. 100, while when the quality difference is maximum the slider should be at the bottom i.e. 0. Participants are unaware if which sequence is the original and which one is the test sequence [46].

Subjective Assessment Methodology for Video Quality (SAMVIQ)

Most of the methodologies discussed above were developed to perform the subjective test in the TV or similar environment. However, the SAMVIQ method can be used for PC and for mobile environments. Participants are shown different versions of the same video sequence and when all the sequences are rated, the following sequence content can be then accessed [48]. In this case, the participants can access any version of the sequence and they can replay, start or stop, change or keep the current score if they prefer. Participants are allowed to view one version at a time and they use the DSCQS method for rating the sequence. Figure 7 illustrates the test organization in the SAMVIQ.

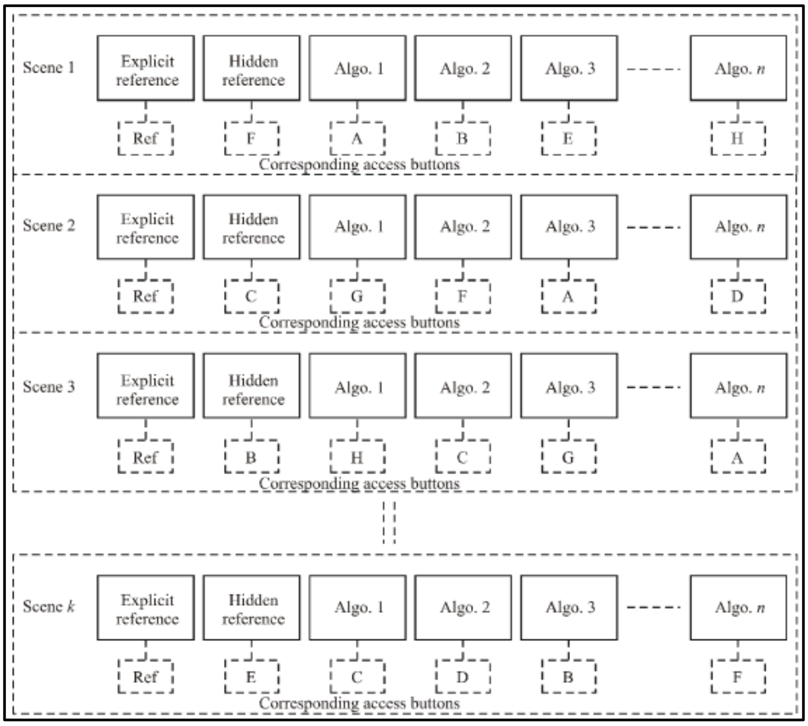

Figure 7: Test organization in SAMVIQ [48]

Pairwise Comparison (PC)

The pairwise comparison methodology is preferred when participants are not always capable to express their view by means of exact rating score. The PC method tries to evaluate all possible combinations of two samples. In the pairwise comparison methodology, participants are shown video sequence pairs one after another. The duration of each video sequence is around 10 seconds. Participants observe the quality of these two video sequences one after another and based on their perception they select the video sequence with higher quality i.e., provide the preference [49]. Preference values are converted to regular quality values (MOS scale) using various algorithms [25].

Results can be stored in a so called comparison matrix (C)

	C1	C2	C3
C1	0	3	1
C2	3	0	2
C3	5	4	0

where, "$C_{ij} = n$" means that condition C_j was preferred over C_i n times.

Table 3 illustrates the comparison between various video subjective test methodologies.

Table 3: Comparison between different video subjective test methodologies [46][49]

	DSIS	DSCQS	SSCQE	SDSCE	SAMVIQ	PC
Explicit reference	Yes	No	No	Yes	Yes	No
Hidden reference	No	Yes	No	No	Yes	No
Scale	Bad to excellent	Bad to excellent	Bad to excellent	Bad to excellent	Bad to excellent	Preference
Sequence length	10s	10s	5min	10s	10s	10s
Picture format	All	All	All	All	All	All
Presentation of test material	I. Once II. Twice in succession	Twice in Succession	Once	Once	Several times	Several times (in pairings)
Minimum participants	15	15	15	15	15	15
Display	Mainly TV	Mainly TV	Mainly TV	Mainly TV	Mainly PC	Mainly PC
Possibility to change score	No	No	No	No	Yes	No
Continuous quality evaluation	No	No	Yes (moving slider)	Yes (moving slider)	No	No

3.5.1.2 Objective QoE estimation

Objective video quality estimation methods are based on media, service, and transmission parameters. These methods mostly rely on mathematical models that can be used when estimating the multimedia QoE based on some objective parameters. The lack of human presence increases error margins for the corresponding estimation. However, the estimation process can be automatic and thus can be performed fast enough. Therefore, service providers and network operators are mostly interested in a tool that can objectively reflect the subjective mean opinion score (MOS) of users with reasonable accuracy.

Objective methods can be classified in three groups [50] according to the availability of the original video sequence for estimation; those are full reference method, reduce reference method, and no reference method.

Full reference (FR)

In the full reference method, the distorted sample is compared with the original sample including per-pixel processing and temporal/spatial alignment. The comparison between the original and distorted samples is then used by measuring device (algorithms) for QoE estimation. Figure 8 illustrates the full reference method.

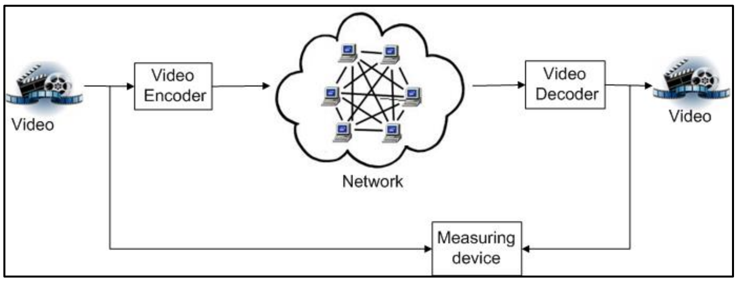

Figure 8: Full reference model

Reduce reference (RF)

The reduce reference method is somehow between the full reference and no reference methods. It is developed to predict the multimedia quality based only on partial information about the original sample. In this case, partial information indicates the features extracted from the original sample. Figure 9 illustrates the reduce reference model.

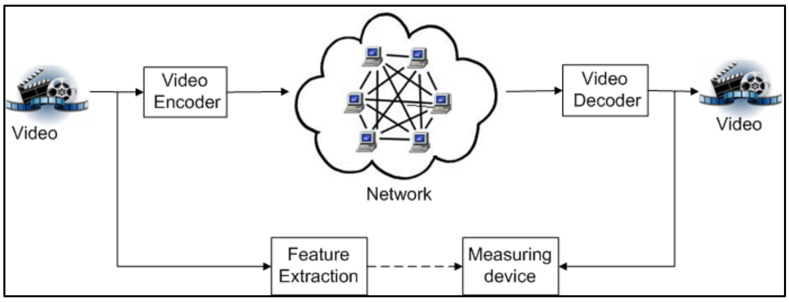

Figure 9: Reduce reference model

No reference (NR)

The no reference method uses a degraded signal for the quality estimation and does not rely on any information about the original reference sequence. Due to the lack of an original sample for the comparison, the accuracy of NR methodology is less than that of the FR or RR methods. Figure 10 shows the no reference model.

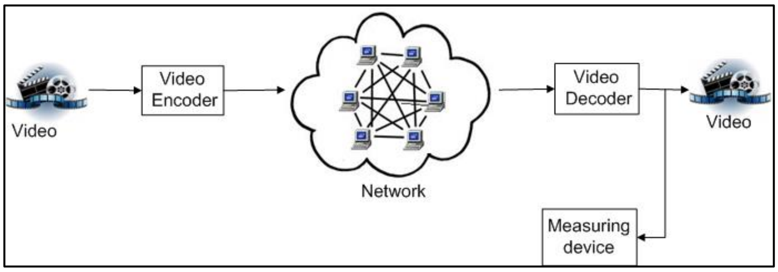

Figure 10: No reference model

3.5.2 Common Objective Techniques used for the multimedia (Video/Audio) QoE estimation

There exist various models for objective quality estimation of a multimedia traffic. These models are classified in five different groups based on the application scenario as well as on the objective input metrics. Different models available for the objective quality estimation assessment are listed below [51].

Media layer model

The media layer model utilizes media signals as an input for the multimedia QoE estimation. This model uses content-dependent features (noise level, interruptions, etc.) for the QoE evaluation. Intrusive speech quality estimation methodology along with FR and RR video quality methodologies falls into the media layer model.

Parametric Packet layer model

The packet layer model relies on the information gathered from IP and RTP headers (IP/RTP packet loss, IP/RTP jitter, etc.) for the multimedia QoE estimation. This model is efficient for estimating the "in-service quality". This is because the computational complexity for this technique is not hard as it does not require media signal to be decoded for the estimation purpose.

Bit stream layer model

The bit layer model exploits the information gathered from the payload bit stream before decoding and the packet header information for estimating the multimedia QoE. This model often needs to decrypt the encrypted multimedia payload therefore, it can have high computation complexity.

Hybrid model

The hybrid model utilizes several or all of the above methodologies for estimating the multimedia QoE. It exploits as much information as possible to estimate the multimedia quality; therefore, it is considered as one of the most effective models for the multimedia QoE estimation.

Parametric Planning model

The parametric model relies on network and terminal quality design and management parameters (for example, coding bit rate, packet loss rate, etc.). These models typically use a mathematical formula, representing the quality estimation as a function of different parameters [52].

Figure 11 illustrates the scope of various objective quality estimation models while Table 4 presents the comparison of different objective quality estimation models with existing standards.

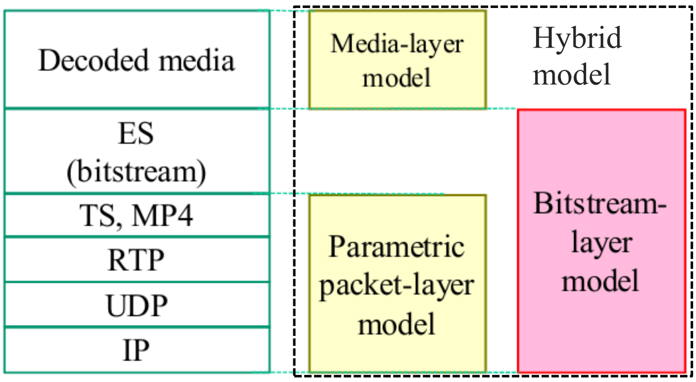

Figure 11: Scope of different objective quality estimation model

Table 4 : Comparison of objective quality estimation models with existing standards [53].

	Media-layer model	**Parametric packet-layer model**	**Parametric planning model**	**Bitstream layer model**	**Hybrid model**
Input information	Media signal	Packet header	Quality design parameters	Packet header and payload	Combination of any
Primary application	Quality benchmarking	In-service nonintrusive monitoring (e.g. network probe)	Network planning, terminal/application designing	In-service nonintrusive monitoring (e.g. terminal-embedded	In-service nonintrusive monitoring

	Media-layer model	Parametric packet-layer model	Parametric planning model	Bitstream layer model	Hybrid model
				operation)	
Video (Standards)	ITU-T J.144 [SD], ITU-T J.vqhdtv [HD], ITU-T J.mm** [PC]	ITU-T P.NAMS [IPTV]	ITU-T G.1070 [videophone], ITU-T G.OMVS [IPTV]	ITU-T P.NBAMS [IPTV]	ITU-T J.bitvqm [IPTV]
Multimedia (Standards)	ITU-T J.148				

3.6 Subjective test environment

Subjective tests involve human participants and require appropriate environment to perform the test. Test environment should be unbiased, should not be affected by external noise, and should be close to real world scenario. Most of the research works use lab environment for subjective tests, however, crowdsourcing environment is getting popular as well [10]. In the following sections, we discuss laboratory and crowdsourcing environments in detail.

3.6.1 Laboratory environment

The laboratory experiment provides a controlled environment for performing subjective tests for evaluating the multimedia quality. Different parameters associated with the test like noise-level, distance between screen and users, screen size, etc. can be easily controlled according to the requirements. However, lab based experiments have limitations such as 1) high cost in terms of time and labor 2) limited participants diversity [22]. A laboratory experiment takes weeks for preparing tests, recruiting users, scheduling time slots for supervising the experiments, etc. Also users need to be physically present in the laboratory to perform the test [22]. Generally lab tests are performed in university or research laboratory so the participants for the test are the either students or researchers. This limits the diversity of participants: most of them are boys which are young and well-educated, etc.

3.6.2 Crowdsourcing environment

In crowdsourcing environment, subjective experiments can be performed from distance and there is little control over the participant's environment. It is computer software assisted method generally performed on a web platform. This methodology mainly involves collecting subjective assessment of quality through ubiquitous streaming via the Internet. Different screen test and cheat detection mechanisms are employed in a crowdsourcing experiment [23] [24] to make the test reliable and resilient. Crowdsourcing based subjective experiments have gained attention to replace the needs of lab-based tests and these experiments offer promising correlation with the former [25]. This allows an investigator to get opinions from a vast variety of subjects; in a time-flexible, test-data size scalable, and swift manner [10]. Different crowdsourcing software available for multimedia subjective test are QualityCrowd [26], PC-Video-Test-Interface [10], etc.

3.7 QoS/QoE learning algorithms

The QoS parameters reflect objective network and service level performance and they do not directly address the user satisfaction of the delivered service or application. On the other hand, QoE become an important indicator, useful for network operators and service providers to help them understand the user acceptability towards a particular service or application. The paradigm is shifting towards user-centric evaluation of a service or application performance. To attract or bind users to a service, real time estimation of QoE is a must for network operators and service providers. Therefore, it is necessary to derive a correlation between the QoS parameters and the QoE, which can be used to identify the impact of different QoS parameters on the QoE of the users and moreover, objectively estimate the QoE. However, the relationship between QoS and QoE is hard to estimate, since this relationship is not linear. Moreover, higher QoS level does not always yield the higher QoE value. In Table 5 we present different QoS/QoE correlation methodologies available in the literature.

Table 5: A summary of the selected evaluation approaches for each model [151]

Aspects/ Models	AQoS Parameters	NQoS Parameters	Video coding Parameters	Subjective Metrics	Objective Metrics	Learning Technique	Simulation / Test bed	Technology	Reference Measurements
Fiedler et al [134]	×	×		×			Simulation /test-bed	Wired	RR
Siller et al [135]		×	×	×	×		test-bed	Wired	FR
Wang et al [55]		×			×		Simulation /test-bed	Wired	NR
Agboma et al [136]	×			×			test-bed	Wireless	NR
Menkovski et al [137]	×			×		×	Simulation /test-bed	Wireless	NR

Aspects/ Models	AQoS Parameters	NQoS Parameters	Video coding Parameters	Subjective Metrics	Objective Metrics	Learning Technique	Simulation / Test bed	Technology	Reference Measurements
Machado et al [138]		×		×	×	×	Simulation	WiMax	FR
Du et al [139]		×		×	×	×	Simulation /test-bed	Wired	FR
Kim et al[140]		×		×			Simulation /test-bed	Wired	NR
Khan et al[141]	×	×			×	×	Simulation	Wireless	NR
Han et al [143]	×	×			×		Simulation	Wireless	NR
Laghari et al [124]	×	×		×		×	test-bed	Wireless	NR
Ramos et al [144]		×	×		×		Simulation /test-bed	Wired	NR
Koumaras et al [145]		×	×		×		Simulation /test-bed	Wired	NR
Frank et al[146]	×	×	×	×		×	Simulation	Wired	NR
Calyam et al [147]	×	×	×	×	×	×	test-bed	Wired	NR
Mok et al [105]	×	×		×			test-bed	Wired	NR
Elkotob et al [148]	×	×		×			test-bed	Wireless	NR
Mushtaq et al [149]		×	×	×		×	Simulation/ test-bed	Wired	NR
Hoßfeld et al [150]	×	×		×			test-bed	Wired	NR
Staelens et al [79]		×	×	×		×	Simulation	-	NR
Kang et al [81]	×	×	×		×	×	Simulation	Wireless	NR
Cherif et al [82]		×		×		×	Test-bed	Wired	NR

AQoS - Application level QoS, **NQoS** -Network level QoS, **Subjective metric** - MOS, **Objective metric** - PSNR, SSIM, VQM, etc.

A number of models have been developed for automatic evaluation of user satisfaction however, the accuracy of these models are still questionable. Moreover, nowadays, people study human cognitive processes very carefully and try to elaborate models that behave similar to brain neurons. Since a human mind is known to be non-deterministic, it is challenging to get a formal algorithm of the human behavior. That is the reason why researchers turn their attention to self-adaptive models and learning algorithms.

Table 6 provides the comparison of estimation techniques in terms of their modeling capabilities [130].

Table 6: Comparison of estimation techniques in terms of modeling capabilities [130]

Technique	Model Free	Can resist outliners	Explains output	Suits Small data sets	Can be adjusted for new data	Reasoning process is visible	Suit complex model	Include known facts
Least Square Regression	N	N	P	N	N	Y	N	P
Robust Regression	N	Y	P	P	N	Y	N	P
Neural Networks	Y	N	N	N	P	N	Y	P
Fuzzy Systems	Y	P	Y	Y	P	Y	Y	Y
Hybrid-Neuro-Fuzzy system	Y	P	Y	P	P	P	Y	Y
Rule Based Systems	N	N/A	Y	N/A	N/A	Y	Y	Y
Case-Based Reasoning	Y	P	Y	P	Y	P	Y	N
Regression Trees	Y	Y	Y	P	Y	P	Y	P
Decision Trees	Y	Y	Y	P	Y	P	Y	P

(Yes = “Y”, No= “N”, Partially = “P”)

Various QoS/QoE learning algorithms can be found, for example, in [79], [80], [81], [137], [138], [139], [141], [142], [146], [147], [149]. These models use different learning algorithms in different use cases and claim to provide higher accuracy. In Table 7, we present some of the models that use learning algorithms for QoS/QoE estimation.

Table 7: QoS/ QoE models using learning algorithms

Models	Learning algorithm	Subjective test	Objective metric	QoS parameters	Correlation	Conclusion
Menkovski et al [137]	Support Vector Machine (SVM), Decision Tree (DT)	Yes		SI, TI, video bitrate, video framerate	-SVM accuracy Mobile-88.59±2.85%, PDA-89.38±2.77%, Laptop-91.45±2.66%; - DT, accuracy Mobile-93.55±1.76%, PDA-90.29±2.61% Laptop-95.46±2.09%	-Classification based on acceptable or unacceptable QoE (2 class)
Machado et al [138]	Artificial neural network	-	MOS from Evalvid tool	jitter, delay, throughput, packet loss	-	-It was observed that the neural network for the scenario had a very good prediction. - The influence of the video dynamics and the amount of nodes is perceptible,
Du et al [139]	BP (Back-Propagation) Neural Network	-	PSNR, DMOS Video Quality Evaluation System	Delay, Jitter, Loss, Burst, Bandwidth, Congestion Period, Disorder Packets	-	-No subjective test -visualizes the relationship between the network technical parameters and DMOS
Khan et al [141]	Adaptive Neural Fuzzy Inference System (ANFIS) & non-linear regression	-	PSNR to MOS conversion	Content Type, video Sender Bitrate, and frame rate, Block Error Rate, Mean Burst Length	ANFIS-0.87 Regression-0.86	-It is possible to predict the video quality if the appropriate parameters are chosen. - The content type has a bigger

Models	Learning algorithm	Subjective test	Objective metric	QoS parameters	Correlation	Conclusion
	analysis			(MBL).		impact on quality than the sender bitrate and frame rate. - They found that the video quality is more sensitive to network level parameters compared to application level parameters.
Frank et al[146]	Artificial neural network	Yes	-	Frame Rate , Bit Rate , Quantization scale, and the size of the group of pictures, Packet Loss Rate , Mean Burst Loss (MBL), and Jitter	4% error rate	-For the evaluated scenarios, the codec related parameters seem to have very little influence on the MOS score. The packet loss rate and jitter are strongly related with the perceived quality, which reinforces the need to have a well dimensioned network and some control (QoS) mechanisms to protect the video streams if we want to have acceptable quality levels.
Calyam et al [147]	Neural network	Yes		Loss, jitter	> 0.9	-The model showed the impact of PFIFO and TFIFO router queuing disciplines on multimedia QoE in IPTV content delivery networks and developed separate models for each of them. -The model speed results suggest

Models	Learning algorithm	Subjective test	Objective metric	QoS parameters	Correlation	Conclusion
						that over 10,000 flows can be handled in < 4ms.
Mushtaq et al [149]	Naives Bayes (NB), 4-Nearest Neighbour (4-NN), Support Vector Machine (SMV), Decision Tree (DT), Random Forest (RF) and Neural Network (NNT)	Yes		gender, frequency of viewing, interest, delay, jitter, loss, conditional loss, motion complexity and resolution	Mean absolute error rate DT= 0.126 RF= 0.136 NNT= 0.17 4-NN= 0.2 NB=0.22 SVM= 0.26	-Decision tree and random forest outperform other methodology.
Kang et al [81]	Neural network (RFB)	-	PSNR to MOS	bit rate, frame rate , resolution, packet loss rate, video content features,screen size of terminal equipment.	0.89	-No-reference, content-based QoE estimation model for video streaming service over wireless network -Higher accuracy
Cherif et al [82]	Neural network (RNN)	Yes	-	Packet loss percentage	Around 0.9	-A_PSQA correlates very well with the subjective scores for almost all video sequences compared to the other Full- and No-ref. tools.
Staelens et al [79]	Decision tree	Yes	-	information extracted from the network and the received encoded (impaired) video stream	Accuracy 83% -plus content classification 86% -with full bitstream processing 85%	-Classification based on visible or invisible impairment -Content classification gives higher accuracy -Errors are less visible in low motion.

Models	Learning algorithm	Subjective test	Objective metric	QoS parameters	Correlation	Conclusion
Laghari et al [124]	Rough set theory (RST)	Yes	-	-Packet loss, packet reorder, delay, bitrate		-The different types of content require different level of QoS support. - Using RST, any set of raw data can be turned into usable date and important core attributes could be found easily with considerable accuracy -Qualitative and quantitative QoE data assessment

In this section, we briefly discuss the three QoS/ QoE learning algorithms that are used in our works.

3.7.1 Neural networks for QoE prediction

Neural networks are widely used for solving various problems in the artificial intelligent area. Neural networks are composed of interconnected processing elements (neurons/nodes) working together to solve specific problems. The network can be divided into different levels (layers) which contain a number of neurons and each level can be treated as a collection of states. Usually, for each neuron there exists a formula (firing rule) that calculates its output according to weighted inputs that is used when coming to the next level via weighted edges. The firing rule decides whether a neuron should be fired for any given input patterns.

The neuron operates in two modes i.e., training mode and use mode. In the training mode, the neurons are trained according to the input patterns, so that the actual output of the network matches the desired output with minimum error. In the use mode, if the network detects known pattern (taught input), its current outputs is an associated output. Otherwise, the firing rule is used to determine whether to fire or not.

In the following section we describe random neural network that is often used for QoS/QoE correlation and QoE estimation.

Random neural network [42]

Generally 3-layer (1 input layer, 1 hidden layer, and 1 output layer) feed forward RNN is considered for the QoS/QoE correlation purpose. The input layer corresponds to the QoS parameters and the output layer represents the QoE score. The output of the network can be represented by a function f (QoS), which is calculated as:

$$f\,(\text{QoS}) = \frac{\sum_{h=1}^{H} Q_h \quad W_{ho}^{+}}{r_o \quad + \quad \sum_{h=1}^{H} Q_h \quad W_{ho}^{-}}$$

$$\text{where, } Q_h = \frac{\sum_{i=1}^{I} C_i r_i^{-1} \quad W_{ih}^{+}}{r_h + \quad \sum_{i=1}^{I} C_i r_i^{-1} \quad W_{ih}^{-}}$$

Here, C represents input QoS parameters C = (C_1, ⋯, C_I), h represents hidden nodes h=1….H, and O represents the output node. Q_h is the activity rate of the hidden neuron h. The strictly positive numbers r_o, r_h for h =1…..H and r_i for i =1…..I correspond to the firing rates of the neurons in the network (respectively, for the output one, the hidden nodes, and the I input ones). The weights are the variables tuned during the learning process. We denote by $w+_{uv}$ (resp. by w^{-}_{uv}) the weight corresponding to an exiting (resp. inhibiting) signal going from neuron u to neuron v (observe that both numbers $w+_{uv}$ and $w-_{uv}$ are ≥ 0). Learning will thus consist of finding the appropriate values of the weights capturing the mapping from ($c^{(k)}$ to the real number $q^{(k)}$ where $q^{(k)}$ is the quality given by a panel of human observers to some audio or video sequence (depending on the application) when the source parameters had the values present in $c^{(k)}$ for k =1..K [42].

3.7.2 Fuzzy logic for the QoE prediction

Fuzzy logic was introduced by Lotfi A. Zadeh [60] in 1956 and it extends the classical set theory to the so called *fuzzy set* theory. In the general set theory, a set *A* can be specified by its *characteristic function* χ, and the set *A* contains an element *a*, i.e., $a \in A$ if $\chi(a) = 1$, otherwise $\chi(a) = 0$ and $a \notin A$. If *A* is a *fuzzy* set then the χ function is not a classical predicate but its value belongs to an interval between 0 and 1, i.e., $\forall\ a\ \chi(a) \in [0, 1]$. In this case, the χ function shows the degree *how likely that the statement* '*a* belongs to *A*' *holds*. In other words, differently from classical Boolean algebra, a partial truth is introduced in the fuzzy logic where the truth degree may range between *completely true* and *completely false* [61].

Fuzzy logic expert systems are one of the well-known estimation/prediction techniques that is used for making decisions based on imprecise/ambiguous information in various fields [62]. For instance, Adeli and Neshat [131] proposed a fuzzy expert system approach for diagnosis of heart diseases, while in [132] a fuzzy expert system is used for a hotel selection. Usually, the aim of a fuzzy logic expert system is to draw a concise result based on ambiguous information. Fuzzy logic expert system has three main components, namely *fuzzifier*, *fuzzy interference engine* and *defuzzifier* as shown in Figure 12 [63].

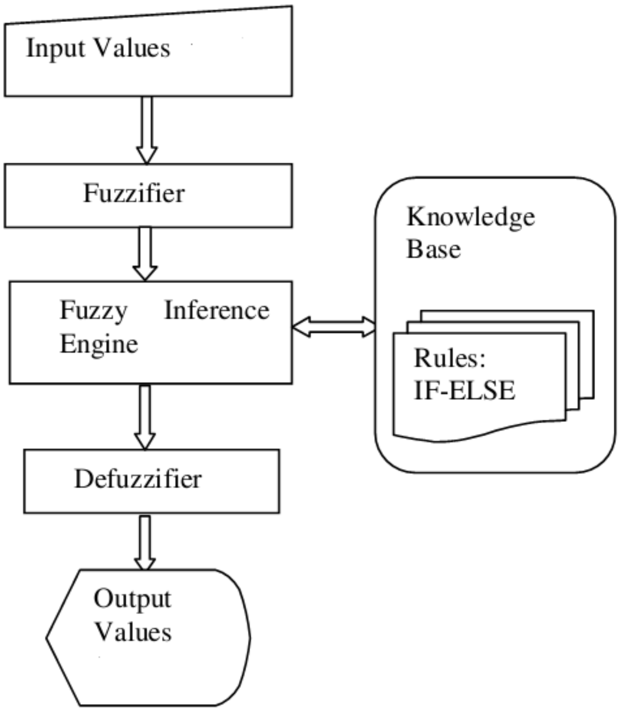

Figure 12: A fuzzy expert system

The *fuzzifier* contains the membership functions (fuzzy sets). In the *fuzzifier*, input parameters (service parameter values) are mapped into membership functions to determine the membership of these parameters to appropriate fuzzy sets. The *fuzzy inference engine* contains a collection of IF-THEN rules, which are obtained from experts or learnt using other intelligent techniques. The inputs taken from the fuzzifier (i.e., membership values) are applied to the antecedents of the fuzzy rules. The obtained value is then applied to the consequent membership function (i.e., the output QoE membership function). In other words, the consequent membership function is clipped or scaled to the level of the truth value of the rule antecedent. If more than one rule is triggered from one set of input parameters, the outputs of all the rules are aggregated into the *aggregated output fuzzy set*. The *defuzzifier* is used to perform a *defuzzification*, namely a single output value is obtained from the defuzzifier with the use of the *aggregated output fuzzy set*. There exist various defuzzification techniques [64] such as the centroid method, the weighted average method, the maximum method, etc.

The fuzzy expert system significantly depends on the membership functions and inference rules. The more accurately the membership functions and inference rules are specified the higher is the prediction ability of the expert system. Therefore, in order to effectively apply the fuzzy logic expert system for the QoE estimation, one has to carefully specify the membership functions as well as the inference rules. These membership functions and inference rules can be derived based on subjective data set; namely statistics gathered from end-users and/or based on experts' knowledge and recommendations.

Once the initial machine (fuzzy expert system or a logic circuit) is derived, one may use it to predict the QoE value. However, the new statistical data can also play a role in 'upgrading' the system by increasing its prediction ability. To improve the fuzzy logic expert system, fuzzy membership functions and inference rules are updated based on the new subjective data set.

In this following, we describe how a fuzzy logic expert system can be learned or trained for predicting the QoE value of multimedia services. Figure 13 represents the basic block diagram of the fuzzy logic expert system construction.

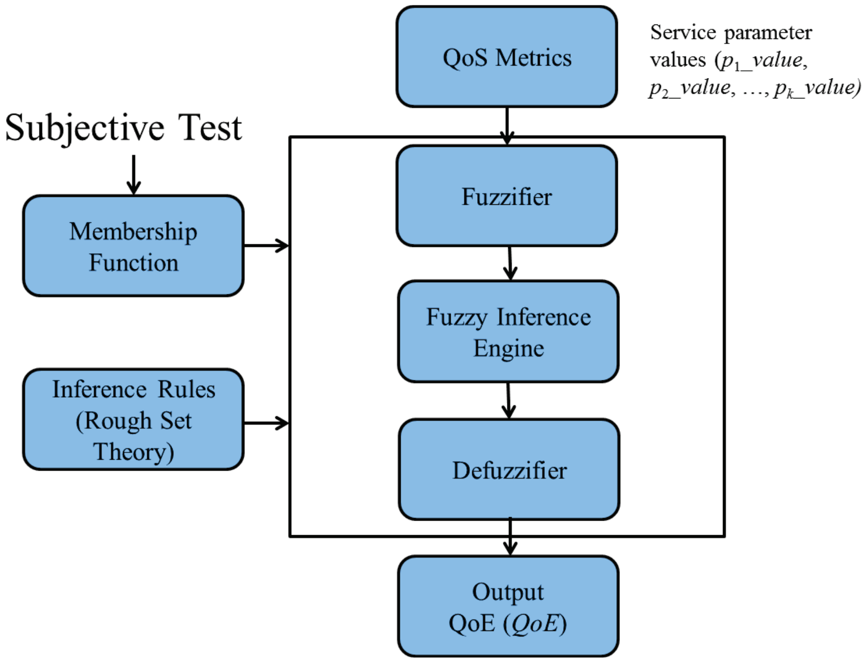

Figure 13: A fuzzy logic expert system for QoE estimation

We further discuss how such a system can be used when estimating the quality of a multimedia service. Given a multimedia service W and a collection of service parameters p_1,

p_2, …, p_k that are used for the QoE evaluation, the output QoE of the multimedia service is represented by the variable *QoE(W)*.

Each service parameter p_i, $i \in \{1, …, k\}$ is classified into A_j classes, $j \in \{1, …, l\}$, and a membership function $\mu^{pi}{}_{Aj}(x)$ is derived for each class of service parameters. In this thesis work, we use the MOS score, thus $l = 5$ and the score belongs to the set {*excellent*, *good*, *fair*, *poor*, *and bad*}. Similar classes are considered for the output QoE. As mentioned above, the membership functions and inference rules are constructed based on the subjective data set, i.e., statistical data provided by end-users and/or by experts.

In order to estimate the QoE of a multimedia service, different patterns (p_1*_value*, p_2*_value*, …, p_k*_value)* are submitted to the fuzzy logic expert system. The output of the system represents the QoE value for a multimedia service. Below, we provide the algorithm for evaluating the QoE value based on the fuzzy logic expert system.

Algorithm for evaluating/predicting the QoE value for multimedia service

Inputs: A multimedia service *W*, service parameters p_1, p_2, …, p_k and their values p_1*_value*, p_2*_value*, …, p_k*_value*;

Inference rules given in the form: IF <antecedent> THEN <consequent>;

Membership functions $\mu^{pi}{}_{Aj}(x)$, $i \in \{1, ..., k\}$, $j \in \{1, ..., l\}$, where x is the service parameter value and A_j represents different classes of corresponding parameter values.

Output: The QoE value (*QoE*)

Procedure:

1. Map the service parameter values p_1*_value*, p_2*_value*, …, p_k*_value* into the membership functions $\mu^{pi}{}_{Aj}(x)$ and get the membership value in the *fuzzifier.*

2. Apply the justified inputs (membership values) taken from the fuzzifier to the fuzzy inference rules in fuzzy inference engine.

Rule: IF p_i is A_j AND p_i is A_j then QoE is A_j.

The output of the rule evaluation is the consequent membership function (*output fuzzy set*) clipped or scaled to the level of the truth value of the rule evaluation. If more than one rule is triggered the outputs of all the rules are aggregated into the *aggregated output fuzzy set*.

3. Defuzzify the aggregated output fuzzy set and get a single QoE value (*QoE*) in the defuzzifier.

4. **Return** *QoE value*.

There exist various defuzzification techniques (Step 3) [64], for example, the centroid method, the weighted average method, the maximum method, etc., and, in this work we use the *centroid* method. The mathematical basis of this method relies on the *center of gravity (COG)* that can be expressed by the following formula.

Center of Gravity (COG)

$$y = \frac{\int \mu_i(x) \bullet x dx}{\int \mu_i(x) dx},$$

where y is the defuzzified output, $\mu_i(x)$ is the aggregated membership function and x is the output variable.

3.7.3 Rough set theory

Rough set theory (RST) is one of the tools for machine learning. It provides algorithms, explanations and some theoretical basis for the research on learning [123]. It was first proposed by Prof Z. Pawlak in 1980. It can be used for extracting knowledge or decision from uncertain and incomplete information.

RST is used for discovering patterns, rules and knowledge in data. RST has many advantages, such as it does not have information loss, is flexible and extendable as compared with other data mining technologies [124]. It has also found application in data mining, policy-making analysis, process control, and pattern recognition.

Basic definitions and concepts of RST

In rough set theory we define data in a form of information system (*S*). Let $S = (U, A, V, f)$ where U and A represent non-empty universal set. Let U has finite n objects $\{x_1, x2,...n\}$ and A is non-empty finite attribute set *(a1,a2…n).*One attribute set corresponds to one equivalence relation, i.e., $A= C \cup D$, and $C \cap D = \emptyset$, where *C* (in our case QoS parameters) is the conditional attribute set and *D* decision attribute set (in our case MOS scores) . Here V_a represents the domain value for the attribute set *a* and similarly, *f* represents the information function.

Some of the important properties of RST which are used to classify and reduce data to important data and achieve CORE influencing factors are listed below.

Properties

Indiscernibility of objects:

If two objects are characterized by same information then they are indiscernible (similar). In RST it is defined as

$$IND(C) = \{(x, y) | (x, y) \in U^2$$
$$, \forall a \in C\ (a\ (x) = a(y))\}$$

Here, the objects x and y are indiscernible with set of condition attributes C denoted by $IND(C)$. That means they are inseparable with condition attributes C. $IND(C)$ splits the given set of users in the survey (U) into a family of equivalence classes $\{X_1, X_2, \ldots X_r\}$ called elementary sets.

Rough set Approximation:

With any rough set a pair of precise sets, called the lower and the upper approximation of the rough set, is associated. Let $P \subseteq A$ is a set of conditional attributes and $X \subseteq U$ is set of users, then

$$P_*X=\{x \in U: [x]_P \subseteq X\}$$
$$P^*X=\{x \in U: [x]_P \cap X \neq \varnothing\}$$

The lower approximation P_*X consists of all objects which surely belong to the set and the upper approximation $P^* X$ contains all objects which possibly belong to the set. The difference between the upper and the lower approximation $P^* X$- P_*X constitutes the boundary region of the rough set [5].

Reduct and CORE:

A reduct is a minimal set of attributes discerning one object from all objects with a different decision. A reduct refers to a decision rule. Finding all reduct is an NP-complete problem. Different algorithms like greedy algorithm and genetic algorithm are used to search for reduct. CORE is the intersection of all the reducts and is the set of indispensably important attributes.

Decision rules:

IF-THEN decision rules are generated by reading values for each attribute in the reduct. With every decision rule two conditional probabilities, called the accuracy (i.e., certainty) and the coverage coefficient, are associated. The accuracy coefficient expresses the conditional probability that an object belongs to the decision class specified by the decision rule, given it satisfies conditions of the rule. The coverage coefficient gives the conditional probability of reasons for a given decision [13]. We calculate support, accuracy and coverage of condition attributes from [13] corresponding to decision rules.

Support of a rule:

$$\sigma_x = {supp_x(C,D)}/{|U|}$$

Accuracy of the rule:

$$cer_x(C,D) = \frac{\sigma_x(C,D)}{\pi(C(x))}$$

Where,

$$\pi\big(C(x)\big) = \frac{|C(x)|}{|U|}$$

Coverage factor of decision rule:

$$cov_x(C,D) = \frac{\sigma_x(C,D)}{\pi(D(x))}$$

Where,

$$\pi\big(D(x)\big) = \frac{|D(x)|}{|U|}$$

3.8 Evaluation metrics for QoE prediction techniques

In this section, we shortly discuss how QoE prediction techniques can be evaluated with respect to two main criteria: prediction accuracy and self-adaptability. The prediction accuracy refers to the ability of the model's estimated score to match that of the subjective QoE. Therefore, for a model to have high prediction accuracy, the score difference should be minimal. On the other hand, self-adaptability refers to the ability of the model to adapt to new dataset. An efficient model should automatically adapt (re-organize) to the new dataset without much complexity and time. However, both accuracy and self-adaptability highly depend on the size and coverage of the learning dataset and the number of considered input parameters.

Furthermore, in spite of the fact that machine learning algorithms are widely used for evaluating user satisfaction, almost all these models rely on statistics provided by high quality experts and/or collected from ordinary users. Collecting these statistics is the most critical task since it requires time and cost. Moreover, when training a model there are a number of additional points specific to the learning algorithms that need to be considered like defining the number of branches of a decision tree or designing membership functions and inference rules of a fuzzy expert system.

3.9 Multimedia QoE standardization bodies

There are different multimedia QoE standardization bodies, industry forums, and others that work on different aspects of multimedia QoE. These include definitions and terms of reference, requirements, recommended practices, test plans, and many more [66]. Some of the most active ones are discussed below.

3.9.1 Video Quality Experts Group (VQEG)

The Video Quality Experts Group (VQEG) works in the field of multimedia QoE assessment (particularly video) [67]. It was founded by ITU-T and ITU-R group members in 1997. The group works in the field of video quality assessment and investigates different new and advanced subjective and objectives methods and measurement techniques (VQEG). Moreover, they plan, test, and validate objective quality estimation. VQEG is an open group and does not require fees, membership applications, or invitation to join. Some of the past projects of VQEG groups are FRTV Phase I, FRTV Phase II, multimedia phase I, RRNR-TV etc. Some of the active projects of VQEG groups are 3DTV, Audiovisual HD (AVHD), HDR (High Dynamic Range Video), Hybrid Perceptual/Bitstream JEG-Hybrid, MOAVI (Monitoring of Audio Visual Quality by Key Indicators), Quality Recognition Tasks (QART), RICE (Real-Time Interactive Communications Evaluation), and Ultra HD (VQEG).

3.9.2 ITU-T

The ITU-T (Telecommunication Standardization Sector of the International Telecommunications Union) is the primary international body working in the field of standardization of telecommunications equipment and systems [68]. ITU-T is a part of ITU (International Telegraph Unit). This organization is based on public-private partnership and requires membership to be involved in. There are different study groups inside ITU-T that work for multimedia quality standardization work. ITU-T Study Group 9 is focused on studies of cable television and quality assessment methods for video and multimedia services. ITU-T, Study Group 12 (SG12) is mainly studying QoE requirements and assessment methods for multimedia services including IPTV [53]. A joint group has been established i.e., Joint Rapporteur's Group on Multimedia Quality Assessment (JRG-MMQA) to harmonize the work of these two study groups. IUT-T works on all areas of multimedia QoE assessment i.e., speech, audio, video, and multimedia (combined) and has developed different standards for a quality assessment of multimedia for example ITU-T J.148, ITU-T P. NBAMS, ITU-T P.862, etc.

3.9.3 ATIS IIF

The ASTIF-IIF [69] is working with Industry segments to define necessary standards and specifications for IPTV network architecture, QoS and QoE, Security, and Interoperability (ATIS-IIF). It requires membership for involvement. Different partners working together in ATIS IIF include service providers and manufacturers, content and entertainment providers, manufacturers, and the entire IPTV industry ecosystem. The Quality of Service Metrics Committee (QOSM) inside ATIS IIF works in different issues related to QoS and QoE in IPTV services. It has issued different documents related to QoS and QoE in IPTV services like ATIS-0800008, ATIS-0800004, etc.

3.10 Future challenges and issues related to multimedia QoE assessment

Multimedia QoE assessment and monitoring is essential to deliver an optimized end to end high QoE service. This requires a deep understanding and efficient identification of different objective and subjective parameters that impact the user experience. Multimedia content delivery is a large and continuously evolving field that involves various actors from content service providers to Internet service providers, and to content consumers (users) themselves. Therefore, a comprehensive QoE assessment requires the understanding, the role, and impact of these actors on multimedia content from delivery till consumption [70]. A multi-disciplinary approach involving different measures at the server, network, application, or user levels for a wide range of objective (QoS) and subjective (user perception) metrics is necessary for building QoE assessment models. A typical process for building such model includes:

- Conducting subjective lab or crowdsourcing tests to evaluate user perception in different scenarios. As the number of impacting parameters is relatively high, the objective of subjective tests is to measure user acceptability with respect to a limited number of parameters like screen size change, player buffering strategy, network conditions change, etc.
- Building correlation model to map between parameters (like QoS parameters) measured during the subjective tests with the QoE scores given by subjects. This phase is considered as the learning phase.
- Evaluating the model against user scores to measure its accuracy.

The QoE assessment model requires the extraction of QoS parameters from different points of the network. For this perspective, the measurement of potential QoS parameters plays a key role in providing the required input data for the quality estimation model. Such measurements can be achieved by installing network monitoring probes on key points in the network infrastructure. These tools deploy Deep Packet Inspection (DPI) techniques to extract relevant parameters from the network traffic. When the traffic is encrypted, DPI needs to decrypt the content in order to extract those relevant parameters. In most of the cases, this operation is simply impossible. Therefore, new models using statistical properties of the network traffic need to be designed.

Furthermore, with the push for personalized and user centric services, there is a pressing need for QoE estimation probes capable of processing high speed network links in real time to extract QoS parameters and correlate them with user perceived QoE. However, the relationship between QoS and QoE is fuzzy and non-linear. To address this issue there are large numbers of intelligent algorithms proposed in the literature however, there is still room for innovative mechanisms to efficiently correlate QoE from QoS, in real time. Finally, efficient QoE management systems aim at reacting before the user even notices the quality degradation [71]. This requires an efficient feedback loop that can detect, locate and react in real time to degraded network conditions by controlling or reconfiguring different components of the QoE estimation system. Big advances have been made in this direction, however, open questions like when, where, and how to react still need to be addressed.

Chapter 4

QoE estimation of VOD services: Network Layer perspective

4.1 Introduction

IPTV (Internet Protocol Television) is a means by which television services are transmitted using IP suite over a packet switched network. According to the International Telecommunication Union focus group on IPTV (ITU-T FG IPTV), "*IPTV is defined as multimedia services such as television/video/audio/text/graphics/data delivered over IP based networks managed to provide the required level of quality of service and experience, security, interactivity and reliability*". IPTV can deliver both Live TV as well as Video-on-Demand (VoD) services.

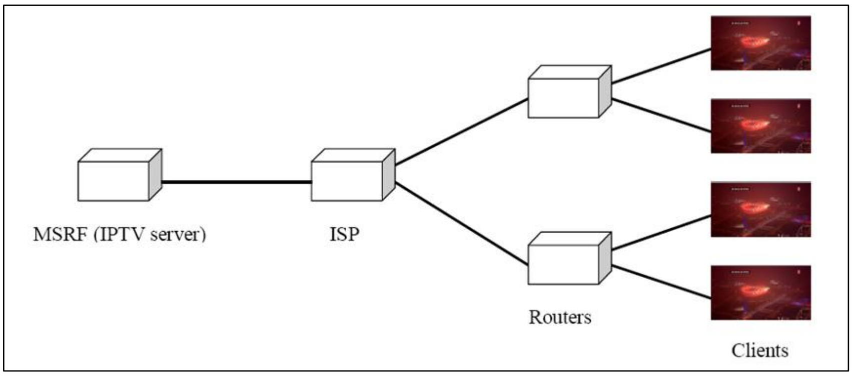

Figure 14: Multicast IPTV (reference)

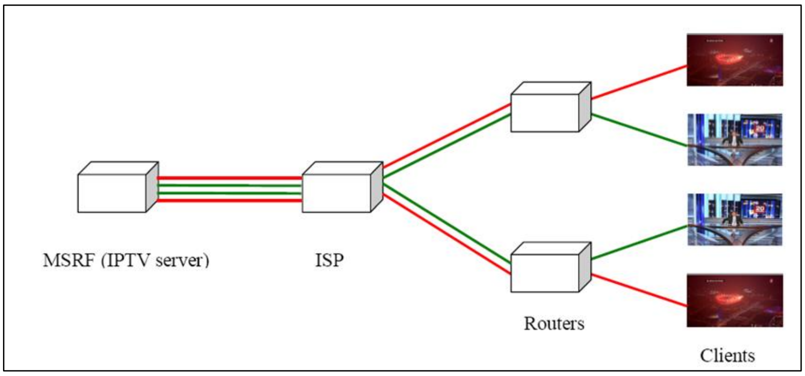

Figure 15: Unicast IPTV (VoD)

IPTV supports both multicast and unicast delivery of content as shown in Figure 14 and Figure 15 respectively. Multicast IPTV enables TV content provider to deliver TV content in a single stream (using broadcast technology) to many customers at the same time. Since only one stream is transmitted over the network, it saves a considerable amount of bandwidth. Using IGMP (Internet Group Management Protocol) protocol [72], clients can receive multicast packets and enable the routing of the broadcast stream to their network device through the network. Unicast IPTV is a point to point delivery. There is a separate content stream on the network for each unicast session for each user. VoD is an example of Unicast IPTV where customers request for particular multimedia content and receive it on their TV sets. Since the server needs to send the content to each user in separate streams, unicast IPTV is usually not bandwidth efficient.

Since IPTV multimedia flows share the same bandwidth with other internet traffic, there is a high possibility of network congestion. Moreover, multimedia contents require larger bandwidth and are very sensitive to timing schedule. Therefore, the lack of network resources can lead to packet losses, jitter, delay, etc. with the corresponding result in multimedia impairments.

RTP is generally used as a multimedia transport protocol since it addresses the time critical requirement of multimedia bit streams. RTP provides a timestamp and sequence number to IP packets containing media data. This can be used by the receiver to synchronize play back and manage buffers minimizing network jitter [73].

In this chapter, we focus on a specific scenario with VoD services that use RTP/UDP as transport protocol and analyze the impact of network perturbation on video QoE. Moreover, we will present a real-time video QoE estimation system based on network QoS parameters. The accurate estimation of video quality requires access to both the original and the received video streams, however this is unsuitable for real-time monitoring. We are aiming at a real time QoE estimation tool that does not require complete decoding of video stream and access to original reference video, it is based on network QoS parameters that accurately estimates video QoE. This permits monitoring to be performed at intermediate measuring points along the networks path.

4.2 Related works

The ability to identify the perceived degree of video impairment due to network perturbation is a key point in the quality estimation of video traffic. Moreover, the impact of network perturbations on video can range from distortion-less to intolerable distortion. Moreover, not all network impairments result necessarily in a visible degradation, which has also been accounted for in ITU-T G.1080 [77] and TR-126 [78]. Therefore, measuring the impact of network perturbation on the quality of the video traffic is a challenging task as shown in several works [74] [75] [76].

There have been a number of research works done with the objective of video traffic QoE estimation based on media parameters (PSNR [127], VQM [114], SSIM [128], and PEVQ [129]) however, very limited work can be found on QoE estimation of video transmission from the network perspective without any media parameters.

Media Delivery Index (MDI) [80] is a scalable metric for assessing the effect of delivery network on the video and can be measured from any point between the video end points. It has two components, delay factor (DF) and media loss rate (MLR), both using packet loss and jitter as predictors of video quality. However, it does not consider users perception in quality estimation and only gives an objective metric for quality evaluation. Furthermore, the work presented in [140] proposed a numerical formula to evaluate QoE using different QoS parameters such as packet loss, burst loss, jitter, delay, GoP (Group of Picture) length. However, it also lacks user perception and experimental and validation results. Therefore, to incorporate the user perception in quality estimation, a combinational (hybrid) approach which involves both objective and subjective approaches is proposed. In this combinational approach, a subjective test is performed to create a dataset which represents the relationship between objective metrics and subjective QoE. Following this, different intelligent systems are trained with this subjective dataset to objectively predict the QoE. These methodologies rely on the correlation between objective metrics and subjective QoE. However, the relationship between objective QoS metrics and QoE is fuzzy and non-linear, and hard to calculate. Therefore, to address this issue there are numbers of intelligent algorithms proposed in the literature however, there is still room for innovative mechanisms to efficiently correlate QoE from QoS in real time. Most of the intelligent algorithms used for this purpose are based on neural networks [81] [82] [83] however, neural networks are computationally complex, requires large training data and more training time, and moreover, the reasoning process is not transparent. Decision tree [79] based learning approach was proposed for multimedia quality estimation based on bit stream information (QoS parameters), however, decision trees only partially suit for small datasets, small variations in the dataset need a regeneration of the tree and also the reasoning process is not completely transparent.

In this thesis work, we propose a no-reference QoE estimation system based on fuzzy logic [84] to estimate the impact of the network conditions on the video quality, i.e., the QoE. An advantage of using fuzzy expert systems is that they are simple, computationally less intensive and reasoning process is transparent. It can be seen in Table 6 in section 37, that fuzzy system outperforms other estimation techniques in terms of modelling capabilities. Moreover, they are good at making decisions with imprecise information. We consider three QoS metrics (packet loss rate, packet loss burstiness, and jitter) as the network condition indicators. However, it must be noted that our methodology can also incorporate additional parameters as well. The variation of these QoS metrics impacts the quality of the delivered video and, consequently, the user satisfaction level. Our objective is to design and implement a method to estimate the variation of the user satisfaction level in function of the network QoS conditions. At first, we performed a set of subjective tests with real participants to measure the correlation between QoS metrics and QoE of video traffic. Second, we proposed a fuzzy expert system which is based on No Reference method that can estimate the video quality based on the network conditions. Figure 16 illustrates the multimedia QoE estimation framework based on network QoS parameters. The correlation between the QoS metrics and the participants' QoE is transformed into fuzzy membership functions using probability distribution functions and curve fitting methods. We also proposed a simple methodology for fuzzy inference rules generation by assigning weights to the video impairment scores. Three different sets of experiments were performed to evaluate our system. In the first one, we simulated our system in MATLAB and compared the estimated QoE output (also called estimated MOS or eMOS) with the subjective QoE obtained from the participants in a controlled test. This experiment validated our methodology showing high correlation between our estimated and the subjective QoE. In the second experiment, we integrated our system as part of a monitoring tool in an industrial IPTV test bed and compared its output with standard Video Quality Monitoring (VQM). The outputs of both video quality estimation methods were also highly correlated. In the third experiment, we performed benchmarking analysis between our monitoring tool and Indra's monitoring tool [162], the experimental results showed that the scores predicted by both the tools were highly correlated. The experimental results show that the proposed video quality estimation method based on fuzzy expert system can effectively measure the network impact on the QoE. In the following sections, we describe our methodology and the experimental results.

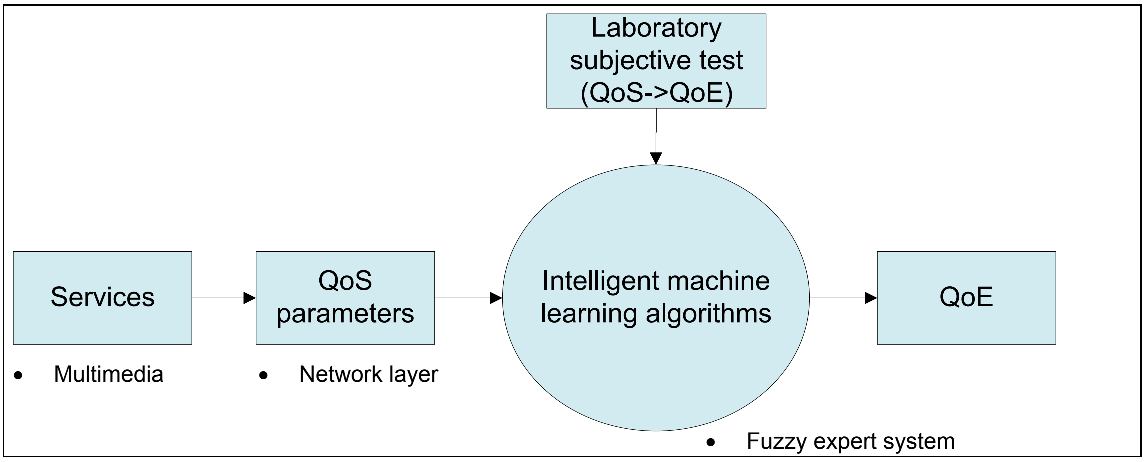

Figure 16: Multimedia QoE estimation Framework (Based on network QoS parameters

4.3 Methodology

In order to develop our video quality estimation technique, we followed a methodology that consists on conducting subjective tests with end user participants in order to build a learning set for correlating objective network QoS metrics with the subjective QoE provided by the participants. This correlation was then used to build the membership functions and inference rules of our fuzzy expert system for video QoE estimation. Figure 17 illustrates the methodology for the proposed multimedia QoE estimation system.

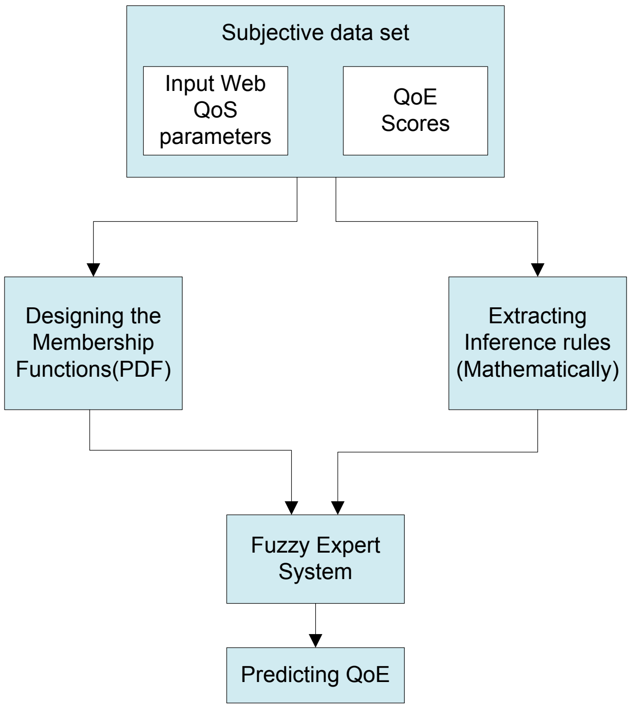

Figure 17: Methodology of proposed video QoE estimation system

4.3.1 Subjective Tests

In the subjective tests, we presented different video clips to 25 participants who rated each video clip according to the perceived impairment giving one of the MOS scores as described in section 2.2.3.We used six video clips of different types (sports, movie, animation, and interview) and generated 228 sample video clips which were constructed with different network level perturbations. The characteristics of these videos are listed in Table 8:

Table 8: Video characteristics

Video Content	Resolution	Frame rate (frame per second)
Movie1	1280x720	47
Animation1	1280x720	50
Interview1	1280x720	50
Sport	1280x720	50
Animation2	1980x818	48
Interview2	1280x720	60

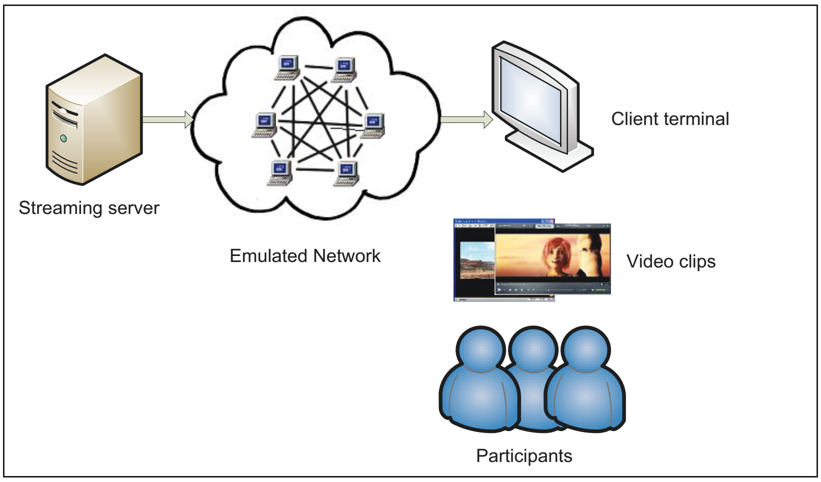

Figure 18: Experimental setup for subjective tests.

We constructed these video clips by streaming from a server to a client and correspondingly introducing perturbation through an emulated network. We selected three QoS parameters for perturbation; packet loss, jitter, and packet loss burstiness, which we considered promising for the mapping of QoS to QoE for video traffic. These constructed video clips along with the original video clips were shown to different participants in a random order in a closed room. For each participant it took around 2.5 hours to perform the test. After watching 20 video clips a pause of 5 minutes was taken. Figure 18 illustrates the experimental environment for the subjective tests.

4.3.2 Correlation between QoS Metrics and QoE

From the subjective test, we built a learning set that consisted of the mapping between the participants' scores and the QoS metrics for each of the considered video clips. We used a probabilistic approach to correlate QoS metrics to the participants' scores. Therefore, for every QoS metric, we built five different probability distribution functions (pdf) (one function per QoE score) that provide the variation of the participants' ratio (%) with the QoS metric for a specific QoE score. This probabilistic information was changed into a fuzzy set by dividing the pdf by its peak value (normalized pdf) [11]. The fuzzy set, which has the same form as that of the original pdf, is converted into an equivalent triangular or trapezoidal fuzzy set by using a curve fitting method [12]. The triangular or trapezoidal fuzzy set represents the membership functions for the different QoS metrics.

Figure 19 illustrates the QoE scores membership functions associated with the packet loss rate QoS metric. For example, the packet loss rate of 0.2% has membership values of 0, 0.6, 0.8, 1 and 0.35 corresponding respectively to the QoE scores 5, 4, 3, 2, and, 1. We note that a membership value of 1 represents a high degree of membership to the corresponding class and decreasing membership value represents deviation from the class. Figure 20 and Figure 21 illustrate the membership functions for packet loss burstiness and jitter metrics respectively. Similarly, in Figure 22, the membership functions for the estimated QoE are defined according to the standard MOS [21] definition.

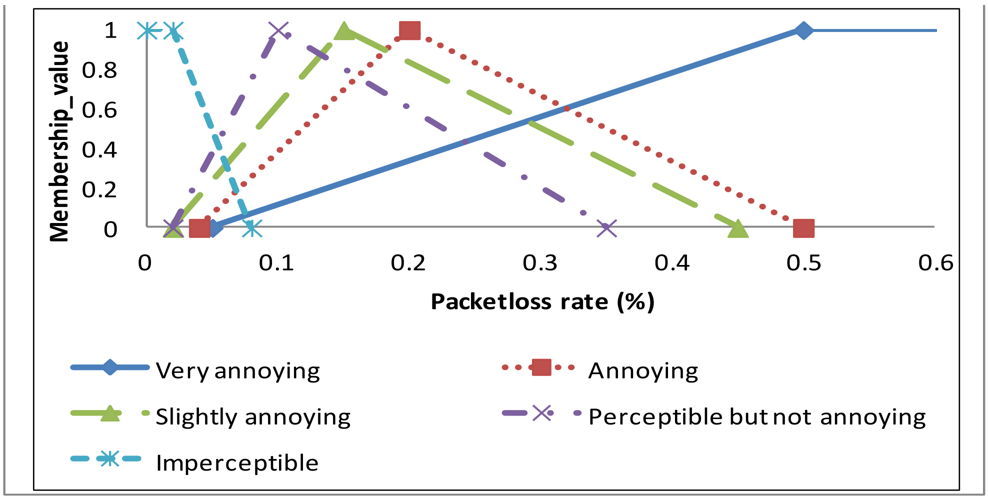

Figure 19: Membership functions for packet loss rate metric.

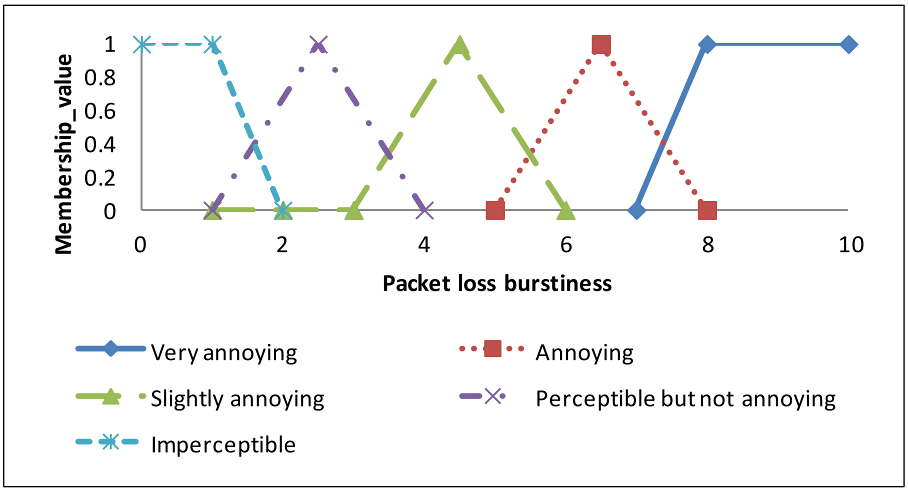

Figure 20: Membership functions for packet loss burstiness metric.

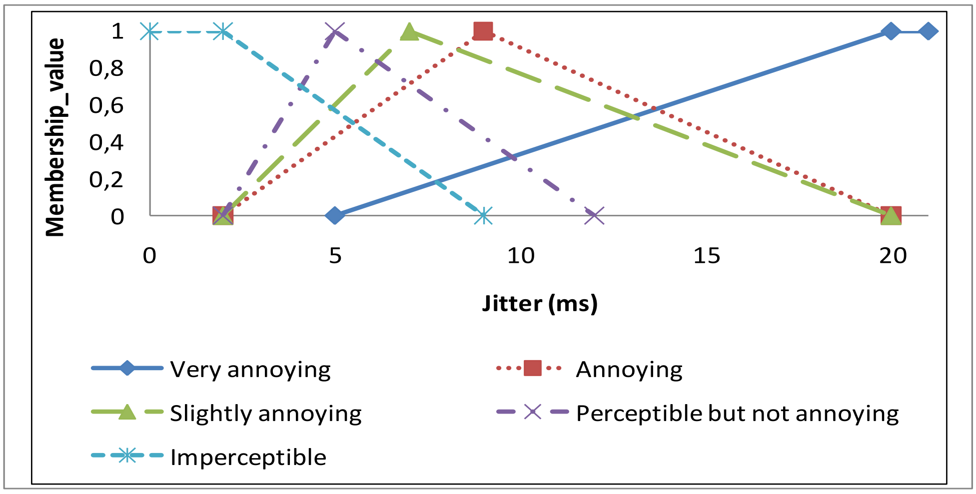

Figure 21: Membership functions for jitter metric.

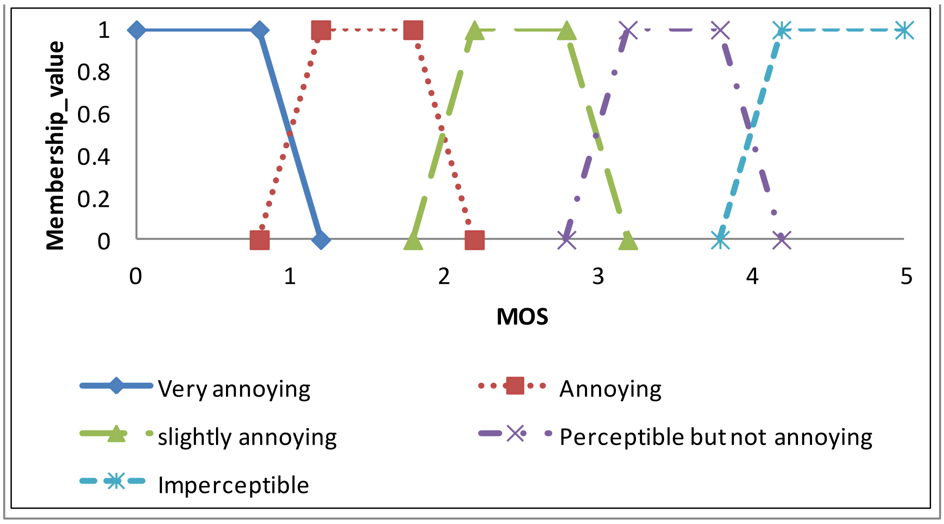

Figure 22: Membership function for the estimated MOS (eMOS)

4.3.3 Video Quality Estimation Fuzzy Inference Rules

The correlation between the QoS metrics and the video quality, described in the subsection 4.3.2, allowed to build five fuzzy membership functions for the three considered QoS metrics (packet loss rate, packet loss burstiness, and jitter). Based on the combinations of QoS metrics and their rating, we have to estimate whether the network impact on the video quality (QoE) is imperceptible (excellent conditions), perceptible but not annoying (good conditions), slightly annoying (fair conditions), annoying (poor conditions), or very annoying (bad conditions). That is, we need to associate an estimated QoE score for the different combination of QoS metrics scores. For example:

IF (Packet loss is very annoying) & (Burst Loss is very annoying) & (Jitter is very annoying)

THEN (the estimated QoE is very annoying).

Considering the combination of the QoS metrics scores, we have a set of 5^3 possible rules. We follow the following methodology to define the rules while at the same time reducing their number. We associate a "weight" to each rating as follows: 0 for "imperceptible", 1 for "perceptible but not annoying", 3 for "slightly annoying" 5 for "annoying", and, 7 for "very annoying". For every combination, we calculate the rule weight as the sum of the weights of the QoS metric scores. The rule output corresponds to the estimated QoE score as defined in Table 9. For instance, if all the QoS metrics scores are imperceptible then rule score is 0 (0+0+0), which corresponds to "imperceptible" QoE score. Likewise, if two QoS metrics scores are "perceptible but not annoying" and one is "slightly annoying", then the rule weight is 5 (1+1+3), which corresponds to an "annoying" QoE score.

Table 9: Rule weight to QoE mapping

Rule weight	Estimated QoE score
0	Imperceptible
1-2	Perceptible but not annoying
3-4	Slightly annoying
5-6	Annoying
7+	Very annoying

4.3.4 QoE Estimation System

Our proposed video QoE estimation system is based on fuzzy logic that is powered with a learned membership function (QoS/QoE correlation) and a set of fuzzy inference rules. Fuzzy logic is a well-known technique that can handle problems with imprecise and incomplete data [84]. Figure 23 illustrates the building blocks of our proposed system that can be placed at any point in the network between the video source and the terminal. The system performs a per-flow analysis. A Deep Packet Inspection (DPI) engine at the entry point of the system inspects the network traffic to identify video flows and extract relevant per-flow QoS metrics (packet loss rate, packet loss burstiness, and jitter). These metrics are constantly fed to the fuzzy expert system that uses the defined membership function and inference rules to estimate the QoE on a per flow basis.

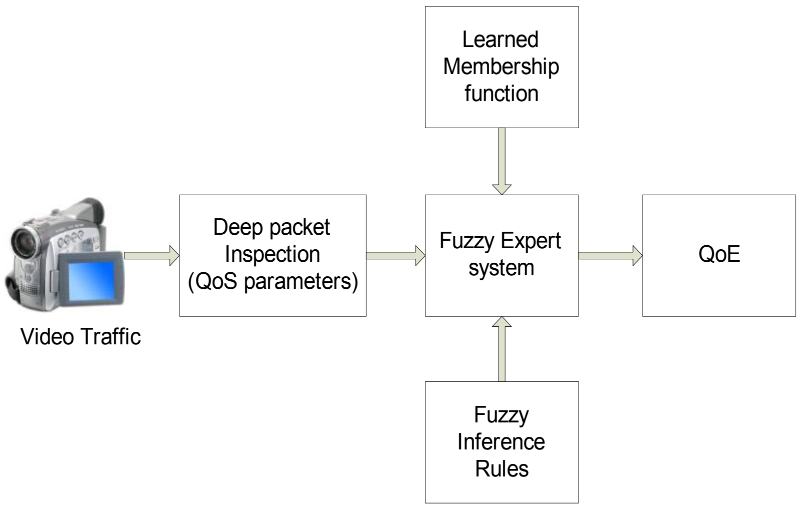

Figure 23: Building blocks of the video QoE estimation system

4.4 Validation and Experimental Results

4.4.1 Validation of the Proposed Methodology

To validate the proposed methodology, we compared the results obtained from the subjective tests in section 4.3.1 with those obtained from our proposed system. For this end, we used the Fuzzy logic toolbox of MATLAB [85] and developed a simulation scenario with our membership functions and rules for validation. Video clips with different QoS metric values were used for validation. For each video clip, we obtained subjective QoE from subjective tests as described in section 4.3.1; and, estimated QoE from our system simulated in MATLAB. Figure 24 represents the comparison between the subjective and estimated MOS. We can see that subjective MOS and estimated MOS have a linear relationship and are highly correlated with correlation coefficient of 0.95. This indicates that the proposed system succeeds in reflecting user's perception. This is also illustrated in Figure 25 that considers the probability distribution of the difference between the participants' subjective scores (MOS) and the estimated scores. We can see that in around 60% of the tests, the score difference were less than 0.5. It reached 83% for score difference less than 1. This means that in 83% of the cases, the difference in the subjective and estimated QoE were at most one score level (e.g., if the participants reported "Imperceptible" video quality, our method would have reported either "imperceptible" or "perceptible but not annoying"). Only 4% of the tests showed a score difference of more than 1.5. This estimation accuracy emphasizes the ability of the proposed system to measure the impact of the network conditions on the user satisfaction.

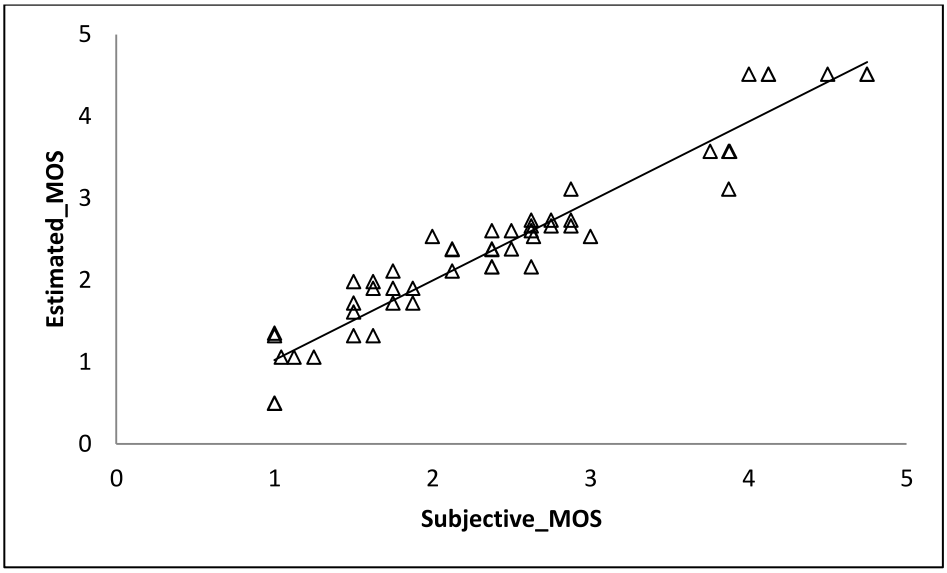

Figure 24: Comparison between the subjective and estimated MOS

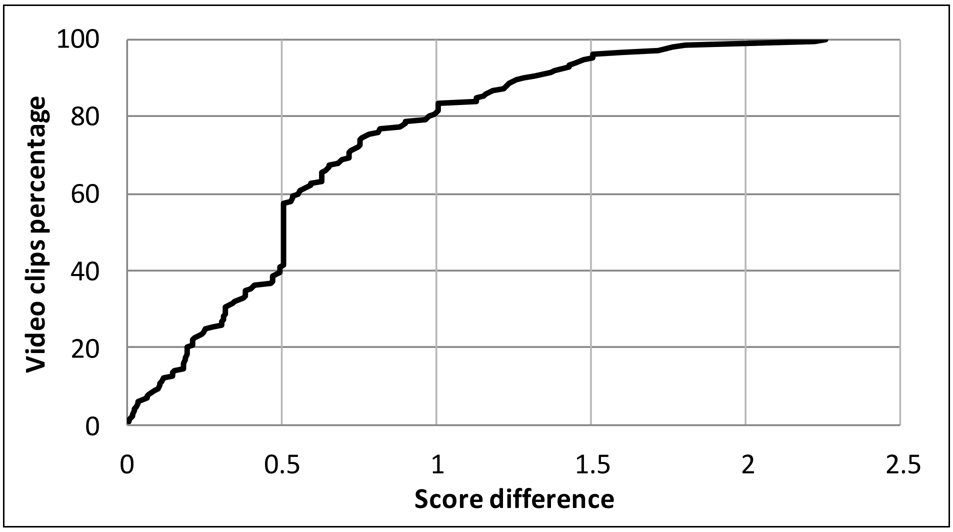

Figure 25: Probability distribution of the subjective and estimated MOS difference

4.5 Real Test Bed Experimentations

To evaluate our system in close to real network conditions, we implemented the proposed fuzzy expert system for video quality estimation as a module in a MMT [86] probe. Then we integrated MMT probe and a VQM probe in an industrial IPTV test bed [87]. The experiments consisted in comparing the results of both probes in the presence of different emulated IP level and copper line perturbations.

In the following sections, we first describe the industrial evaluation test bed; then we present an analysis of the experimental results.

4.5.1 Test Bed for QoE Measurement

Figure 26 presents the building blocks of the Vierling experimental test bed, which is an industrial evaluation test bed for QoE assessment in Digital Subscriber Line technology (xDSL).

The test bed is constituted of an IPTV server that uses VLC to stream video clips over RTP/UDP transport. The streaming server is connected to a Digital Subscriber Line Access Multiplexer (DSLAM) modem emulator at the central office (CO). At this point of the network, an IP traffic perturbation tool is used to emulate network conditions by introducing a variety of network impairments such as packet loss, delay, and jitter. These perturbations are used for quantifying the network impact on the video QoE. The DSLAM modem is connected via kilometers length copper lines to the Customer Premises Equipment (CPE) xDSL modem. Furthermore, the test bed allows performing measurements on multiple copper line configurations at the network access level and allows adding xDSL disturbers with the help of remote controlled HF signal and noise generator. The test bed can reproduce a typical DSL network configuration: (i) the IPTV server represents the service provider network, and (ii) the DSLAM modem emulator and CPE (Customer Premises Equipment) modem correspond to the access network on a physical copper line with thousands of copper configurations.

In addition, external network probes can be attached to the modem at the CPE side to sniff the video traffic and analyze the collected data in order to correlate QoS parameters and estimate the corresponding QoE level.

We implemented our QoE estimation system as a module in the MMT to facilitate its deployment in IP based networks. Now, the MMT method refers to our QoE estimation method. This monitoring tool uses DPI to detect video flows and extract the QoS metrics of interest to be used for the QoE estimation.

At the CPE side, we installed the MMT probe along with a VQM probe in order to compare the results of the two methods. The QoE measurements consisted in analyzing the network traffic of the broadcasted video clips. Two different types of deterioration factors were simulated: xDSL noise and copper faults on the copper cables and, packet loss and message delay/jitter on the IP level on the DSLAM modem.

We compare the results of the MMT method with those obtained using the VQM method.

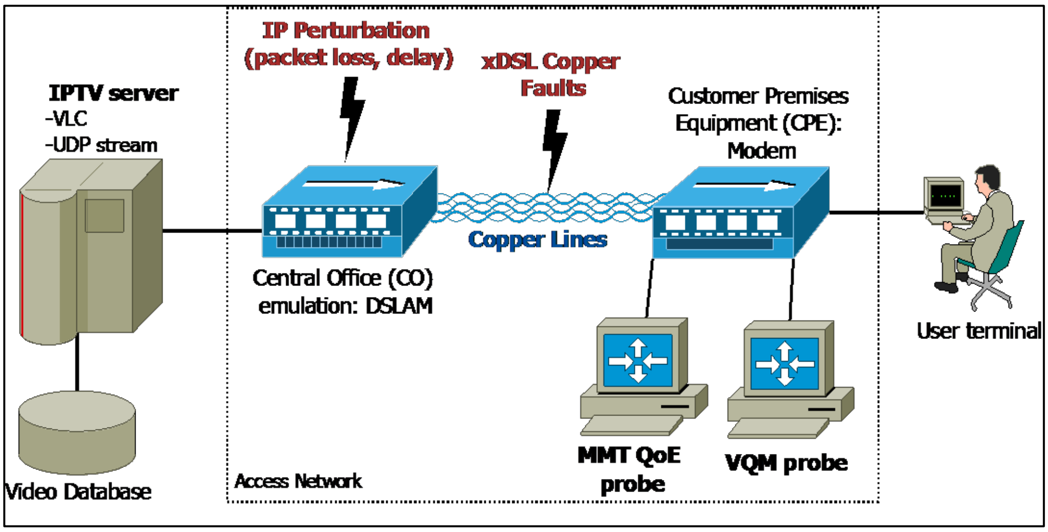

Figure 26: Vierling experimental test bed for QoE measurements

4.5.2 Experimental Results

A total number of around 600 video streams were analyzed. For every stream, we collected the applied perturbation type, the QoS parameters, and the estimated MOS given by the MMT method and by the VQM method. Figure 27 illustrates the variation with the packet loss rate of the estimated Mean Opinion Score (eMOS) given by MMT and by VQM. For both methods, we can see that the eMOS values decrease with the increase of packet loss. The decrease, however, is sharper using MMT method. This is due to the fact that MMT estimation model is mainly designed for high definition video quality. In fact, the subjective test described in section 4.3.1 used exclusively HD video clips. In the test bed experiments, average quality videos were used. This fact made the transmitted video clips less sensitive to packet loss than HD videos. As the packet loss rate increases, the eMOS of both methods decreases to reach a point where the estimated quality becomes very bad. We should note here that MMT eMOS has a maximum value of 4.51 and a minimum value of 0.494. This is due to the centroid method used for defuzzification in the fuzzy expert system [84].

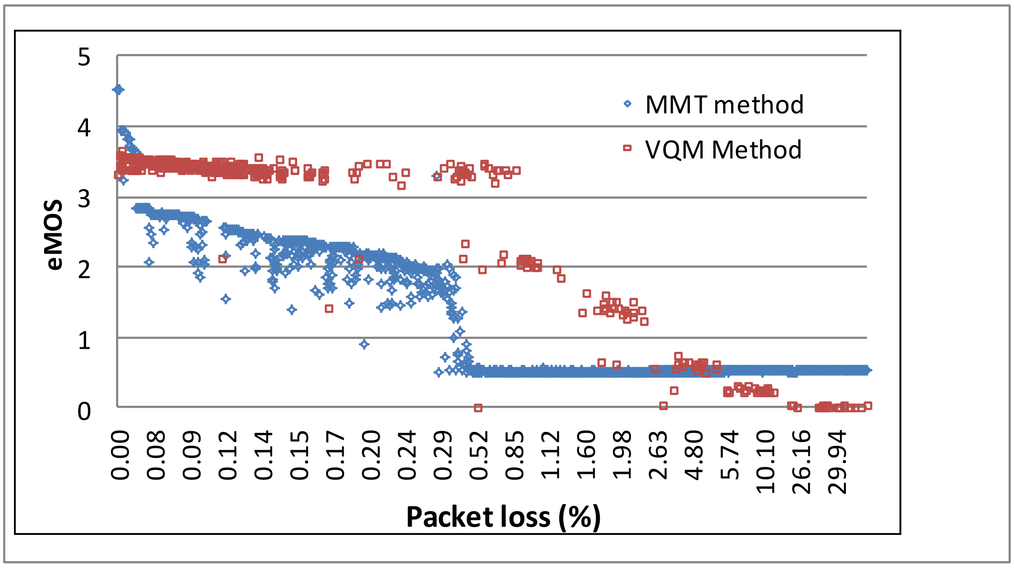

Figure 27: Variation of eMOS with packet loss rate

Figure 28 compares the results of MMT method and VQM method for different copper line perturbations (noise of different amplitudes). We can see that the results of both methods are highly correlated.

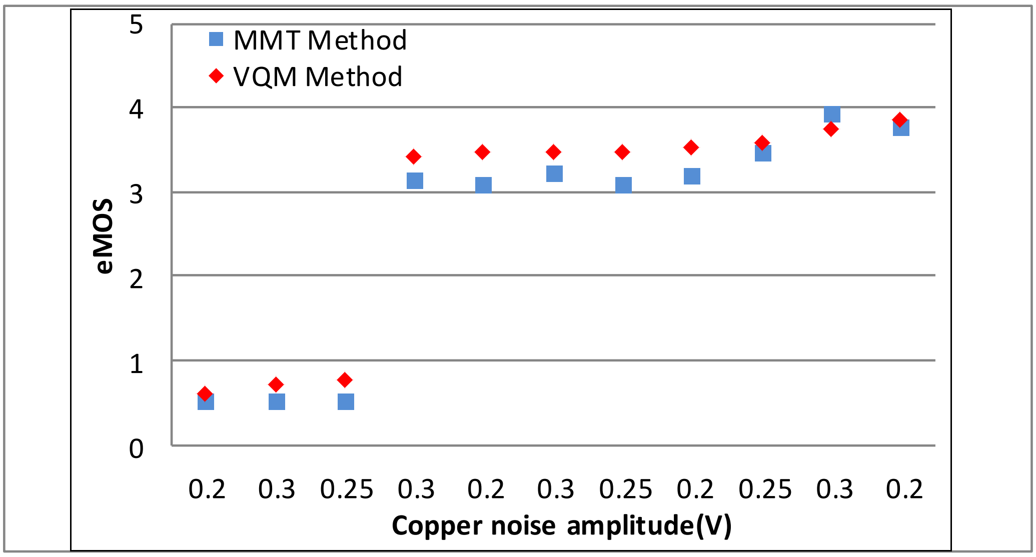

Figure 28: Estimated MOS in different copper line perturbations

Furthermore, the VQM method is calculated based on media signal parameters therefore, it is expected to yield higher accuracy. On the other hand, the MMT method is based on network parameters and not all network impairments result necessarily in a visible degradation [77] [78] [79]. However, the results of the test bed evaluation show that MMT proposed video quality estimation succeeds in reflecting the impact of network perturbations on the video quality with reasonable accuracy.

4.6 Benchmarking

For benchmarking analysis, we compared the result obtained from our MMT probe with the Indra's probe [162] for QoE estimation in the real IPTV test bed. Video traces from different network monitoring probe with different level of network impairments were used for this purpose. After that, for different level of network impairment we compared the estimated MOS obtained from MMT probe and Indra's probe. In Figure 29 we can see that both probes behave in similar way and their MOS scores are highly correlated with correlation index of 0.86.

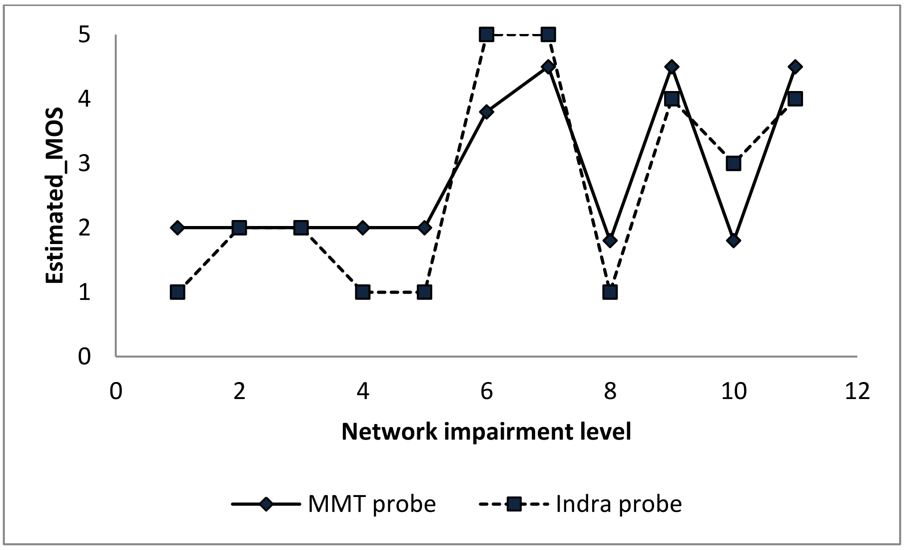

Figure 29: Benchmarking MMT probe with Indra's probe

4.7 Conclusion

In this chapter, we have analyzed the effect of different network QoS parameters on video QoE. From the subjective test, we have found that video QoE is very sensitive to network perturbations and different values of QoS parameters (due to network perturbation) correspond to different levels of video impairments with different QoE values (MOS scores). Therefore it is very hard to accurately estimate video QoE based on observations or simple mathematics as the relationship between QoS and QoE is fuzzy. To address this problem we have proposed a video quality estimation method based on a fuzzy expert system to measure the impact of network condition (QoS parameters) on the user perceived satisfaction level (QoE) of video services. We have performed a set of subjective tests with real participants in order to correlate network QoS parameters levels with the user perceived video quality. The defined fuzzy membership functions were derived from the QoS/QoE correlation using probability distribution functions. We have also proposed a simple method to build estimation inference rules by assigning weights to the different video impairment scores. The proposed methodology has been validated versus the results of subjective tests. The proposed system has been integrated and tested in an industrial IPTV evaluation test bed. Moreover, we performed a benchmarking analysis between our QoE estimation system and Indra's QoE estimation system. The validation and experimental results have shown that our QoE estimation method is correlated to both the participants' subjective QoE scores as well as to the estimated MOS given by the VQM method and Indra's method. However, to increase the accuracy in estimation/prediction it is necessary to incorporate other parameters like codec parameters, MAC level parameters, users profile, etc. as the network parameters alone cannot precisely evaluate the QoE at the user end.

Chapter 5

QoE estimation of VoD services: MAC Layer Perspective

5.1 Introduction

The use of IEEE 802.11 based WLANs (Wireless Local Area Networks) is immense. The reason for this paradigm shift is the tremendous increase of Wi-Fi enabled gadgets, smart phones and tablets in our daily lifestyle. There is an equal amount of increase in wireless multimedia services like Voice over IP (VOIP), IPTV, telemedicine and internet gaming for both home and professional use. The reason of becoming one of the preferred access technology is also fuelled because of its easy deployment, availability and cost effectiveness. Moreover, the advancement of this technology has attracted the operators to use it as a supplement to enhance their service and business coverage.

Although there is a huge advancement in wireless technologies, the issues related to QoS provisioning still remain challenging. The legacy IEEE 802.11 a/b/g are capable of supporting bandwidth intensive applications such as audio/video streaming and interactive gaming, but they cannot guarantee QoS when traffic load increases.

Recently, in order to meet multimedia service requirements, the IEEE 802.11 working groups have specified the IEEE 802.11n standard with many new promising enhancements such as MIMO (Multiple Input Multiple Output), OFDM (Orthogonal Frequency Division Multiplexing), frame aggregation and Block Acknowledgement (BA). The high data rate that 802.11n allows has enabled the adoption of demanding applications like high definition TV, VOD (Video on Demand), video conferencing, etc. It is interesting to investigate the performance of these multimedia services over such wireless networks. Furthermore, when it comes to wireless networks, MAC (Medium Access Control) protocols play a crucial role in the performance of the whole system as it governs how resources (channel) get allocated to different stations.

For the quality estimation of video traffic in wireless networks, the ability to identify the perceived degree of video impairment with respect to MAC-level parameters is important.

Moreover, the impact of MAC-level parameters on video can be significant and may completely degrade the video quality. Therefore, careful parameterization of these parameters can increase overall performance of video transmission. Many research were conducted regarding quality estimation of video traffic based on network and upper level parameters, however, very few research has been done from the MAC layer perspective.

Generally, QoS parameters are used to determine the service quality, which do not necessarily reflect the user's satisfaction towards a particular service. QoS parameters reflect network and service level performance; however, they do not address user's reaction to the service or application. Therefore, it is necessary to estimate QoE to indicate service quality.

In this chapter, we focus on mapping MAC-level parameters to user perceived QoE in order to provide tools to measure the impact of the MAC-level parameters on the video QoE. In order to achieve this objective, intensive video subjective test with a number of participants were performed. In the video subjective tests, participants were asked to provide MOS [21] (Mean Opinion Score) values to video clips with impairments (varied MAC-level parameters). These MOS scores reflect the user satisfaction towards video quality. As a first contribution, we evaluate the performance of video transmission over 802.11n from MAC level perspective. Furthermore as a second contribution, we propose a QoE estimation system based on Random Neural Networks (RNN) [42] to estimate the impact of some MAC-level parameters on video quality.

5.2 **Related works**

The IEEE 802.11n standard is a relatively new WLAN technology that can operate in both 2.4 GHz and 5 GHz band. It brings multiple technological enhancements to improve speed, reliability, and range. The increase in data rate is about 11 times as compared to the former IEEE 802.11g technology (i.e., 600 Mbps versus 54 Mbps). MIMO (Multiple Input Multiple Output)/SDM (Spatial Division Multiplexing) is a key feature introduced at the PHY layer. Additional technologies like OFDM, 40 MHz channel width operation, and short guard intervals have also contributed to the increase in physical data rate [88]. At the MAC layer, enhancements are made to the earlier IEEE 802.11e MAC, like frame aggregation and Block Acknowledgement (BA) to reduce the overhead, increase the useful throughput, and improve the overall efficiency.

With frame aggregation multiple data units or sub frames are concatenated with one leading physical header. The objective is to reduce the inter-frame space and the preamble all together into to a single radio preamble. This reduces the overhead and resources wastage as compared to the traditional single frame transmission. There are different aggregation types, mainly A-MPDU (Message Protocol Data Unit aggregation), A-MSDU (MAC Service Data Units aggregation), and two-level aggregation. In A-MSDU multiple MSDUs are aggregated and transmitted at once as a single MPDU. In A-MPDU multiple MPDU frames are aggregated into larger A-MPDU frames with only one physical header. The difference in A-MPDU frame aggregation is that aggregation is done after the MAC layer sends the MPDU frame [88]. In this work, we consider A-MPDU as our aggregation scheme as it performs better for videos [89]. The BA mechanism was introduced in IEEE 802.11e and has been modified in IEEE 802.11n to support the A-MPDU aggregation. With a single BA frame, multiple received frames can be acknowledged instead of a single ACK frame for individual data frames. When an A-MPDU with errors is detected, the receiver transmits a BA to selectively acknowledge the uncorrupted MPDUs. The sender is then able to re-transmit only the non-acknowledged MPDUs. This mechanism can only be used with the A-MPDU aggregation mode and is very efficient in error prone environments [90].

Most of the work in the area of video QoE estimation is from the network layer perspective. They measure the impact of network parameters on quality of the video traffic, i.e., QoE [7] [74] [75] [76]. However, few efforts have been done to analyze the effect of MAC-level parameters on video QoE. Moreover most of the existing works analyses the video QoE based on individual MAC-level parameters [91, 92, 93, 94, and 8].

In [91], authors have evaluated the performance of VOIP services in IEEE 802.11n WLAN environments. They considered IEEE 802.11n specific parameters and Bit Error rate (BER) for analyzing audio quality. However, we are using video traffic and a different set of MAC parameters. Authors in [93] presented the impact of IEEE 802.11n frame aggregation mechanisms on video streaming. They used PSNR (Peak Signal-to-Noise Ratio) to represent the quality of video. Authors in [94] analyzed the impact of background traffic on video quality in IEEE 802.11b WLAN environment. In their work, video quality assessment was performed objectively with PSNR, Video Quality Metric (VQM), and Structural SIMilarity (SSIM). However we differ from [93] and [94] since we used subjective evaluation as compared to their objective approach. The work presented in [92] estimated QoE of audio-video services from MAC-level QoS based on IEEE 802.11e and used its associated parameters such as distance between the nodes and TXOP limit. Our work share some similarities with [92] but differs from them since we performed our experiments over IEEE 802.11n wireless network and considered different MAC-level parameters for evaluation.

Moreover, our work analyzes the combinational impact of various MAC-level parameters on video QoE in IEEE 802.11n wireless network.

5.3 Performance Evaluation (Effect of individual MAC layer parameters on video QoE)

In this section, we analyzed the performance of video transmission which is affected by different MAC level parameters. In order to achieve this objective, intensive video subjective tests with a number of participants were performed. In these video subjective tests, the participants were asked to provide MOS [21] (Mean Opinion Score) values as defined in section 2.2.3 to video clips with impairments (varied MAC-level parameters). These MOS scores reflect the user satisfaction towards video quality. Furthermore, we also analyzed how the perceived QoE varies for different video content types.

5.3.1 Potential MAC-layer parameters

In this thesis work, we have taken into account different MAC-level parameters (including some of the enhanced 802.11n MAC parameters) to evaluate their effects on video QoE. These parameters are either directly associated to MAC layer (e.g., queue size, aggregation, retransmission limit) or indirectly associated (e.g., number of competing station, BER). They all play a crucial role and can impact the quality of the video traffic. Each of these MAC parameters is described below:

5.3.1.1 Bit Error Rate

Wireless networks suffer due to the error prone wireless channel characteristics. This results in the variation of the Bit Error Rate (BER) that can cause the MAC frame to be received with errors and trigger retransmissions that can impact the overall performances of the system [91].

5.3.1.2 Aggregation Size

Frame aggregation, here A-MPDU (MAC Protocol Data Unit) aggregation, allows combining many MAC frames into one larger aggregated frame. Streaming applications can take benefit from frame aggregation and the mean aggregate size can increase with both resolution and the wireless channel conditions. Limiting the maximum frame aggregation length can severely impact both the average delay and the quality of a video stream. However in high BER, aggregation size should be limited to reduce larger frame drops [90].

5.3.1.3 Number of competing stations

The performance of the wireless network degrades with increasing number of users. Indeed, as the number of competing stations increases in the system, stations have to struggle most of the time to find a free channel (due to high contention).This may lead to higher collision rates, larger delays, and packets dropping .

5.3.1.4 Queue length

Many applications are sensitive to queue characteristics like queue length, waiting time, service time, etc. A large queue can hold more packets to avoid higher packet drop. However, this needs a larger memory which is not feasible in already deployed physical implementations and may also result in outdated information. On the other hand small queue length can increase packet drop, which is more likely to happen in saturated conditions.

5.3.1.5 Retransmission Limit (Maximum number of retransmissions)

Retransmission is a good policy to deal with corrupted packets at the MAC layer. However, there are delay bounds for each traffic type and when packets arrive after a certain deadline, they become useless. When channel conditions are not highly error prone, using different values for maximum retransmission limits in different channel conditions can enhance the system performance.

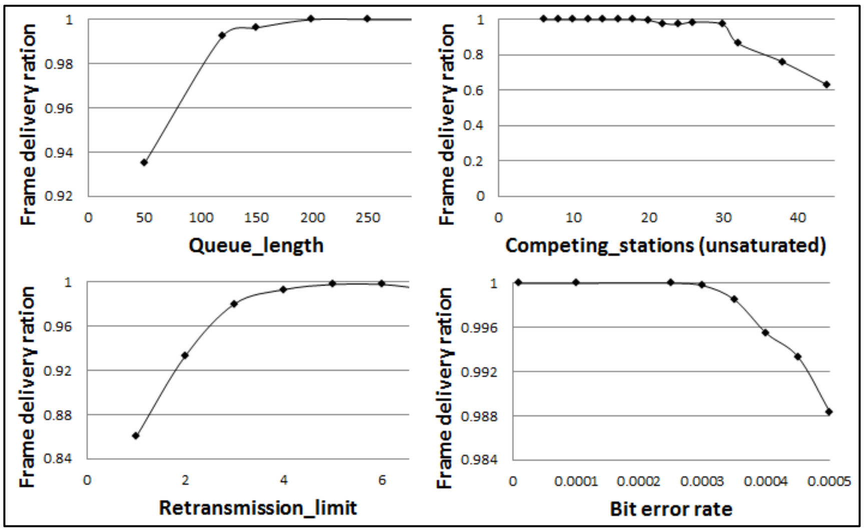

Figure 30: Frame delivery ratio vs different MAC-level parameters

Figure 30 shows the effect of these MAC-level parameters on video frame delivery ration. In this figure, we presented the case for "movie" content type, however similar affect was observed for other content types as well. We can see that at different values of these MAC-level parameters, the value of frame delivery ration changes. Since, video quality is extremely sensitive to packet drops/ frame drops, a decrease in frame delivery ratio consequently degrades the video QoE. This justifies the importance of the chosen MAC-level parameters

5.3.2 Simulation Platform

For simulations, we use NS2 simulator embedded with IEEE 802.11n module. We considered a network topology consisting of static stations where each station transmits traffic to its paired station. Two types of traffic flow i.e., video and data are used in the simulation. The video flow consisted of H.264/AVC video traces that were fed as input using the Evalvid tool [95] and the data flow consisted of CBR (Constant Bit Rate) traffic. The simulation was conducted for three different types of video content, as shown in Table 10.

Table 10: Video characteristics

Type	Bit rates (kbps)	Resolution	Frame per second	Time (Sec)
Movie	2549	640 x 360	24	15
Animation	2508	640 x 360	24	15
Sports	2686	640 x 360	24	15

During our simulation, we have designed different scenarios to examine the performance of video under different MAC-level parameters. Some of the simulation parameters that are common to all scenarios are presented in
Table 11 along with their respective values.

In each scenario, each video clip was streamed from a source node to a destination node and correspondingly different MAC-level parameters were varied. This was repeated for all three video content types. This combination resulted in a total of 200 video clips with different perturbations.

Thereafter, for the subjective test the constructed video clips along with the original video clips were shown to different participants in a closed room, in a random order. Participants rated each video clip according to their perceived impairment giving one of the following MOS scores. On average each participant rated 100 video clips. After watching 20 video clips a pause of 5 minutes was allowed. Figure 31 illustrates the experimental environment for the subjective tests. A total of 40 users registered for the test, which is considered reasonable for this kind of subjective tests [96].

Table 11: Simulation parameters

Parameters	Values
Simulation Area	250 m X 250 m
Simulation time	30 s
Node mobility	Static
Bandwidth	96 Mbps
MAC layer	IEEE 802.11n
Video packet size	1052 bytes
CBR packet size	1000 bytes
Time slot	20 s
SIFS	10 s
TXOP limit	3.264 ms
Antenna Type	Omni Directional
Network Interface Type	Wireless/Physical/MIMO
Interface Queue Type	Aggregation queue
Number of antennas	4
Transmission protocol	UDP

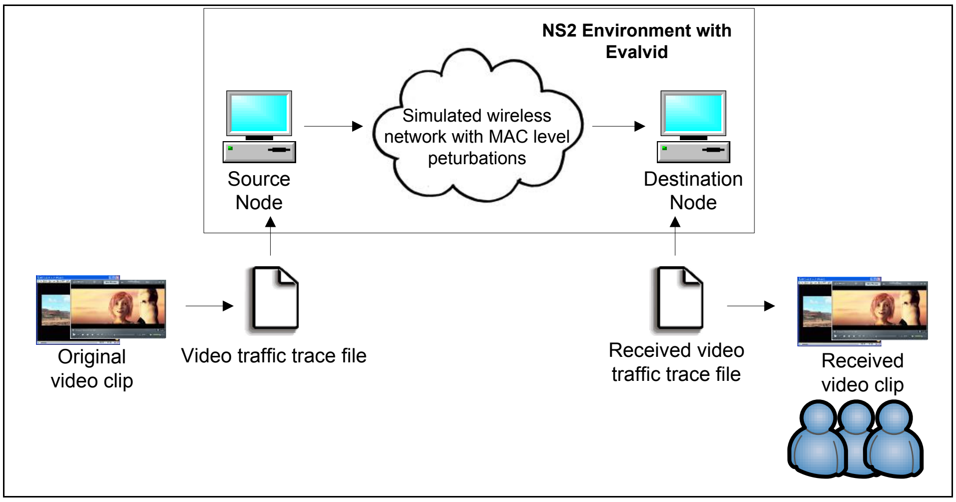

Figure 31: Subjective test Environment

5.3.3 Experimental Results

As a first set of experiments, we analyzed the effect of individual MAC level parameters (aggregation, BER, number of competing stations, and load) on the MAC QoS parameters.

The MAC QoS parameters considered were MPDU drops, collision rate, and jitter. We varied each individual MAC-level parameter and calculated the three QoS values for the respective cases.

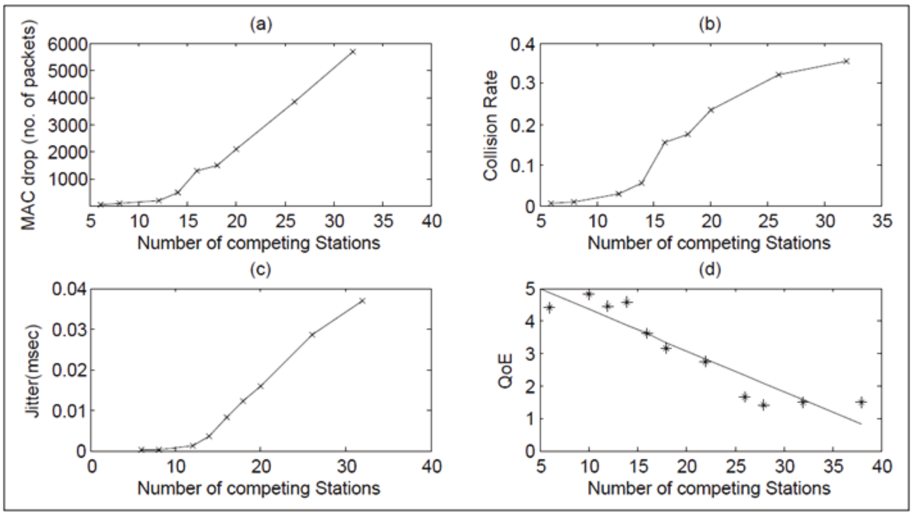

Figure 32: QoS and QoE versus number of competing stations

Figure 32 shows one of the cases with varying number of competing stations. We can see that as the number of competing stations increases there is an increase in MPDU drop, collision rate, and jitter. This signifies the effect of the increasing number of competing stations on the MAC QoS parameters. We can also see in Figure 32(d), the QoE values that were subjectively obtained for the same variation of the number of competing stations. We can observe that there is a similar degradation in the video quality (QoE) suggesting that the degradation in MAC layer QoS has a subsequent effect over video QoE. We conducted a similar type of simulations varying the other stated MAC-level parameters and each of them followed a similar behavior in QoS and QoE relation.

For the second set of experiments, we analyzed the individual effect of different MAC level parameters. As mentioned earlier, IEEE 802.11n introduces two main new MAC parameters that are frame aggregation and BA. BA is designed in order to combat error prone environments when large aggregated frames are transmitted. From our initial set of experimental results, Figure 33 and Figure 34 show the behavior of aggregation in the absence and presence of BA respectively. In Figure 33, we can see that in the absence of BA the performance of the system degrades drastically even at a slight increase of BER. This is not the case with BA as in Figure 34 where it tries to combat the effect of BER to a certain extent (2.5×10^{-4}) .We can also observe that when BA is not used, the performance degradation increases with the value of aggregation and this is shown by lower MOS values. This is due to the loss of entire A-MPDU frame which is not the case with BA where only the corrupted sub-frame is lost. For the rest of the simulation we consider BA as enabled. Moreover, from Figure 34 we can see that for the case with BA, different levels of aggregation value follow similar trend of performance degradation with an increase in BER. This is because video data rate is not high enough to aggregate sufficient frames for higher aggregation values. However, to get the optimal result we used the highest values of aggregation (64000 bytes) for the rest of our performance evaluation.

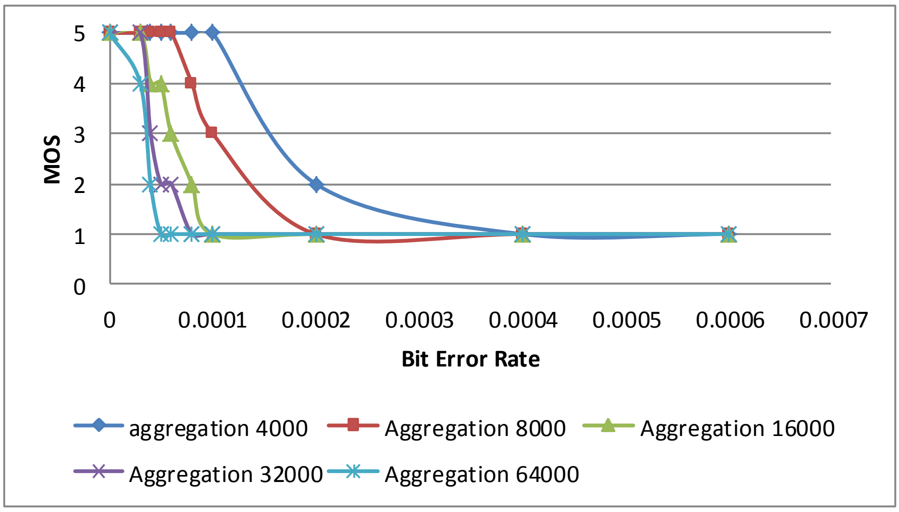

Figure 33: Aggregation without Block Ack

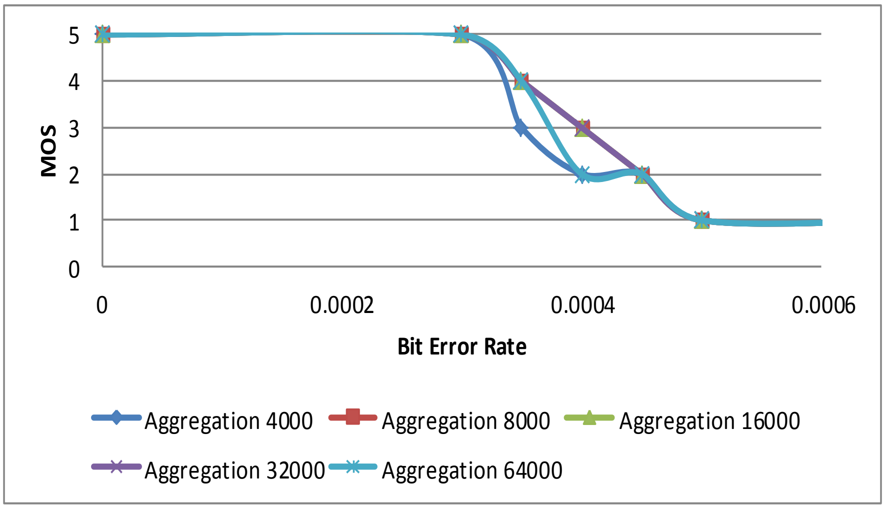

Figure 34: Aggregation with Block Ack

In the following results we showed how different MAC-level parameters impact user perceived video QoE for different video contents. Each point in the figures below represents the average MOS score given by the participants for different video clips and the line graph represents the trend these scores follow. It was observed that the impact is more on sport and animation contents than on movie content. This is because video contents chosen for sport and animation had more objects movements as compared to that in movie. Therefore, video service vendors/operator should be more cautious while providing QoS and QoE to different type of video contents.

5.3.3.1 Impact of Bit Error Rate on video QoE

Figure 35 shows the impact of BER on the video QoE. As BER increases, MOS value decreases exponentially. This is because an increase in BER will increase the number of corrupted video frames due to errors in transmission, which finally leads to the dropping of the video frame. From the subjective test results, BER value less than $1x10^{-4}$ has almost no effect on video transmission with MOS score of 5. However, if the BER value increases above $2x10^{-4}$ video transmission error increases causing degradation in video quality.

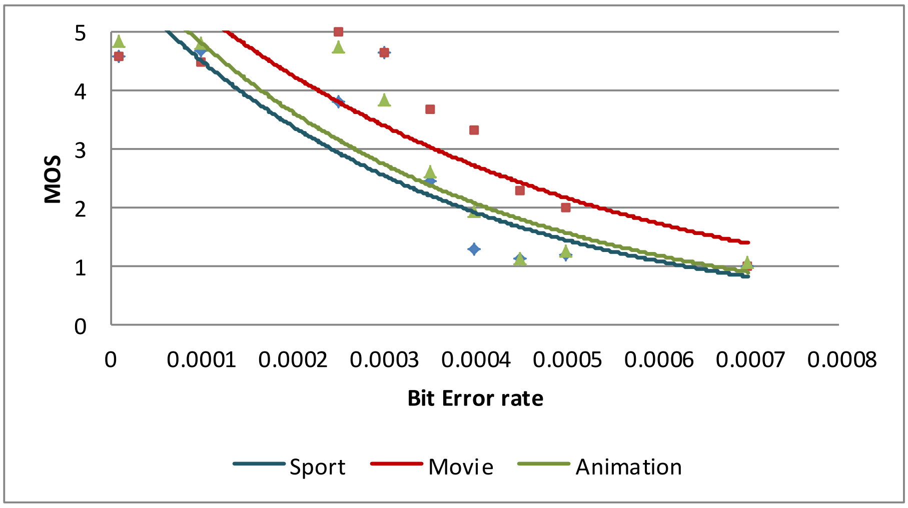

Figure 35: Impact of BER on video QoE

5.3.3.2 Impact of number of competing stations on video QoE

As the number of competing stations in the system increases, stations need to compete more to find the channel free. This situation can be worse when the total traffic in the system gets saturated. So it is necessary to analyze the impact of the number of competing stations over video QoE in both saturated and unsaturated conditions.

Impact of number of competing stations in unsaturated condition

In this scenario, overall channel capacity is unsaturated. To analyze the impact of the number of competing stations on video QoE, the number of stations with CBR traffic flows in the system is increased. When the total traffic rate in the system is low, packet loss will be less due to less congestion. However, as the number of competing stations increases, the collision probability increases due to limitations in the access mechanism and as a result, frame drop probability increases. Figure 36 shows that as the number of competing stations increases, the user perceived video QoE decreases. The result of subjective tests shows that video transmission is not affected much if the competing station in the system is less than 10. However, as the competing stations in the system increase above 10, video quality degrades which is shown by the exponentially decreasing curves of the MOS values.

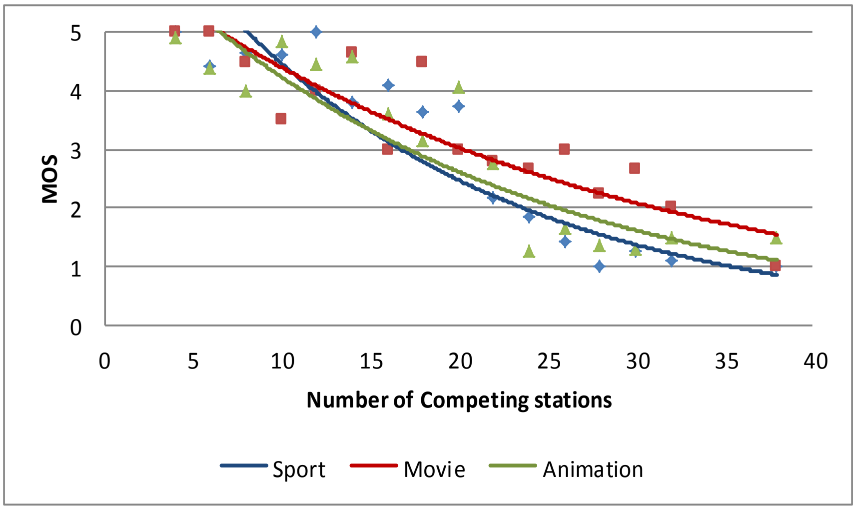

Figure 36: Impact of number of competing stations on video QoE in unsaturated conditions.

Impact of competing stations in saturation condition

In this scenario, the overall channel capacity is saturated. To analyze the impact of the number of competing stations on video QoE, the number of stations with CBR traffic flows is increased. As the total traffic rate in the system is very high, the increase in competing stations increases both congestion and collision probability. Figure 37 shows that the increase in the number of competing stations decreases the user perceived video QoE. From the subjective tests results in saturated conditions, it can be seen that when the competing stations in the system is less than 3 there is no effect on video transmission. However, as the competing stations increases above 3, the video quality starts degrading which is shown by the exponentially decreasing curves of MOS values.

From the above results we can conclude that the effect of increasing the competing stations is more when the total traffic is saturated in the system. For the saturated case, the video QoE drops even when the number of competing station in the system is 3 as compared to 10 when the traffic is unsaturated.

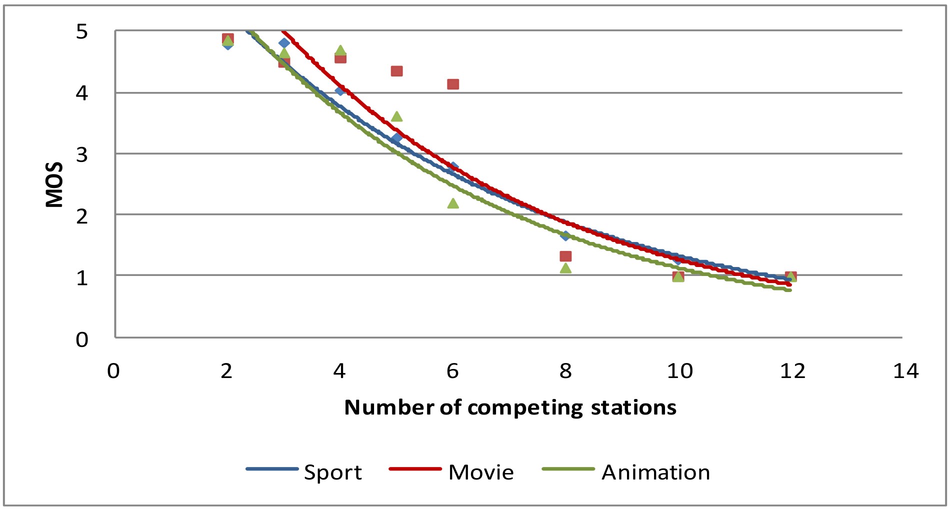

Figure 37: Impact of number of competing stations on video QoE in saturation condition

5.3.3.3 Impact of Queue length on video QoE

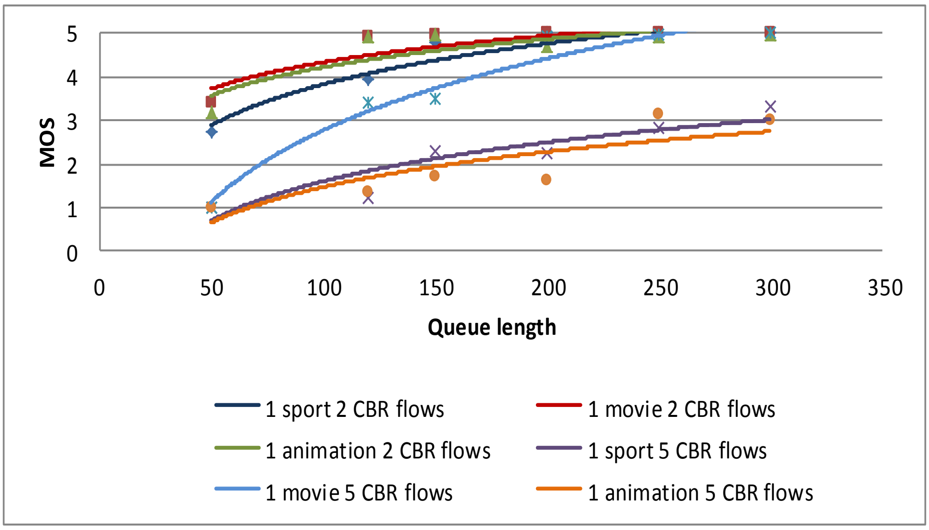

Figure 38: Impact of queue length on video QoE

Figure 38 shows the impact of queue length on the video QoE. As the queue associated with the MAC layer gets full, frame dropping probability increases. The MAC interface queue gets full when packet arrival from upper layer is higher than the transmitted rate. It gets worse when the number of competing stations increases because the station has to struggle more to find a free channel. This results in queue overflow and hence leads to frame drops. Therefore, we analyzed the effect of different queue lengths on video QoE with different number of competing stations in saturated traffic conditions. We can observe that, as the queue length decreases, the video QoE decreases correspondingly. The effect of queue length is higher when the number of competing stations is high i.e., 5 CBR flows. For example, for the sport clip with queue length 200 and 2 competing stations, the MOS value is almost 5; however, for the same queue length with 5 competing stations MOS value is 2.5.

Therefore, to achieve higher value of QoE, sufficient queue length should be assigned to the stations. However, we cannot assign higher value of queue since it requires larger memory which is not feasible in physical implementations and may also result in outdated information if packets are stored in the queue for a longer time.

5.3.3.4 Impact of Retransmission Limit on video QoE

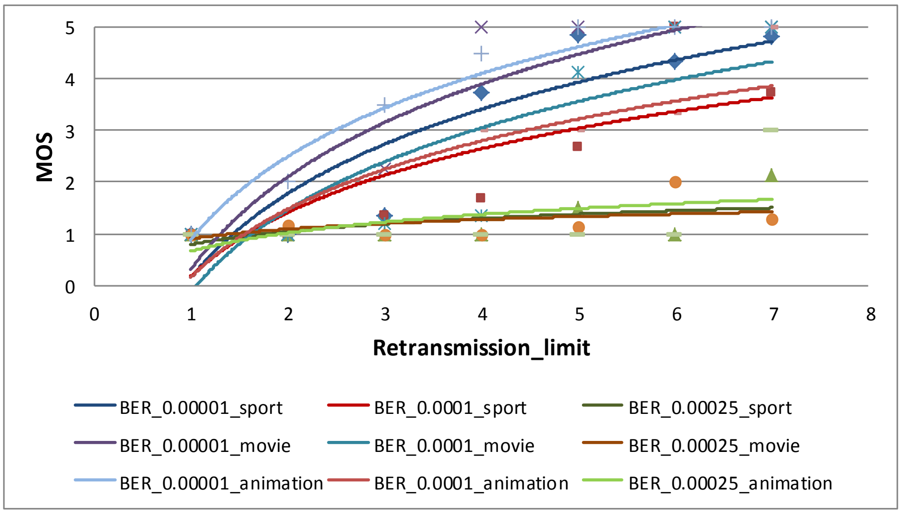

Figure 39: Impact of retransmission limit on video QoE

Figure 39 shows the impact of MAC-level retransmission limit on video QoE. The MAC protocol implements retransmission to avoid losing packets during transmission on air. Using different retransmission limit values in different channel conditions can increase the performance of video transmission. To analyze this effect, different values of BER were introduced in the channel. We can see that as the retransmission limit increases, the video QoE increases for both high and low BER case. However, the MOS value does not improve after certain level of BER and only increasing the retransmission limit value cannot improve the video QoE in high BER. This is caused by numerous frame corruptions due to high error prone channel conditions. For instance, we can see that at high BER, the MOS values are always between 1 and 2.

From the experimental analysis, we can say that video service vendors/operators should be more cautious while providing QoS/QoE to different type of video contents. It was observed that the impact is more on sport and animation contents than on movie content. Furthermore, to improve the video quality in highly loaded conditions, video traffic should be offloaded over less loaded channels. Moreover, to maintain better video quality, video traffic should be transmitted over channel with low BER. Additionally, queue length and retransmission limit value should be wisely selected in order to achieve high video QoE.

5.4 Objective QoE estimation based on MAC level parameters

In this section, we propose a QoE estimation system based on Random Neural Networks (RNN) [42] to estimate the impact of some MAC-level parameters on video quality. Since the nature of the relationship between MAC-level parameters and video QoE is non-linear, RNN learning method is used in this thesis work. Moreover RNN is considered appropriate for the identification of QoE perceived parameters in multimedia applications [97]. Here, we considered four MAC-level parameters: Bit Error Rate (BER), frame aggregation, number of competing stations, and traffic load. However, it must be noted that our methodology can also incorporate additional parameters as well. Since the variation of these attributes impacts the quality of the delivered video and consequently the level of user satisfaction, our objective is to design and implement a method to estimate the variation of user satisfaction level as a function of these different MAC-level parameters, both individually and combined. Figure 40, illustrates the video QoE estimation framework based on MAC level QoS parameters.

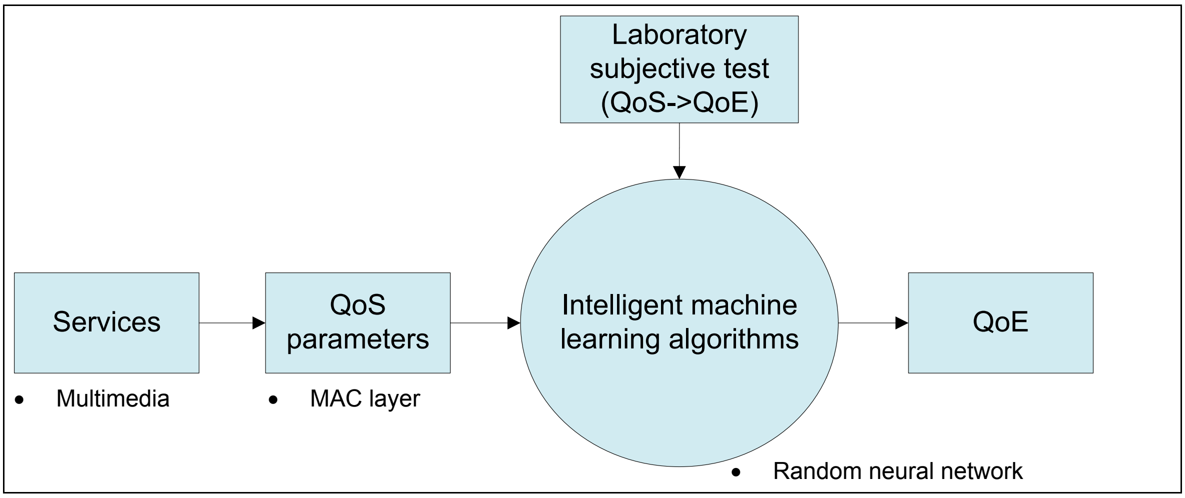

Figure 40: Video QoE estimation Framework (Based on MAC level parameters)

At first, we perform a set of subjective tests with real participants to measure the correlation between MAC-level parameters and QoE of video. Secondly, we propose an algorithm based on RNN that can estimate the video quality based on the MAC-level parameters. We simulated our system using “qoe-rnn” [97] and compared the estimated QoE output (estimated MOS or eMOS) with the subjective QoE obtained from the participants in a controlled test. The experiments validate our methodology showing acceptable correlation between our estimated and subjective QoE. This shows that the proposed video quality estimation method can effectively measure the impact of MAC-level parameters on the video QoE.

5.4.1 Methodology

In order to develop our video quality estimation technique, we followed a methodology that consists of conducting subjective tests with real end user participants in order to build a learning set and correlate objective MAC-layer metrics with the subjective QoE provided by the participants. This data set was then used to train the RNN for video QoE estimation.

5.4.2 Simulation Environment

For the simulations, we use the NS2 simulation environment embedded with the IEEE 802.11n module. A specific network topology was considered which consisted of static stations, where each station sends out traffic to its paired station. We used two types of traffic flows i.e., video and data. For video flow, we used the H.264/AVC video traces that were supplied as input using the Evalvid tool [95] and for the data flow, CBR (Constant Bit Rate) traffic was used. The characteristics of the video traffic are shown in Table 12.

Table 12: Video characteristics

Type	Bit rate (kbps)	Resolution	Frame per sec	Duration (Sec)
Video clip	2508	640 x 360	24	15

Different simulation scenarios were adopted to examine the performance of video under the combined effect of different MAC-level parameters. In our experiments, we have considered BER, number of competing stations, load, and aggregation as the most influencing MAC parameters. The value of the retransmission limit and the queue size are considered to be constant as in practice they are not changed frequently. Aggregation is also considered as one of the important parameters as its effects can be seen in different BER conditions and since it is considered as one of the fundamental parameters in the IEEE 802.11n standard [91].Simulation parameters that are common for all the scenarios are presented in Table 13.

In order to construct different video clips for the subjective test, we streamed the original video clip from the source node to the destination node and simultaneously varied the MAC-level parameters. This resulted in a variety of video clips at the destination, ranging from distorted to distortion less. In total 390 of such video clips were constructed.

Table 13: Simulation parameters

Parameters	**Values**
Simulation Area	250 m X 250 m
Simulation time	30 s
Node mobility	Static
Bandwidth	96 Mbps
MAC layer	IEEE 802.11n
Video packet size	1052 bytes
CBR packet size	1000 bytes
Time slot	20 s
SIFS	10 s
Antenna Type	Omni Directional
Network Interface Type	Wireless/Physical/MIMO
Interface Queue Type	Aggregation queue
Number of antennas	4
MAC Retransmission limit	7
Queue Length	200 packets
Transmission protocol	UDP

Thereafter, for the subjective test, both the constructed and the original video clips were shown to the participants in a random order. The subjective test was performed in a closed room to reduce the effect of external noise and light. Participants rated each video clip according to the perceived impairment giving one of the following MOS scores as described in section 2.2.3.

During the subjective test, on average, each participant rated 40 video clips. After rating 20 video clips, they were allowed to take a pause of 5 minutes. A total of 25 users were registered for the test, which is considered reasonable for this kind of subjective tests [96]. The dataset collected from the subjective test was divided into training set and test set for training and validating the RNN respectively. Figure 41 illustrates the experimental environment for the subjective tests.

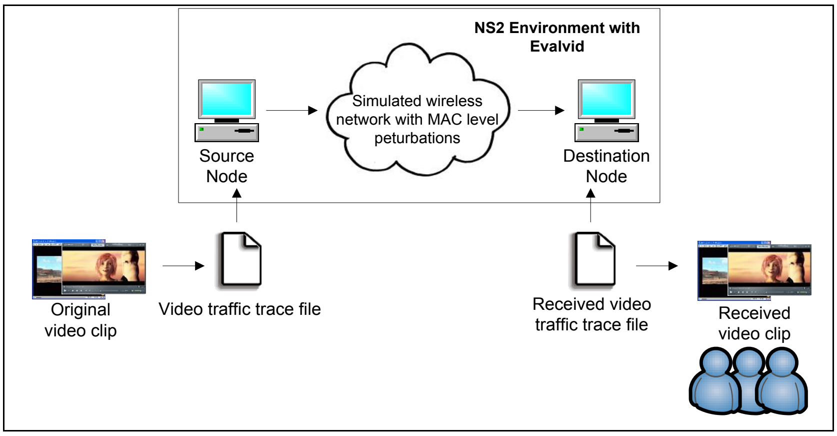

Figure 41: Experimental environment

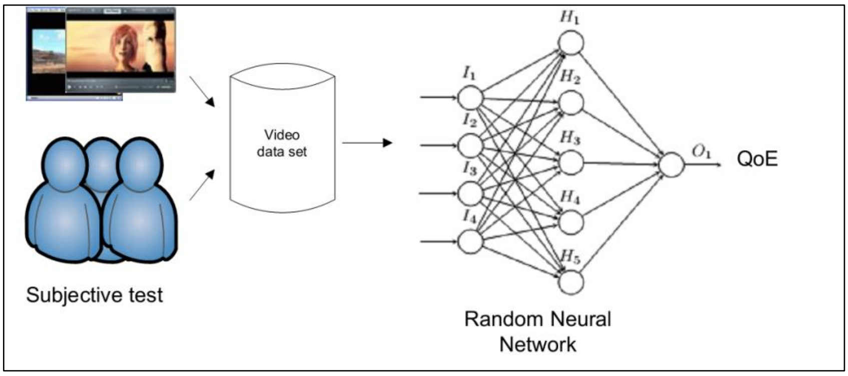

Figure 42: QoE estimation Module

5.4.3 QoE estimation module with Random Neural Network

Figure 42 presents the QoE estimation module with RNN. The subjective data set (training set) that contained the MOS score for each distorted videos were used to train the RNN. The RNN used for QoE estimation has 4 input nodes (MAC-level parameters), 5 hidden nodes (selected based on the best root mean square error) and 1 output node (QoE). In Figure 42, 'I' represents input node, 'H' represents hidden node and 'O' represents output node. Once the RNN has been trained, it can be used for real time QoE estimation without any human participation. The output of RNN is a MOS score which is as close as possible to those voted by users.

5.4.4 Experimental results

For the experiments, we analyzed the combined effect of different MAC-level parameters on video QoE. Following the simulation process explained in Section 5.4.2, we obtained different sets of videos containing different levels and types of MAC-layer perturbations. The subjective QoE values of each video were obtained from different participants. The estimated QoE were then derived from these subjective QoE dataset (test set) using our proposed RNN system. In order to validate the proposed methodology based on RNN, we compared the results of subjective tests with those obtained from our proposed system.

In Figure 43, we present the variation of subjective QoE from users and objective QoE from RNN for a video content. Each red point in the figure represents the objective QoE of a particular video clip and the blue points represent the subjective MOS. The Pearson correlation coefficient between subjective and objective QoE was found to be 0.89, which indicates high correlation. In Figure 44, we also present the probability distribution of the difference between the participants' subjective scores (MOS) and the estimated scores. We can see that in around 58% of the tests the score differences were less than 0.5. It reached 78% for score differences less than 1. This estimation accuracy thus emphasizes the ability of the proposed system to measure the impact of the network conditions on the user satisfaction.

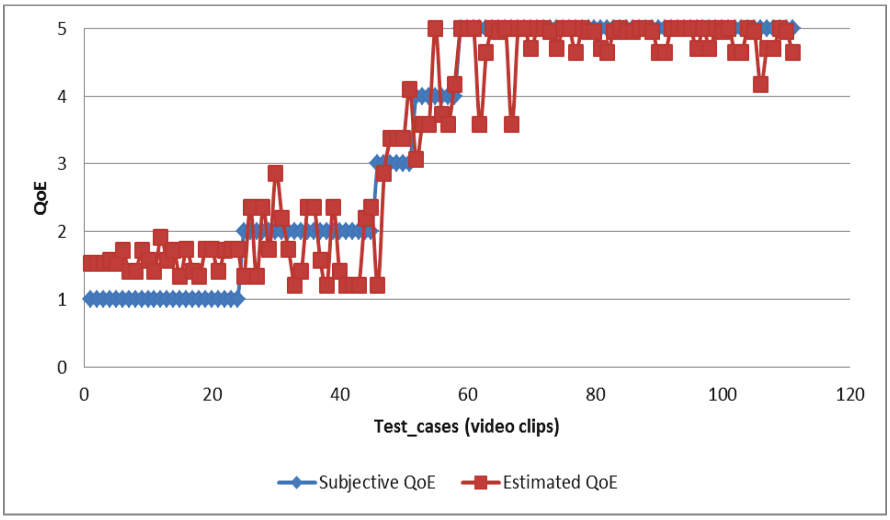

Figure 43: Comparison between subjective and estimated QoE

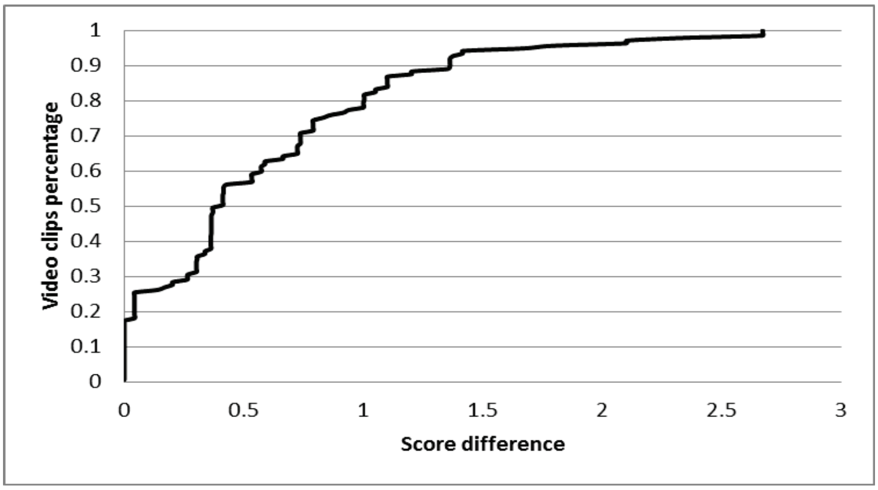

Figure 44: Probability distribution of the subjective and estimated MOS difference

In addition, the trained random neural network was tested with different values of the MAC parameters. We gave different sets of inputs that were not present during the training phase to analyze the variation of QoE in different MAC conditions.

In Figure 45, the variation of QoE with respect to BER and the number of competing stations is presented. Each station has 1 Mbps of traffic and aggregation length of 64000 was used for the purpose. The QoE of video degrades with both the increasing number of users and the channel error (BER) in the system. Even in a very good channel conditions (BER 0.00001) the quality of video can degrade below a MOS value of 4 when the number of competing stations increases beyond 16. On the other hand for the same 16 stations, the MOS can fall below 3 in medium error conditions (BER 0.00007) and may fall below 1 in high error conditions (BER 0.00095). In addition, high channel error (BER 0.00095) can severely affect the performance even when there are just two stations competing in the system.

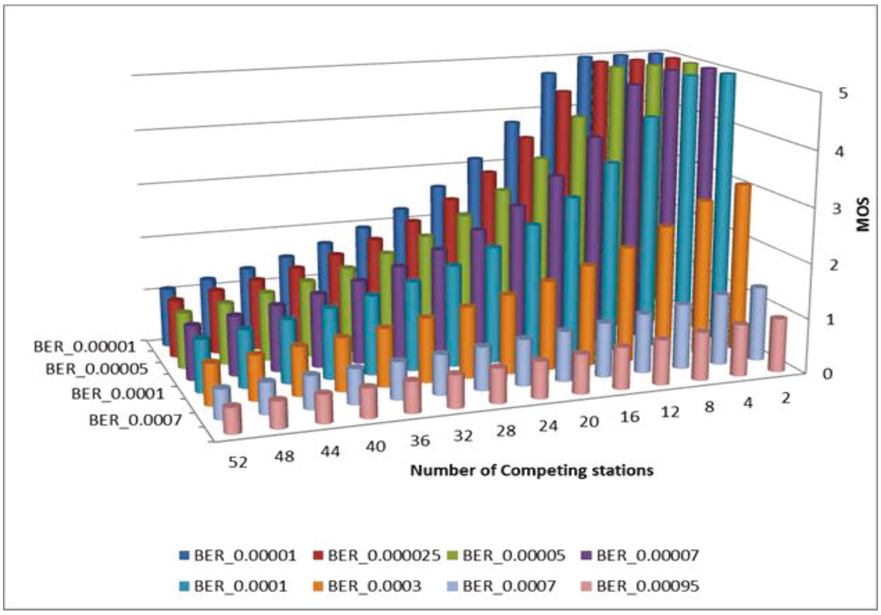

Figure 45: Variation of QoE with respect to BER and the number of competing stations

In Figure 46, the variation of QoE with respect to BER and the aggregation length is presented. 5 stations transmitting each 1 Mbps of traffic were used for the purpose. We can see that in low BER conditions (BER 0.00001, 0.000025, 0.00005, 0.00007) it is better to use high aggregation values than lower ones so as to provide better QoE values (MOS 4,5). However in the case of high BER (BER 0.0003, 0.0007, 0.00095), even aggregation cannot help to improve the QoE of the videos. This is true because in high BER scenario most of the frames are lost and even retransmissions cannot help to improve the quality.

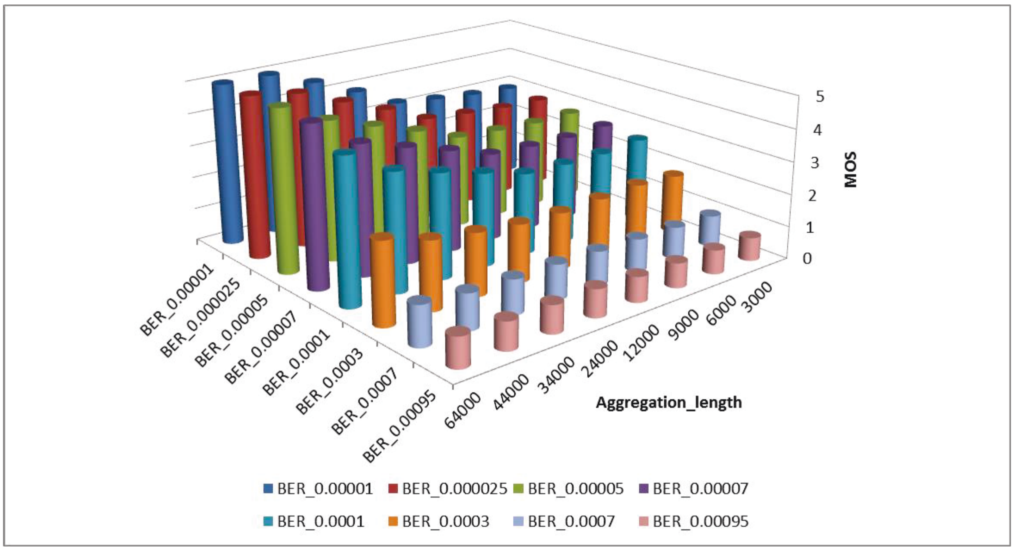

Figure 46: Variation of QoE in different BER and aggregation lengths

Therefore, we can see that using our proposed methodology we can predict the impact of different MAC level parameters on video QoE. This kind of module can be used in the design on Soft MAC techniques that can adapt parameters according to the MAC conditions in order to provide better QoE.

5.5 Conclusion

In this chapter, we analyzed the impact of different MAC-level parameters on the QoE of video traffic over IEEE 802.11n wireless network. We used H.264/AVC video traffic to evaluate the performance. We considered different MAC-level parameters like BER, number of competing stations, queue length, and retransmission limit and evaluated the video QoE. Intensive subjective tests were performed to analyze the effect of these parameters on video QoE. The experimental results show the mapping between MAC-level parameters and video quality. The study shows that careful parameterization of MAC-level parameters can provide an improved QoE for video streaming applications when used over a wireless network. We then proposed a methodology based on Random Neural Network (RNN) for objective QoE estimation. The RNN was trained with a subjective dataset for QoE estimation. The experimental results show that the estimated QoE correlates with subjective QoE with reasonable accuracy. This estimation system can be used in the design on Soft MAC techniques that can adapt parameters according to the MAC conditions in order to provide better QoE.

Chapter 6

QoE estimation of OTT services

6.1 Introduction

OTT services refer to “Over The Top” services, that provide services over the Internet and are not regulated by traditional distributors (ISP or system operator). From a service provider perspective it is a normal Internet traffic and they are only responsible for transporting IP packets and has no right to control, view, copyrights, and/or make other redistribution of the content [98]. These are the third party content that operates on top of the Internet and are lower in cost than traditional services like cable television, IPTV, VOD, etc. Most the OTT applications that we find in Internet today are related to media and communication like NowTV, Netflix, WhereverTV, Hulu, whatsapp, MyTV, etc. Moreover, OTT Video streaming shares the biggest part of Internet video demand [156].

Users are subscribed to telecommunication operators and service providers and they expect to have best overall performance. With the increase of OTT-based services, high bandwidth is needed to serve them. This bandwidth scarcity will have negative impact on user experience. Even though service providers may not own OTT services, they are still expected to provide superior user experience and performance [99].

Nowadays, most of the OTT video streaming applications use adaptive HTTP/TCP as a transport mechanism for multimedia streaming. In adaptive HTTP based video delivery, video contents are encoded in different data rates and stored as small fragments (a few seconds each) on the server. This will allow applications to switch between different data rates based upon different network conditions and other variables [39]. This will eliminate the unwanted freezing events at the client terminal however, reducing the multimedia quality, i.e., the resolution (bit rates). Also, the adaptive bitrate switching (adaptive HTTP/TCP) which is deployed to avoid freezing can cause irritation to users because of frequent change in multimedia quality as bit rate changes. Different implementations of adaptive HTTP based video delivery are Apple’s HTTP Live Streaming (HLS), Microsoft’s Silverlight Smooth Streaming (MSS), Adobe’s HTTP Dynamic Streaming (HDS), and Dynamic Adaptive Streaming over HTTP (MPEG-DASH) [39].

Subjective experiments are considered to be the most valid methodology to assess the QoE. Subjective experiments are typically conducted in a controlled laboratory environment. Objective or computer software assisted methods have been largely seen as an alternative approach, to get around the complications involved in the lab-based subjective experiments. Crowdsourcing based subjective experiments have gained attention to replace the needs of lab-based tests and these experiments offer promising correlation with the later [26]. This methodology mainly involves collecting subjective assessment of quality through ubiquitous streaming via the Internet. This enables the investigators to receive opinion from a vast variety of subjects; in a time-flexible, test-data size scalable, and swift manner.

This chapter presents an insight into a crowdsourcing-based subjective perceptual preference of various adaptation scenarios. Specifically, paired-comparison of the test videos has been performed [49] and a suitable technique has been used to convert the obtained preferences data into the regular mean opinion score (MOS), as usually obtained in an absolute category ranking.

6.2 Related Work

Based on the subjective test, there has been some work done on estimating the quality of adaptive video streaming. The work presented in [157], studied the optimum number of coding quality levels that could be used in adaptive video system in mobile video content. Authors in [158] studied the effects of frame rate and resolution on the user's perceived QoE. In [159] authors studied the impact of quality variation on the users' QoE. Different scenarios such as rapid and gradual bit rate drop as well as oscillating the quality were considered as a use case. Furthermore in [160], authors presented the impact of fluency, bitrate distribution, startup bitrate level, bitrate switching frequency, etc. on the users perceived QoE. Similarly, in our work we studied the impact of different adaptation scenarios (bit rate switching, freezing) on adaptive bitrate streaming but using a crowdsourcing based subjective test. We analyzed whether users prefers bit rate switching or freezing in different adaptive scenarios. A novel web based application for performing subjective test is developed for this purpose.

6.3 Methodology

The methodology for QoE assessment of OTT video services involves performing a crowdsourcing-based video subjective test. A novel web based application was developed for performing subjective tests which was embedded inside a crowdsourcing platform. In the subjective test, we used PC (Pair wise comparison) methodology where users provide score viewing two video sources and expressing their preference. The PC scores obtained were converted to mean opinion scores using Bradley-Terry model [100]. The framework for QoE estimation of OTT services is shown in Figure 47.

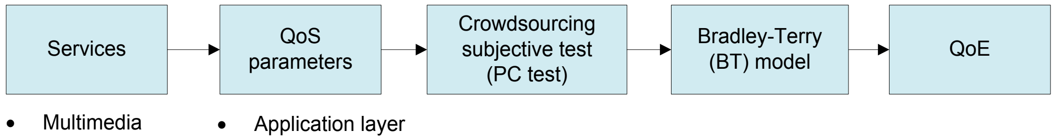

Figure 47: Framework for QoE estimation of OTT video services

6.4 Test Background

The videos for our subjective test are originally from the subjective lab experiment detailed in [101]. The original videos were all in 1280x720 resolution with a frame rate of 24 fps and encoded using the high profile for H.264/AVC at 4 different bitrates: 600 kbps, 1 Mbps, 3 Mbps, and 5 Mbps. Seven different sources were used, three sources were taken from entertainment movies and the rest of the content were from a soccer match, a sports documentary, a newscast, and a concert.

Several adaptation scenarios for the videos were produced in the original experiment, such as going from a high to a low bitrate in a stepwise manner as shown in Figure 48. In our subjective experiment we used the following scenarios from the original experiment: Gradual Decreasing (GD), Rapid Decreasing (RD), constant 600 kbps (N600), constant 1 Mbps (N1), constant 3 Mbps (N3), and constant 5 Mbps (N5). Additionally, we introduced new buffering scenarios to test the quality perception in relation to the aforementioned scenarios. The buffering scenarios include: 1 Freezing event for 2 seconds in the constant 3 Mbps video (1F3M), 2 Freezing events for 1 second each in the constant 3 Mbps video (2F3M), and 1 Freezing event for 2 seconds in the constant 1 Mbps video (1F1M). In total 9 different scenarios were used, resulting in a total of 63 stimuli.

Code	Characteristic of scenarios
DGR2	Gradually decreasing the quality (5-3-1-600) with 2 seconds chunk
DRP2	Rapid decreasing the quality (5-600) with 2 seconds chunk
1F3M	1 Freezing event for 2 seconds in the constant 3 Mbps video
2F3M	2 Freezing events for 1 second each in the constant 3 Mbps video
1F1M	1 Freezing event for 2 seconds in the constant 1Mbps video
N5	No degradation- The whole segment at 5Mbps
N3	No degradation- The whole segment at 3Mbps
N1	No degradation- The whole segment at 1Mbps
N600	No degradation- The whole segment at 600kbps

Figure 48: Adaptation Scenarios

Crowdsourcing experiments should be as simple as possible for the subject; therefore we decided to follow the Paired Comparison (PC) methodology [49]. We used the optimized square design [25] based on our assumptions of the quality levels to get reliable measurements and reduce the number of pairings. Using this method, our total test set consisted of a total of 126 pairings. These pairing were divided into 14 tasks with 3 videos from 3 different contents, i.e., 9 videos for each task. The participants were mainly from European countries and United States. We used screen tests [23] prior to the subjective test to filter out potential malicious workers. In total 266 workers participated in the experiment.

6.4.1 Crowdsourcing based subjective test

Crowdsourcing-based video subjective test uses a web-based application to perform a subjective test using a crowdsourcing platform. The crowdsourcing platform helps to perform the test with a global worker pool. There exist different crowdsourcing platforms in the market such as Mechanical Turk [154], microworkers [153], clickworker [155], etc. Some of the web-based multimedia quality assessment applications include [26] [102] [103] [161]. Crowdsourcing experiments can be vulnerable to cheating; therefore different screen test mechanisms are employed in a crowdsourcing experiment [23] to make the test reliable and resilient.

In this work, we developed a new web based application for performing subjective test and used the microworkers [153] crowdsourcing platform to perform the test [10].

6.4.2 Screen test

Screen tests are used to find the end user watching conditions for use in the video and image assessment [152]. If the watching conditions at the workers end are not favorable then the test scores available are not reliable. In our web based application we applied two screen test mechanisms as described in [24].

In screen test 1 as shown Figure 49, participants were asked to select highest and lowest numbers in faint white color from white background. The numbers are moving in the screen and not every number has equal visibility. Visibility of these numbers also depends on screen orientation, screen resolution, screen brightness, screen color combination, and participant's eyesight. The numbers between 1 and 7 are visible to the participants. The higher the number is, the lower the visibility.

In screen test 2 as shown in Figure 50, participants were asked to click visible stars in faint black color from black background. These stars are stable in the screen and not every star has equal visibility. Visibility of these stars also depends on screen orientation, screen resolution, screen brightness, screen color combination and participant's eyesight. Star numbers between 1 to 7 are visible to the participants. The higher the number is, the lower the visibility.

From the result obtained from screen test, we selected 215 workers out of 266 workers.

Figure 49: Screen test 1

Figure 50: Screen test 2

6.4.3 Pairwise comparison and Bradley-Terry (BT) model:

The preference score obtained from the users are processed for preference matrix. The preference matrix contains the probabilities i.e., each element of the rows in the preference matrices corresponds to the amount of times that a video i is preferred over a video j. The actual number of objects that are compared is nine however; they are not compared to all possible combinations but according to the optimized square design. Table 14 shows the preference matrix for soccer video source. For example we can see that video (1F1M) is preferred 5 times to video (1F3M).

Table 14: Soccer comparison matrix

i/j	1F1M	1F3M	2F3M	DGR2	DRP2	N1	N3	N5	N600
1F1M	0	5	0	0	0	4	0	6	10
1F3M	11	0	11	0	0	6	4	0	0
2F3M	0	5	0	0	9	0	4	0	8
DGR2	0	0	0	0	7	10	5	6	0
DRP2	0	0	7	7	0	4	0	0	10
N1	12	9	0	5	10	0	0	0	0
N3	0	12	12	11	0	0	0	7	0
N5	10	0	0	10	0	0	9	0	12
N600	6	0	8	0	6	0	0	4	0

In order to convert this preference value to quality score or MOS scale we use the Bradley-Terry-Luce [100] model.

If P_{ij} represents the probability of choosing stimuli i over j, the BTL model can be represented as:

$$P_{ij} = \frac{\pi_i}{\pi_i + \pi_j}$$

Where, π_i represents the quality score for stimuli i and can take a value from 0 to 1. This expression can be reformulated using preference values, i.e.,

$$P_{ij} = \frac{m_{ij}}{m_{ij} + m_{ji}}$$

Here m_{ij} is the frequency of stimuli *i* being preferred over j. Now, the π_i can be computed by maximizing the log-likelihood function, which is represented as (for our 63 stimuli):

$$L(\pi_1, \pi_2, \pi_3 \ldots \ldots \pi_{63}) = \sum_{i=1}^{63} \sum_{\substack{j=1 \\ j \neq i}}^{63} P_{ij} \ln\left(\frac{\pi_i}{\pi_i + \pi_j}\right)$$

To solve this equation we used the optimization routine in a software package i.e. Matlab function OptiPt.m as mentioned in [104]. Lastly, the obtained estimations of the probability values have been normalized to the regular scale of 1-5 MOS.

6.5 Experimental Results

In order to validate the results obtained from the crowdsourcing experiment, we compared the opinion scores obtained from the lab experiment to the crowdsourcing experiment as shown in Figure 51. The results in Figure 52 shows that the opinion scores obtained from both the experiments are strongly correlated, though not to the degree one of would expect. This can be due to the differences in the test setup, such as evaluation method, viewing environment, and the introduction of new distortions. Our experiments verify the results from earlier studies, e.g., [105], that buffering events has a high impact on the QoE. Due to this, users generally prefer viewing videos at lower bitrates than having buffering events in videos at higher bitrates.

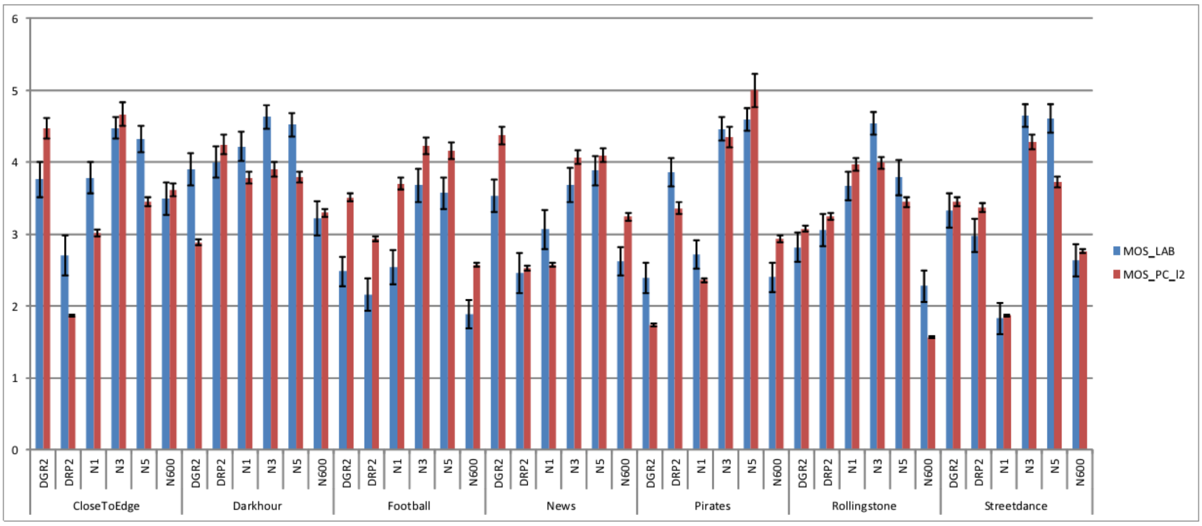

Figure 51: Comparison between laboratory and crowdsourcing subjective test

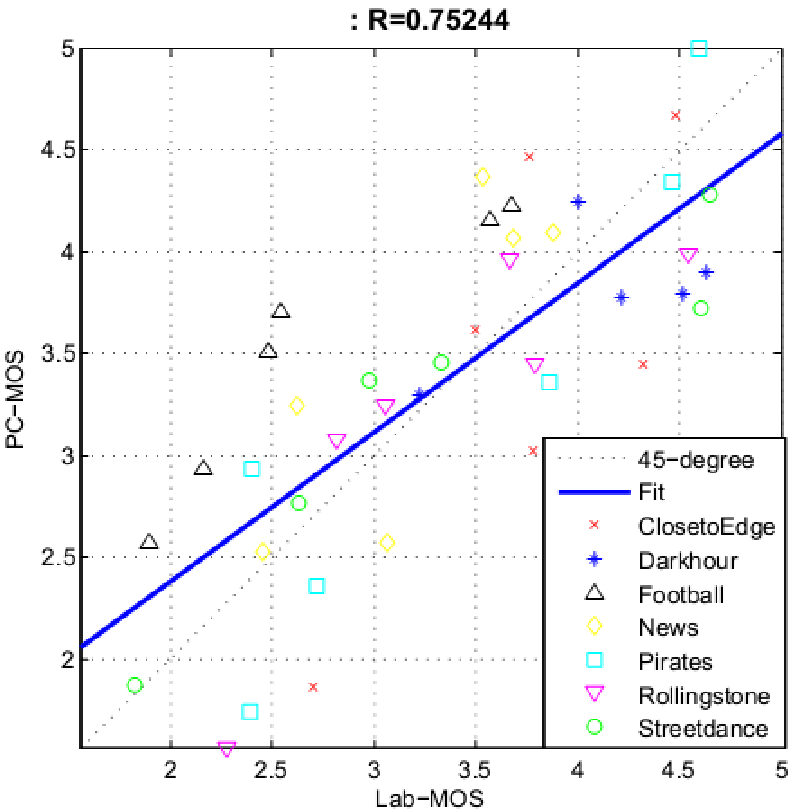

Figure 52: Correlation between laboratory and crowdsourcing subjective test

The quality of the videos can also be compared versus the average video bitrate. This has been illustrated in Figure 53, where the mean of the subjective scores has been calculated over the video contents. Generally, users prefer videos at higher bitrates, i.e., 3 or 5 Mbps and the difference between them is probably more due to the difference in content than the difference in compression levels. Users dislike buffering events and it seems that the frequency is more important than the total duration of these events (both videos at 3 Mbps with buffering has a total buffering time of 2s), which is in line with earlier studies such as the one presented in [106]. But if the bitrate is high enough and the frequency of the buffering events is low enough, e.g., the 1F3M video, this seems to be a viable alternative to decreasing the bitrate of the video or having a constant low bitrate, e.g., 600 kbps or 1 Mbps.

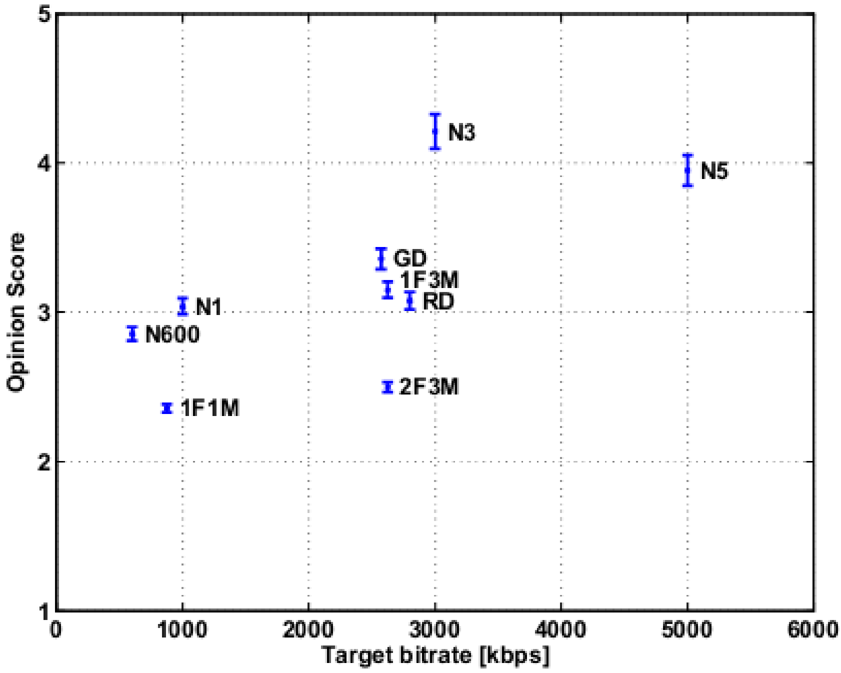

Figure 53: MOS score vs target bitrate

6.6 Conclusion

In this chapter, we have analyzed the different adaptive scenarios based on crowdsourcing-based subjective tests. We also have developed our own web based application for subjective test assessment which has been integrated in the crowdsourcing platform. The results obtained from the crowdsourcing tests were satisfactorily correlated to lab based test which validates our crowdsourcing experiments. Moreover, by analyzing different adaptive scenarios, we found that users prefer video at higher bit rates, however, users are very sensitive to buffering events and their experience decreases with increase in frequency of buffering events. Moreover, if the bitrate is high enough and the frequency of the buffering events is low enough, e.g., the1F3M video, this seems to be a viable alternative to decreasing the bitrate of the video or having a constant low bitrate, e.g., 600 kbps or 1 Mbps.

Chapter 7

QoE estimation for web services

7.1 Introduction

Web Services (WSs) are self-contained software systems that can be published, advertised, located, and invoked through the web, usually relying on standardized XML technologies (REST, SOAP,WSDL, and UDDI [107]) for description and publication, and on Internet Protocols for invocation[108]. Popularity of web services is growing rapidly which leads to a large number of web services or application with similar features. This offers users a number of options and introduces a higher demand on price, response time, availability, reliability, service performance, and other non-functional attributes for selecting a web service.

The availability of a large number of web services providing similar functionalities and features has increased the need for sophisticated discovery and selection processes that can better meet the user's needs. The discovery process is a process of identifying or locating a web service that fulfills certain functional properties. On the other hand, the selection process refers to evaluating and ranking the discovered web services for selecting the one that fulfill a set of non-functional properties [109]. As indicated in [109], the "*functional properties describe what the service can do and the non-functional properties depict how the service can do it*". Non-functional properties involve qualitative or quantitative features such as, throughput, latency, response time, integrity, availability, security, etc. [110], [111]. However, a selection process which relies only on a partial set of non-functional properties can be misleading as this will not necessarily reflect the user's satisfaction. Thus, as we propose here, we need a methodology that considers several parameters to estimate the expected user experience, with each having a greater or lesser impact on the resulting estimation.

Users' demand and expectation for web technology is accelerating with time. Users' gets intolerant if the content is not served in expected time and easily switch to other options if their needs are not fulfilled [112]. About 90% of the people do not want to complain for the low service quality. They just leave the service and move to another ones [113]. So service providers and operators should not wait for user feedback for improving the service quality;

instead, they should continuously monitor QoE and improve it as required. They should provide users with services that can offer high QoE values.

Generally, QoS parameters are used for selecting web services, which do not necessarily reflect the user's satisfaction towards a particular web service. QoS parameters reflect network and service level performance; however, they do not address the user's reaction to the service or application. Therefore, it is necessary to derive a correlation between the QoS parameters and the subjective QoE, so that it is possible to identify the impact of different QoS parameters on the QoE experienced by the users. This motivates research communities for further studies in quality estimation of web services. In recent years, a high amount of research work has been done on QoE assessments for voice and video services. However, little has been done on QoE assessment of web technology.

In this chapter, we propose a methodology to estimate the quality of web services based on a fuzzy-rough hybrid algorithm. Figure 54 shows the web QoE estimation framework based on web QoS parameters. Fuzzy expert systems [84] are good at making decision with imprecise information; however, they cannot automatically formulate the rules that they require for making the decisions. Therefore, a fuzzy-rough hybrid expert system is proposed where rough set theory is used to define the rules necessary for the fuzzy expert system. We consider three web QoS parameters: execution time, availability, and reliability as important indicators for QoE estimation. These parameters have been selected because their variation affects the efficiency of web services and the overall user experience; but, it must be noted that our method can also easily integrate more or other parameters. At first, we conducted subjective tests in a controlled environment with real users to correlate QoS parameters to subjective QoE, i.e., Mean Opinion Score (MOS). Based on the results from these subjective tests, we derived membership functions and rules for the fuzzy system. The probabilistic approach is used for deriving membership functions and the Rough Set Theory [13] is used to derive rules from the subjective tests. We simulated our system in MATLAB and compared the estimated QoE output with the subjective QoE obtained from the participants in the controlled tests. The results from the experiments validated our methodology showing high correlation between our estimated QoE and the subjective QoE.

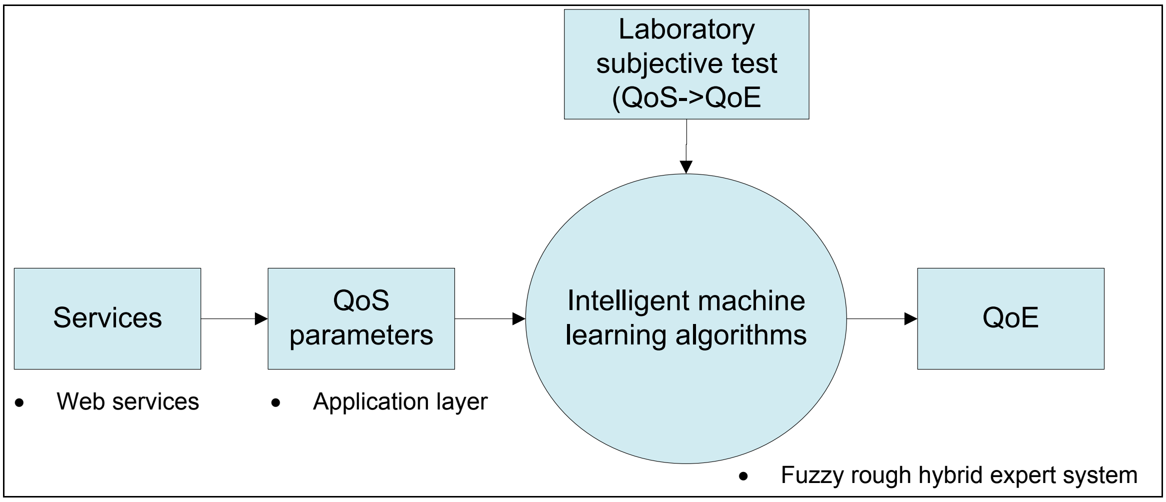

Figure 54: Web QoE estimation system (QoS parameters)

7.2 **Related Work**

The World Wide Web (WWW) has dominated the Internet since it was commercialized in 1995. It has allowed interconnecting the world by sharing information related to daily life activities, such as education, business, commerce, science, social networking, and entertainment. This has fuelled the increase, in the past years, of an enormous amount of web services and applications based on web technologies.

QoE was first defined in the context of multimedia services. A high amount of research attention has been placed in estimating QoE and correlating network QoS with QoE of multimedia services as shown in [7], [114], and [82]. In the case of web services, user satisfaction is often measured in terms of response time. If users need to wait a long time during web service session, it will be perceived negatively. Due to the limitation of resources in mobile networks, the situation with this respect can become even more critical. The effect of response time on user behavior in the web is presented in [115], [116], and [117]. Response time is one of the important parameters; however, it is not sufficient to evaluate web services quality [118]. Two practical approaches to measure QoE are presented in [113], where a service level approach using statistical samples and a network management system approach using QoS parameters is used. Here the authors identify different key performance indicators for mobile services based on reliability (i.e., service availability, service accessibility, service access time, and continuity of service) and comfort (i.e., quality of session, ease of use, and level of support). However, it does not provide any methodology that could map QoS parameters to QoE. The work presented in [119] proposed models that allow selecting web services based on client constraints and QoS information gathered by the service providers at runtime. A new scheme for QoS-aware web services selection which exploits fuzzy logic to locate and select the right service based on the customer's preference or satisfaction degree is presented in [110]. However, the works presented in [119] and [110] lack any experiments and validation of results. Authors in [120], present a web service selection methodology based on a context-based ontology and quality of service measurements. In [121], the authors proposed a dynamic QoS computation model for web service selection based on generic and business specific criteria. A generic quality criterion includes execution price, execution duration and reputation, and business specific criteria including usability. In the work [122], the authors presented QoS-based web service selection criteria, where they propose introducing web service ping operations in all web services for measuring web service latency and service availability. All the above web service quality estimation and selection methods ([119], [120], [121], [122], and [110]) are based on QoS parameters that lack end user participation and do not classify estimated quality into MOS scores. Web service QoE estimation method based on a correlation function between web QoS (execution time, reliability, and availability) and QoE is presented in [108]. It uses a regression analysis tool to calculate indexes of a correlation function from the subjective test data; however, the quality estimated by this method has high ΔMOS error margin.

The main contribution of our research is to present a novel method based on fuzzy-rough hybrid model to estimate the QoE of web services. Experiments performed show that this method correctly reflects the expected customer's preferences and satisfaction degree; and, thus, can be useful for selecting the right web service. The proposed web service quality estimation method is based on QoS-QoE correlation which is obtained through subjective tests. A fuzzy-rough hybrid expert system takes into consideration QoS-QoE correlation to rate each web services with a QoE score (which is in the range of 1 to 5 as in the case of MOS scores). The QoE scores can effectively represent the level of the user's satisfaction (excellent, good, fair, poor, or bad) towards a particular web service. Correspondingly, the scores can be used to rate different service providers. It can also be used for improving the service experience by distributing web clients towards different web service provider's to obtain, for instance, that high priority web clients are served with excellent quality web services, and low priority web clients with lower quality services.

The methodology that we propose relies on subjective tests. Subjective data are strongly influenced by the customer's feeling and experience. So, the correlation between QoS parameters and the participants' QoE remains imprecise, uncertain or ambiguous due to various human mental states and profiles, making the preferences over the criteria hard to quantify. The fuzzy approach [84] can deal with the consumers' imprecision by creating preference relations through the use of fuzzy sets and inference rules. An advantage of using fuzzy expert systems is that they are simple and computationally less intensive. Fuzzy expert systems are good at making decisions with imprecise information; however, they cannot automatically acquire the rules they require for making the decisions. Therefore, a fuzzy-rough hybrid expert system is proposed where rough set theory is used to acquire the rules for the fuzzy expert system. Rough Set Theory is used for discovering patterns, rules and knowledge from the datasets as in [123] and [13]. Rough Set Theory has many advantages as, for instance, that it does not have information loss and it is flexible and extendable as compared with other data mining technologies [124].

7.3 Methodology

In order to develop our web quality estimation technique, we followed a methodology that consists in conducting subjective tests with end user participants in order to build a learning set that correlates web QoS parameters with the subjective QoE. This correlation was then used to build the membership functions and inference rules of our fuzzy expert system for web QoE estimation. The methodology for designing web QoE estimation is shown in the Figure 55.

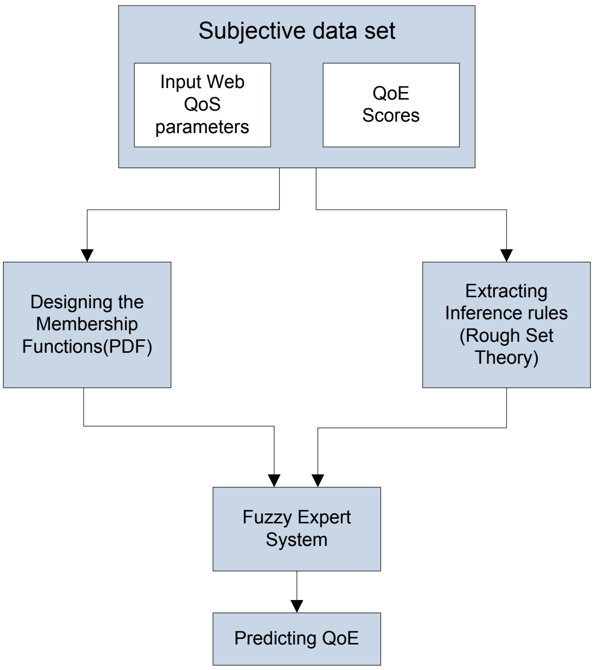

Figure 55: Methodology for Web QoE estimation

7.3.1 Subjective test

The subjective test platform used in [108] has been established to perform the subjective tests. In the experiment, an interactive web application has been developed to simulate typical web service architecture. Each user is asked to use this web application. From the users' responses, a MOS score is obtained. Three QoS parameters: reliability, execution time (in seconds), and availability (in seconds) are measured during the performance of the tests.

- Execution time is measured as a delay at the server level.
- Availability is measured as a period of server downtime, in which the service responds with "*Service unavailable, please retry again in a few moments*" after each request until a certain time has elapsed.
- Reliability is measured as a number of consecutive erroneous responses, in which the service responds with "*An error has occurred, please try again*", until the subject has retried the number of times defined by the test variable.

A total of 88 users were registered for the test which is considered reasonable for this kind of subjective tests [96]. The details of the experiments can be found in [108].

7.3.2 Membership functions

The membership functions required for the fuzzy expert system are designed using the subjective data set. The curve values of the membership functions represent the degree to which a particular QoS parameter value belongs to different QoE scores. We used a Probabilistic Distribution Function (PDF) to derive a membership function as described in [7] and [12]. For every QoS parameter, we built different probability distribution functions (PDF) (one function per QoE score) that provide the variation of the participants' ratio (%) with the QoS metric for a specific QoE score. This probabilistic information was changed into a fuzzy set by dividing the PDF by its peak value (normalized PDF) [11]. The triangular or trapezoidal fuzzy set represents the membership functions for the different QoS metrics. In our case, we reduce the five scale MOS classes to three scale MOS classes (low, medium, and high) for QoS parameters because it was very difficult to find the boundary region between fair and good, and bad and poor, which has thus been replaced respectively by medium and high classes.

Figure 56 illustrates the QoE scores membership functions associated with the execution time QoS parameter. For example, the execution time of 1.5 seconds has membership values of

0.5, and 0.5 respectively to the QoE scores low, and, medium. We note that a membership value of 1 represents a high degree of membership to the corresponding class and a decreasing membership value represents deviation from the class. Figure 57 and Figure 58 illustrate the membership functions for availability and reliability parameters respectively. Similarly, in Figure 59, the membership functions for the estimated QoE are defined according to the standard MOS definition.

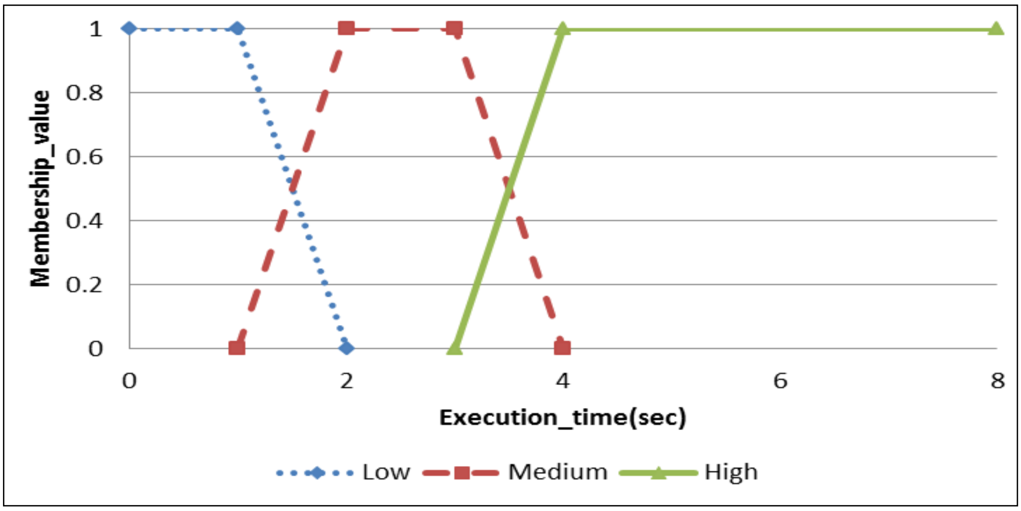

Figure 56: Membership function for execution time

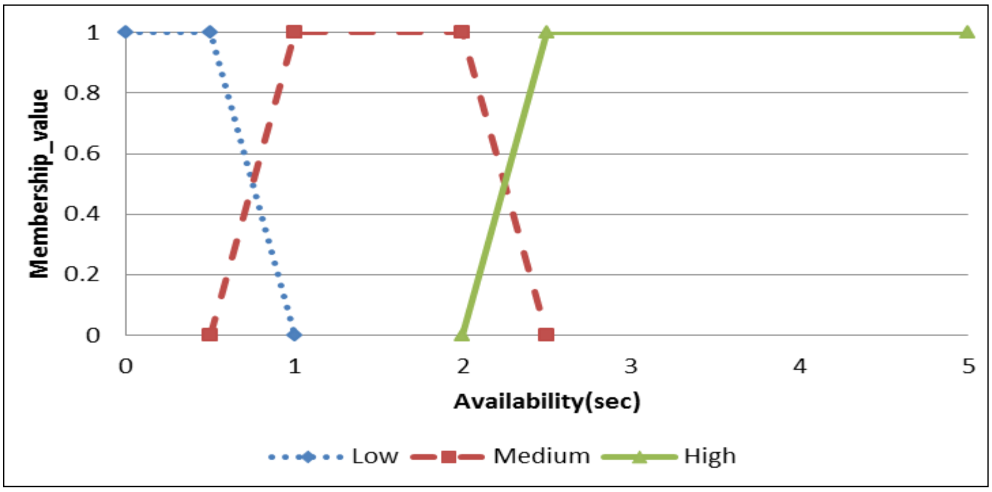

Figure 57: Membership function for availability

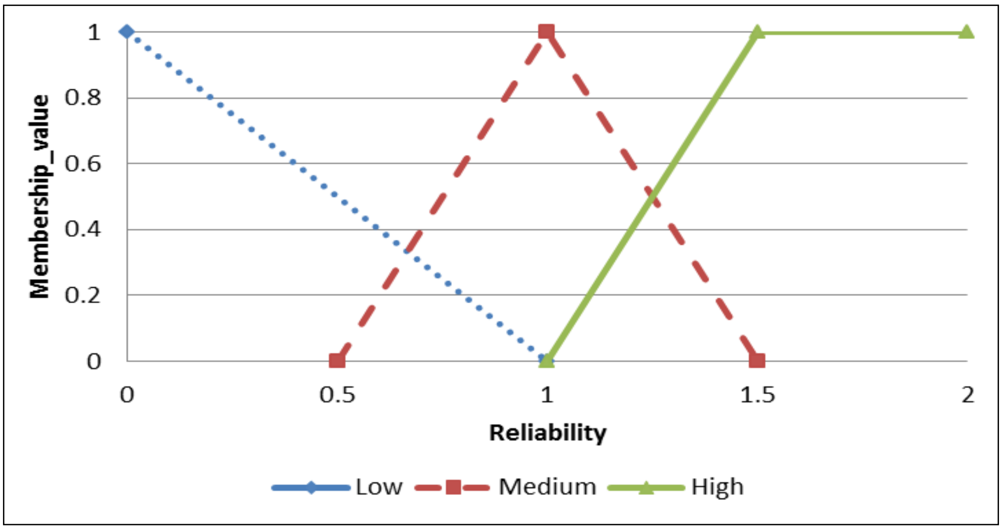

Figure 58: Membership function for reliability

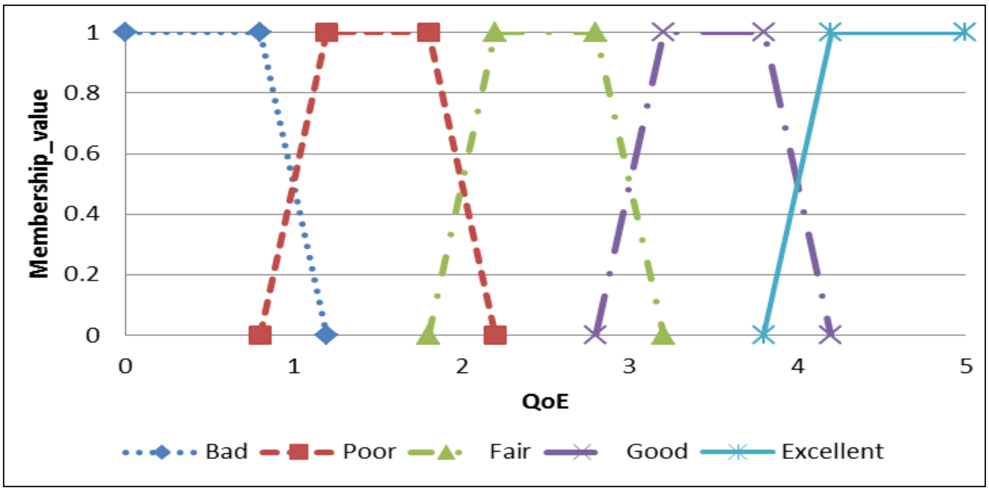

Figure 59: Membership function for QoE score

7.3.3 Inference Rules

We used the subjective test data set to derive the inference rules for the fuzzy expert system. Rough Set Theory is one of the well-known data mining techniques to generate classification/inference rules from the subjective data set [123] [124]. To apply Rough Set Theory on the subjective data set, we represented the subjective data set in the form of a conditional attribute set and a decision attribute set, and processed it through discretization arithmetic. Here the QoS parameters represent the conditional attribute set and the QoE score represents the decision attribute set. Table 15 is an original subjective dataset showing test results. Due to space limitation only ten objects among them are listed in the table.

Table 15: Subjective Test Results

Execution time	Availability	Reliability	QoE
3	0	0	4
0	0	0	5
4	0	2	2
2	2.5	0	3
1.5	2.5	0	3
2	5	2	1
1	8	1	2
4	0	0	4
8	2.5	0	3
8	0	1	1

Table 16: Discretization Table

Low	Medium	High
Execution time < 2 s	2 sec =< Execution time < 4 s	Execution time >= 4 s
Availability < 1 s	1<= Availability <= 2.5 s	Availability >=2.5 s
Reliability =0	Reliability =1	Reliability >=2

Table 17 represents QoE index decision making table derived from Table 15 and Table 16. Attribute values for QoS parameters have been processed through discretization arithmetic as shown in Table 16. Criteria for different QoS parameters in Table 16 where selected from the subjective dataset, where each QoS parameter were evaluated independently of other QoS parameters. The criteria's value represents the average value of all those particular subjective test cases.

Table 17: QoE Index Decision Making Table

Execution time	Availability	Reliability	QoE
Medium	Low	Low	Good
Low	Low	Low	Excellent
High	Low	High	Poor
Medium	Medium	Low	Fair
Low	Medium	Low	Fair
Medium	High	High	Bad
Low	High	Medium	Poor
High	Low	Low	Good
High	Medium	Low	Fair
High	Low	Medium	Bad

We used the Rosetta software [125] which is a Rough Set toolkit for analysis of datasets to generate inference rules. The Johnson's greedy algorithm [126] was used to find the reduct (explained in section 3.7.3). After the reduct is performed, different decision rules were extracted from the dataset. If a rule predicts more than one QoE class then the QoE class with the highest accuracy (highest percentage of votes) is considered to resolve the conflict between the rules. In this way, 15 rules were selected, which were used by the fuzzy expert system for estimating the QoE. The rules are the following:

1. If (Execution_time is Low) and (Availability is high) and (Reliability is Low) then (QoE is Poor)
2. If (Execution_time is Medium) and (Availability is Low) and (Reliability is Low) then (QoE is Good)
3. If (Execution_time is Low) and (Availability is Low) and (Reliability is Low) then (QoE is Excellent)
4. If (Execution_time is Medium) and (Availability is High) and (Reliability is Low) then (QoE is Bad)
5. If (Execution_time is Medium) and (Availability is Medium) and (Reliability is Low) then (QoE is Fair)

6. If (Execution_time is High) and (Availability is Low) and (Reliability is Low) then (QoE is Good)

7. If (Execution_time is High) and (Availability is Medium) and (Reliability is Low) then (QoE is Fair)

8. If (Execution_time is High) and (Availability is High) and (Reliability is Low) then (QoE is Fair)

9. If (Execution_time is Low) and (Availability is Medium) and (Reliability is Low) then (QoE is Poor)

10. If (Availability is High) and (Reliability is Medium) then (QoE is Poor)

11. If (Availability is Low) and (Reliability is High) then (QoE is Bad)

12. If (Availability is High) and (Reliability is High) then (QoE is Bad)

13. If (Availability is Medium) and (Reliability is High) then (QoE is Bad)

14. If (Availability is Low) and (Reliability is Medium) then (QoE is Bad)

15. If (Availability is Medium) and (Reliability is Medium) then (QoE is Bad)

7.4 Web QoE Estimation system

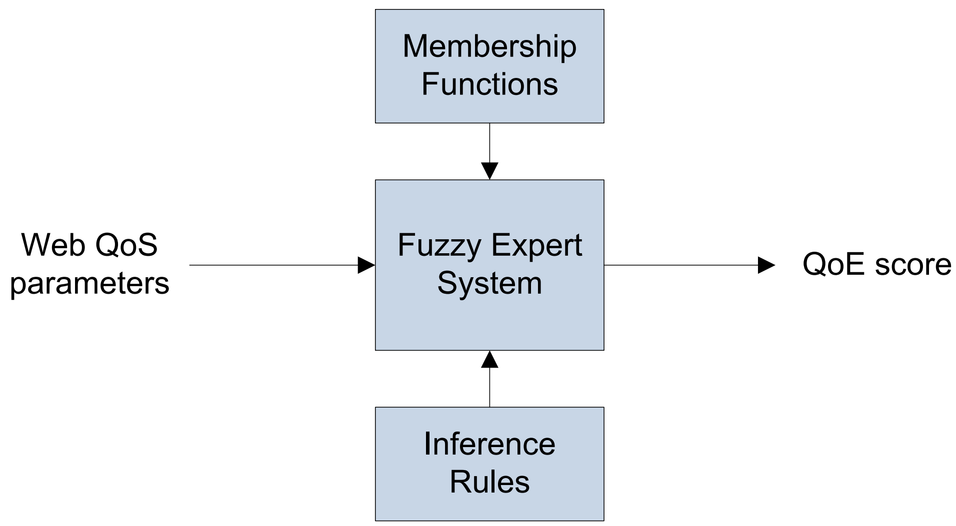

Figure 60: Web QoE estimation system

Figure 60 illustrates the web QoE estimation system. The fuzzy expert system with pre-defined membership function and inference rules (Section 7.3.2 and Section 7.3.3) acts as an intelligent system for web QoE estimation. The web QoS parameters are constantly fed to the fuzzy expert system that uses the pre-defined membership function and inference rules to estimate the web QoE.

7.5 Experimental Results

To validate the proposed methodology, we compared the results obtained from the subjective tests with those obtained from our proposed system. For this, we used the Fuzzy logic toolbox of MATLAB [85] and developed a simulation scenario with our membership function and rules for validation. Test cases with different input values were used for validation. For each test case, we obtained subjective QoE from subjective tests; and, estimated the QoE using our system simulated in MATLAB. Figure 61 represents the comparison between the subjective MOS and estimation MOS. We can see that subjective MOS and estimated MOS have a linear relationship and are highly correlated with correlation coefficient of 0.84. This indicates that the proposed system succeeds in reflecting the user's perception. We should note here that the estimated QoE obtained by our system has a maximum value of 4.51 and a minimum value of 0.523. This is due to the centroid method used for defuzzification in the fuzzy expert system [84]. The results indicate that the proposed system succeeds in reflecting the user's perception. This is also illustrated in Figure 62 that considers the probability distribution of the difference between the subjective scores (QoE) and the estimated QoE scores. We can see that in around 83% of the test cases the score differences were less than 0.5. This estimation accuracy emphasizes the ability of the proposed system to measure the QoE of web services.

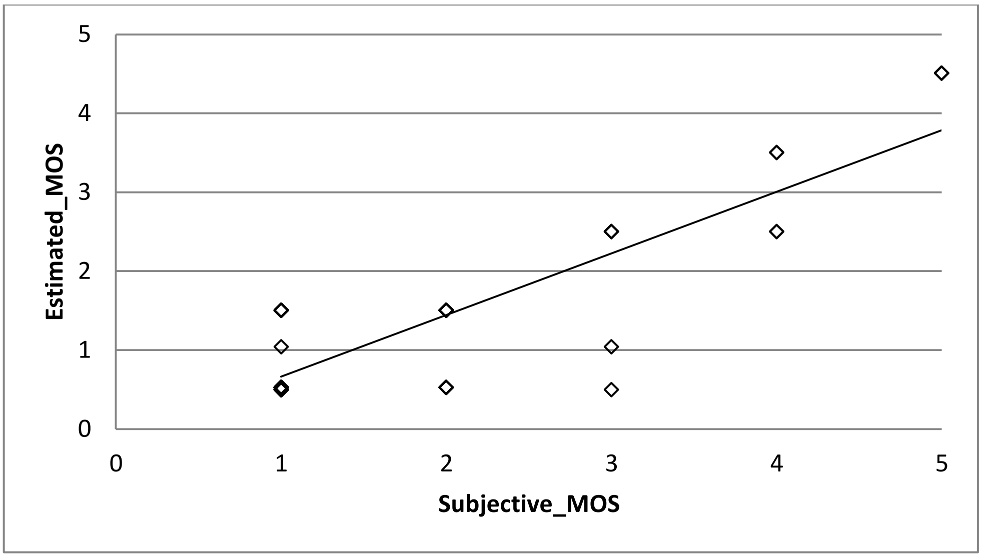

Figure 61: Comparison between subjective and objective QoE

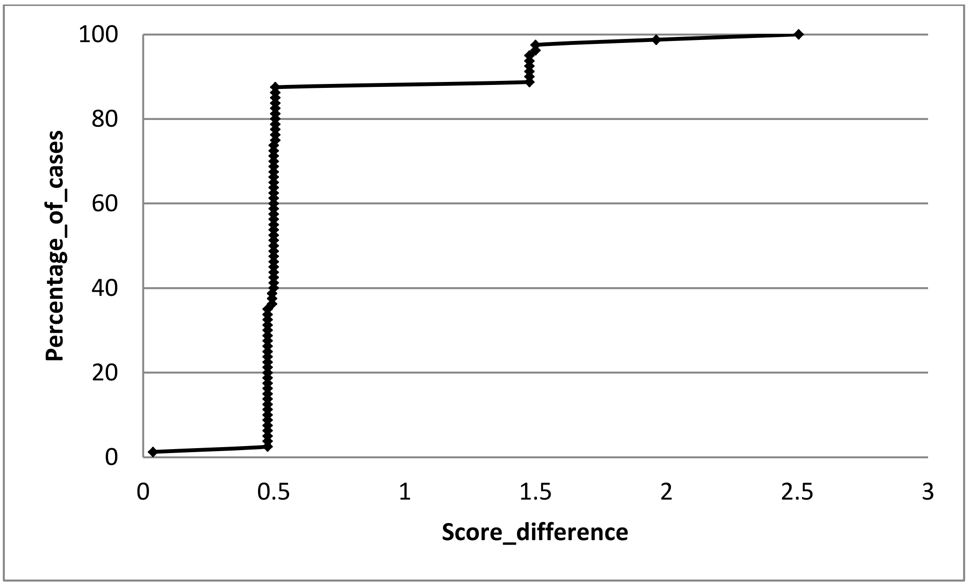

Figure 62: Probability distribution of the subjective and estimated QoE difference

7.6 Conclusion

In this chapter, we proposed a novel method based on fuzzy-rough hybrid expert system for estimating QoE of web services. The estimated QoE can be used as a selection criterion for web services. We have performed a set of subjective tests with real participants in order to correlate web QoS parameters levels with the user perceived quality. The defined fuzzy membership functions were derived from the QoS/QoE correlation using probability distribution functions. We have used a Rough Set Theory to generate estimation inference rules. The proposed methodology has been validated against the results of subjective tests. The validation results have shown that our QoE estimation method is highly correlated to the participants' subjective QoE scores.

Chapter 8

Conclusion and Future work

8.1 Conclusion

In this thesis, our work was mostly dedicated to analyzing the impact of QoS parameters on different internet services and applications. We considered two services as use cases: multimedia VoD and OTT services and web applications. We identified different QoS parameters associated to different layers of the OSI model, which could directly or indirectly impact the service quality. After that, intensive subjective tests were performed to correlate QoS parameters to QoE. We also proposed a QoE estimation system based on different machine learning techniques that use subjective database to train and predict or estimate the QoE. In the following, we summarize the contributions of this thesis work listed in Table 18.

Table 18: Thesis contributions

Application	QoS parameters	Subjective test	Machine Learning Algorithms	Validation
VoD (Chapter 4 and Chapter 5)	Network • Packet loss rate • Burst packet loss, • Jitter	1. Designing laboratory test	Fuzzy expert system	1. Comparing subjective and objective quality score
		2. Mapping QoS/QoE		2. Industrial test bed (VQM vs MMT probe)
		3. Subjective database		3. Benchmarking (comparing MMT probe vs Indra's probe)
	MAC • Aggregation • Number of stations • Bit error rate • Load	1. Designing laboratory test	Random neural network	1. Comparing subjective and objective quality score
		2. Mapping QoS/QoE		
		3. Subjective database		
OTT video (Chapter 6)	Application • Freezing • Bit rate switching • Number of	1. Designing web base application for crowdsourcing		1. Comparing crowdsourcing and laboratory quality score
		2. Pairwise comparison ,		

Application	QoS parameters	Subjective test	Machine Learning Algorithms	Validation
	freezes	Bradley-Terry (BT) model		
		3. Subjective database		
Web service (Chapter 7)	Application • Execution time • Availability • Reliability	1. Designing laboratory test	Fuzzy rough expert system	1. Comparing subjective and objective quality score
		2. Mapping QoS/QoE		
		3. Subjective database		

8.1.1 QoE estimation of VoD services

We analyzed the impact of network conditions on the video QoE and proposed a QoE estimation system for VoD services. For this, we performed a lab based subjective test and used fuzzy expert system as a machine learning algorithm for mapping QoS to QoE. The QoE estimation system extracts the relevant QoS parameters i.e. packet loss ratio, packet loss burstiness and jitter from the network and estimates the video QoE. To validate our QoE estimation system we performed three sets of experiments.

- In the first set, we compared the results of subjective tests to the estimated scores from our proposed QoE estimation system. The comparison results were highly correlated with correlation coefficient of 0.95 and our estimated scores followed the subjective MOS scores.
- In the second set, we implemented our QoE estimation system in an industrial test bed and compared its results with VQM system. VQM considers media parameters for the QoE estimation while our estimation system only considers network parameters. Despite this difference, the results obtained from both systems where highly correlated proving that QoE can be effectively estimated by considering network level parameters.
- In the third set, we performed benchmarking analysis where we compared the results obtained from our MMT probe with the Indra's probe for QoE estimation in the real IPTV test bed. The results show that both probes behave in similar way and their MOS score are highly correlated with correlation index of 0.86.

These experiments validate our system and demonstrate the effectiveness of our QoE estimation system.

After considering network level parameters, we were interested in measuring the MAC level disruptions on the video delivery in 802.11n wireless network. An intensive lab based subjective tests were performed to correlated MAC layer parameters with video QoE. We considered different MAC layer parameters like aggregation, Bit error rate, Queue length, retransmission limit and number of competing stations and analyzed their effect on video QoE. The experiment results suggest that careful parameterization (tuning) of these MAC layer parameters can increase the overall QoE of video services. Moreover, we proposed a QoE estimation system based on random neural networks that learns the relationship between MAC layer parameters and video QoE from the subjective test results. We validated our QoE estimated system by comparing the subjective test results with the results from our proposed QoE estimation system, the results showed high correlation with correlation coefficient of 0.89.

In the above contributions, video QoE was estimated based on the network and MAC layer parameters. The main objective of this work was to provide a QoE estimation tool that can objectively predict the subjective mean opinion of users with reasonable accuracy from the network perspective. The proposed methodology provided higher accuracy, however; the coding and media parameters are equally important and needs to be considered. For example, if a media packet is lost, it is necessary to identify whether I, P or B frame is lost. Since, the impact of I frame loss is more important than P or B frame loss in visual quality.

8.1.2 QoE estimation of OTT services

We analyze the effect of different adaptive scenarios (freezing and bitrate switching) on adaptive bitrate video streaming for OTT video services. A crowdsourcing platform was used for performing the subjective tests. We developed a novel web based subjective test application that can be easily integrated with crowdsourcing platform to perform the test. We compared the results obtained from crowdsourcing test with those obtained from laboratory tests. The comparison results were correlated satisfactorily with correlation coefficient of 0.75, which demonstrates the efficiency of oou platform. Moreover, the subjective test revealed that users dislike buffering events and it seems that frequency is more important than the total duration of these events. However, if the bitrate is high enough and the frequency of the buffering events is low enough, then it seems to a viable alternative to decreasing the bitrate of the video or having a constant low bitrate.

8.1.3 QoE estimation of Web services

We proposed an objective QoE estimation system for web service selection using web QoS parameters. The selected web QoS parameters were execution time, reliability and availability. A subjective test was performed in laboratory using a web application to correlate web QoS parameters with web QoE. The subjective dataset obtained from the subjective test was used to train a novel machine learning algorithm i.e. fuzzy rough hybrid expert system to estimate the web QoE objectively. To validate our system we compared the results of the subjective test with our proposed objective web QoE system. The results showed high correlation between the subjective and objective scores with correlation coefficient 0.84, which validates our proposed system.

8.2 Future Work

This thesis has several contributions and has provided a framework for monitoring and estimating the QoE of different application and services. At the same time, it unlocks some research issues and future directions. In the following, we remark several open issues that can be explored in the future as a continuation of our work.

We will like to extend our QoE estimation framework to other services such as gaming, OTT video etc. We will also investigate the usage of the feedback of our system to implement corrective actions on network and application levels to recover the QoE to satisfactory levels. Furthermore, we will incorporate feature selection methodologies that could automatically select the most impacting QoS parameters. We also plan to make technical analysis of different methodologies proposed in our system in terms of accuracy, processing speed, computational overhead, etc. and compare with other solutions available in the market.

We will extend our VoD QoE estimation system to enrich the model with relevant media and codec parameters and measure the QoE estimation improvement we get. Special care will be taken to measure the computational overhead as we target real-time estimation of video QoE over high speed data links (10+ Gbps).

We will like to exploit the data set obtained from crowdsourcing based subjective test to objectively estimate the QoE for adaptive bitrate streaming. It can be used to train machine learning algorithms for the QoE estimation, as done in the previous contributions. Additionally, we will also investigate on how to adapt our estimation tool in different streaming implementations such as Dynamic Adaptive Streaming over HTTP (MPEG-DASH), Apple HTTP live streaming (HLS), Microsoft Smooth Streaming, etc.

We will like to use web QoE as a metric for web service composition. We also plan to consider the contextual parameters for QoE estimation, as well as implement the solution in different web service frameworks.

References

[1] ITU-T, “Quality of experience requirements for IPTV services,” ITU-T Recommendation G.1080, 2008.

[2] K. Brunnström, S. A. Beker, K. De Moor, A. Dooms, S. Egger, M. N. Garcia,... & A. Zgank, “Qualinet white paper on definitions of quality of experience”, 2013.

[3] Cisco, “Cisco Visual Networking Index: Forecast and Methodology, 2013-2018”, Whitepaper, 2014.

[4] Cisco, “Cisco VNI Service Adoption Forecast, 2013–2018”, Whitepaper, 2014.

[5] J. Pokhrel, F. Lalanne, A. Cavalli, and W. Mallouli, “QoE estimation for web service selection using a Fuzzy-Rough hybrid expert system”, in the 28th IEEE International Conference on Advanced Information Networking and Applications (AINA-2014) Victoria, Canada, May 13-16, 2014.

[6] N. Kushik, N. Yevtushenko, A. Cavalli, W. Mallouli, and J. Pokhrel, “Evaluating web service QoE by learning logic networks”, 10th International Conference on Web Information Systems and Technologies (WEBIST). (pp. 3–5). Barcelona, Spain, 2014.

[7] J. Pokhrel, B. Wehbi, A. Morais, A. Cavalli, and E. Allilaire, “Estimation of QoE of video traffic using a fuzzy expert system,” in10th annual IEEE Consumer Communications & Network-ing Conference (CCNC’13), Las Vegas, Nevada,2013.

[8] J. Pokhrel, I. Paudel, B. Wehbi, A. R. Cavalli and B. Jouaber, “Performance Evaluation of Video transmission over 802.11n wireless network: A MAC layer perspective”, to appear in SaCoNeT - IWQM 2014, June 18-20, 2014, Vilanova I la Geltrù, Spain.

[9] I. Paudel, J. Pokhrel, B. Wehbi, A. Cavalli and B. Jouaber, “Estimation of video QoE from MAC parameters in wireless network: A Random Neural Network approach”, ISCIT, 2014, IEEE, Korea

[10] M. shahid, J. Søgaard, J. Pokhrel, K. Brunnström, K. Wang, S. Tavakol, and N. Garcia “Crowdsourcing based subjective quality assessment of adaptive video streaming”, to appear in IEEE QoMEX 2014, Singapore (presented).

[11] B. S. de Lima and N. F. Ebecken, “A comparison of models for uncertainty analysis by the finite element method,” Finite Elements in Analysis and Design, vol. 34, no. 2, pp. 211 – 232, 2000.

[12] M.B. Anoop, K. Balaji Rao, S. Gopalakrishnan, “Conversion of probabilistic information into fuzzy sets for engineering decision analysis,” Computers & Structures, Volume 84, Issues 3–4, January 2006, Pages 141-155, ISSN 0045-7949.

[13] Z. Pawlak, “Rough set theory and its applications,” Journal of Telecommunications and Information Technology, 2001.

[14] M. van der Schaar, and P. A. Chou, “Multimedia Networking and Communication: Principles and Challenges. Multimedia over IP and Wireless Networks: Compression, Networking, and Systems”, Academic press, 2011.

[15] F. Curbera, M. Duftler, R. Khalaf, W. Nagy, N. Mukhi, and S. Weer-awarana, “Unraveling the web services web: an introduction to soap,wsdl, and uddi,” IEEE Internet Computing, vol. 6, pp. 88–93, 2002.

[16] F. Lalanne, A. Cavalli, and S. Maag, “Quality of experience as aselection criterion for web services,” Signal Image Technology andInternet Based Systems (SITIS), 2012 Eighth International Conference on, pp. 519–526, 2012.

[17] S. Lohr, “For impatient web users, an eye blink is just too long to wait,” New York Times, 2012.

[18] Nokia, “Quality of experience of mobile services Can it be measured and improved?” Nokia white paper, 2004

[19] V. Mehta, and D. N. Gupta, “Performance Analysis of QoS Parameters for Wimax Networks,” International Journal of Engineering and Innovative Technology (IJEIT) Volume, 1, 105-110, 2012.

[20] R. A. Talwalkar and M. Ilyas, "Analysis of Quality of Service (QoS) in WiMAX networks," Networks, 2008. ICON 2008. 16th IEEE International Conference on , vol., no., pp.1,8, 12-14 Dec. 2008.

[21] ITU-T, “Mean opinion Score (MOS) terminology”, ITU-T Recommendation P.800.1, 2006.

[22] Y. C. Yen, C. Y. Chu, S. L. Yeh, H. H. Chu, and P. Huang, “Lab experiment vs. crowdsourcing: a comparative user study on Skype call quality,” In Proceedings of the 9th Asian Internet Engineering Conference (pp. 65-72). ACM, 2013.

[23] T. Hoßfeld, C. Keimel, M. Hirth, B. Gardlo, J. Habigt, K. Diepold, and P. Tran-Gia, “ Best practices for QoE crowdtesting: QoE assessment with crowdsourcing,” In Multimedia, IEEE Transactions on, 2013.

[24] B. Gardlo, S. Egger, M. Seufert, and R. Schatz, "Crowdsourcing 2.0: Enhancing execution speed and reliability of web-based QoE testing," In IEEE International Conference on Communications (ICC 2014), Sydney, Australia, 2014.

[25] J. Li, M. Barkowsky, and P. Le Callet, "Subjective assessment methodology for preference of experience in 3DTV," In IEEE IVMSP Workshop, 2013.

[26] C. Keimel, J. Habigt, C. Horch, and K. Diepold, "Qualitycrowd—a framework for crowd-based quality evaluation," In Picture Coding Symposium (PCS), pp. 245-248, IEEE, Krakow, 2012.

[27] R. Schapire, "Machine Learning algorithms for classification", Princeton University, slides, 2006.

[28] Motorola. The Evolution of OTT video. White paper, 2012.

[29] Cisco. Cisco Visual Networking Index: Forecast and Methodology 2010–2015. Whitepaper, 2011.

[30] Accenture. Hearts, Minds and Wallets, Accenture Video-Over-Internet. Accenture, Consumer Survey 2012, 2012.

[31] H. G. Msakni, and H. Youssef, " Impact of non-QoS related parameters on video QoE" 4th International Workshop on Perceptual Quality of Systems, Vienna, Australia, 2013.

[32] A. Khan, L. Sun, J. Fajardo, I. Taboada, F. Liberal, and E. Ifeachor, " Impact of end devices on subjective video quality assessment for qcif video sequences," In Quality of Multimedia Experience (QoMEX), 2011 Third International Workshop on (pp. 177-182). IEEE, 2011.

[33] R. Bernhaupt, M. Obrist, and M. Tscheligi, "Usability and usage of iTV services: lessons learned in an Austrian field trial." Computers in Entertainment (CIE), 5(2), 2007.

[34] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, RFC 3550. RTP: a transport protocol for real-time applications, 7, 2003.

[35] Avermedia, HTTP/TCP vs. RTP/UDP. Retrieved August 5, 2014, from http://solutions.avermedia.com/?q=en/support/knowledge-base/internet-basic/http-tcp-vs-rtp-udp, 2013.

[36] G. McGath, Basics of streaming protocol. Retrieved August 5, 2014, from http://www.garymcgath.com/streamingprotocols.html, 2009.

[37] A. Biernacki, F. Metzger, and K. Tutschku, "On the influence of network impairments on YouTube video streaming," Journal of Telecommunications and Information Technology (JTIT), 2012 (03). Warsaw, Poland.

[38] M. Seyedebrahimi, C. Bailey and X. H. Peng, "Model and Performance of a No-Reference Quality Assessment Metric for Video Streaming," IEEE Trans. Circuits Syst. Video Techn. 23(12): 2034-2043 (2013).

[39] RGB Networks, Comparing Adaptive HTTP Streaming Technologies. Retrieved August 5, 2014, from http://www.rgbnetworks.com/pdfs/RGB_Adaptive_HTTP_Streaming_Comparison_1211-01.pdf, 2011.

[40] F. Sonnati, New buffering strategies in Flash Player 9 and Flash Media Server 3. Retrieved August 5, 2014, from http://www.adobe.com/devnet/adobe-media-server/articles/fms_buffering_strategies.html.

[41] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, RFC 3550.RTP: a transport protocol for real-time applications, 2003.

[42] G. Rubino, P. Tirilly, and M. Varela, "Evaluating users' satisfaction in packet networks using random neural networks," Artificial Neural Networks–ICANN. (pp. 303-312). Springer Berlin, Heidelberg, 2006.

[43] ITU-T. Methods for objective and subjective assessment of Quality. Recommendation P.800, 1996.

[44] T. H. Falk, Q. Xu, and W. Y. Chan, "Non-Intrusive GMM-Based Speech Quality Measurement," In Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05), (pp. 125-128), 2005.

[45] N. Côté, Integral and Diagnostic Intrusive Prediction of Speech Quality, T-Labs Series in Telecommunication Services, Springer, 2011.

[46] ITU-R, Methodology for the subjective assessment of the quality of television pictures. Recommendation BT.500.13, 2012.

[47] T. Alpert, and J. P. Evain, "Subjective quality evaluation: the SSCQE and DSCQE methodologies," EBU Technical Review, 12-20, 1997.

[48] ITU-R, Methodology for the subjective assessment of video quality in multimedia applications. Recommendation BT.1788, 2007.

[49] ITU-T, Subjective video quality assessment methods for multimedia applications. Recommendation P.910, 1999.

[50] ITU-T, User requirements for objective perceptual video quality measurements in digital cable television. Recommendation J.143, 2000.

[51] NTT. Video quality assessment methods. Retrieved August 5, 2014, from http://www.ntt.co.jp/qos/qoe/eng/technology/visual/index.html.

[52] J. Joskowicz, and J. C. Lopez Ardao, "A general parametric model for perceptual video quality estimation," In Communications Quality and Reliability (CQR), 2010 IEEE International Workshop Technical Committee on, (pp. 1-6), Vancouver, Canada, 2010.

[53] A. Takahashi, D. Hands, V. Barriac, "Standardization activities in the ITU for a QoE assessment of IPTV," Communications Magazine, IEEE , vol.46, no.2, pp.78,84, February 2008.

[54] M. S. Mushtaq, B. Augustin, and A. Mellouk, A. "Empirical study based on machine learning approach to assess the QoS/QoE correlation," 17th European Conference on Networks and Optical Communications (NOC). IEEE, (pp. 1-7). Vilanova i la Geltru, Spain, 2012.

[55] T. Wang, A. Pervez, and H. Zou, "VQM-based QoS/QoE mapping for streaming video," 3rd IEEE International Conference on Broadband Network and Multimedia Technology (IC-BNMT). IEEE, pp. 807-812, Beijing, China, 2010.

[56] M. Wijnants, S. Agten, P. Quax, and W. Lamotte, "Investigating the relationship between QoS and QoE in a mixed desktop/handheld gaming setting," Proceedings of the 5th international student workshop on Emerging networking experiments and technologies. ACM, (pp. 29-30). Rome, Italy, 2009.

[57] T. M. Mitchell, Machine learning. 1997. Burr Ridge, IL: McGraw Hill, 45, 1997.

[58] E. J. McCluskey, Introduction to the theory of switching circuits. McGraw-Hill, 1965.

[59] E. Al-Masri, and Q. H. Mahmoud, "Discovering the best web service: A neural network-based solution," In Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on. pp. 4250-4255. IEEE. San Antonio, TX, 2009.

[60] L. A. Zadeh, Fuzzy sets. Information and Control, 8(3), 338–353,1965.

[61] V. Novák, I. Perfilieva, and J. Močkoř, Mathematical principles of fuzzy logic. Springer, 1999.

[62] A. Baghel, and T. Sharma, "Survey on Fuzzy Expert System," International Journal of Emerging Technology and Advanced Engineering, 230-233, 2013.

[63] H. Attaullah, F. Iqbal, and M. Y. Javed, "Intelligent vertical handover decision model to improve QoS," Digital Information Management (ICDIM). (pp. 119-124). London, UK, 2008.

[64] T. J. Ross, Fuzzy Logic with Engineering Applications. McGraw-Hill Inc, 1995.

[65] N. Kushik, J. Pokhrel, N. Yevtushenko, A. Cavalli, and W. Mallouli, "QoE prediction for multimedia services: Comparing fuzzy and logic network approaches," To appear in International Journal of Organizational and Collective Intelligence (IJOCI), 2014.

[66] S. Winkler, "Video Quality Measurement Standards –Current Status and Trends," In IEEE Information, Communications and Signal Processing, Macau, 2009.

[67] VQEG. Video Quality Expert Group - Motivation, Objectives and Rules. Retrieved August 5, 2014, from http://www.its.bldrdoc.gov/vqeg/about-vqeg.aspx.

[68] ITU, ITU Telecommunication Standardization Sector. Retrieved August 5, 2014, from http://www.itu.int/en/about/Pages/history.aspx.

[69] ATIS-IIF. The IPTV Interoperability Forum. Retrieved August 5, 2014, from http://www.atis.org/iif/.

[70] S. Baraković, and L. Skorin-Kapov, "Survey and challenges of QoE management issues in wireless networks," Journal of Computer Networks and Communications, 2013.

[71] T. Hobfeld, R. Schatz, M. Varela, and C. Timmerer," Challenges of QoE management for cloud applications,"Communications Magazine, IEEE, 50(4), 28-36, 2012.

[72] W. Fenner, RFC 2236: Internet Group Management Protocol, version 2, November 1997.

[73] "What is IPTV: Unicast vs multicast". Retrieved September 2014, from http://iptvpavilion.com/features/iptv-unicast-multicast-0319/.

[74] S. Winkler and R. Campos, "Video quality evaluation for Internet streaming applications," Human Vision and Electronic Imaging, vol. 3, pp. 104–115, 2003.

[75] D. Wijesekera, I. Srivastava, A. Nerode, and M. Foresti,"Experimental evaluation of loss perception in continuous media," Mulimedia Systems, vol. 7, pp. 486-499, 1999.

[76] O. Issa, F. Speranza, W. Li and H. Liu; , “Estimation of time varying QoE for high definition IPTV distribution,”Consumer Communications and Networking Conference (CCNC), 2012 IEEE, vol., no., pp.326-330, 14-17 Jan. 2012.

[77] ITU, Quality of experience requirements for IPTV services. ITU-T Recommendation G.1080, 2008.

[78] Technical Report TR-126, Triple-play Services Quality of Experience (QoE) Requirements. DSL Forum, Tech. Rep., 2006, http://www. broadbandforum. org.

[79] N. Staelens, G. Van Wallendael, K. Crombecq, N. Vercammen, J. De Cock, B. Vermeulen, R. Van De Walle, T. Dhaene, and P. Demeester, "No-Reference Bitstream-Based Visual Quality Impairment Detection for High Definition H.264/AVC Encoded Video Sequences," Broadcasting, IEEE Transactions on , vol.58, no.2, pp.187,199, June 2012.

[80] Agilent, “Whitepaper: IPTV QoE: Understanding and interpreting MDI values,” Agilent technologies, 2008.

[81] Y. Kang, H. Chen, and L. Xie, "An artificial-neural-network-based QoE estimation model for Video streaming over wireless networks," Communications in China (ICCC), 2013 IEEE/CIC International Conference on , vol., no., pp.264,269, 12-14 Aug. 2013.

[82] W. Cherif, A. Ksentini, D. Negru, M. Sidibe, "A_PSQA: Efficient real-time video streaming QoE tool in a future media internet context," Multimedia and Expo (ICME), 2011 IEEE International Conference on , vol., no., pp.1,6, 11-15, July 2011

[83] V. A. Machado, C. N. Silva, R. S. Oliveira, AM. Melo, M. Silva, C.R.L. Frances, J. C W A Costa, N.L. Vijaykumar, and C.M. Hirata "A new proposal to provide estimation of QoS and QoE over WiMAX networks: An approach based on computational intelligence and discrete-event simulation," Communications (LATINCOM), 2011 IEEE Latin-American Conference on , vol., no., pp.1,6, 24-26 Oct. 2011.

[84] M. Negnevitsky, “Artificial Intelligence, A Guide to Intelligent System,” Addision-Weslay, Pearson Education, 2002.

[85] Fuzzy Logic Toolbox, “MATLAB tools for designing, and simulating systems based on fuzzy logic”, [Online], available at http://www.mathworks.fr/products/fuzzy-logic/

[86] “Montimage monitoring tool (MMT)”, [online] available at http://www.montimage.com/products.html

[87] A. Cuadra-Sanchez, “D2.2. Definition of requirements of the management systems to keep up with QoE expectations based on QoS and traffic monitoring,” IPNQSIS Project Report, [Online], available at http://ipnqsis.org/index.php?title=Deliverables.

[88] E. Perahia, and R. Stacey. "Next generation wireless LANs." Throughput, robustness and reliability in 802.11 n, Cambridge, 2008.

[89] H. Zheng, G. Chen, and L. Yu. "Video transmission over IEEE 802.11 n WLAN with adaptive aggregation scheme." Broadband Multimedia Systems and Broadcasting (BMSB), 2010 IEEE International Symposium on. IEEE, 2010.

[90] I. Paudel, F. Outay, and B. Jouaber. "QoS-HAN: QoS provisioning in home automated networks over IEEE 802.11 n." Computers and Communications (ISCC), 2012 IEEE Symposium on. IEEE, 2012.

[91] C.-Y. Wang, H.-Y. Wei, “IEEE 802.11n MAC enhancement and performance evaluation. Mobile Networks and Applications”, 14(6), 760–771.

[92] T. Suzuki, T. Kutsuna, and S. Tasaka, “Qoe estimation from mac-level QoS in audio-video transmission with ieee 802.11e edca,” in Proc. IEEE PIMRC (Workshop), Cannes, France, 2008, pp. 1–6.

[93] G. Barreira, and J. Ascenso. "Impact of the IEEE 802.11 n frame aggregation mechanisms on video streaming quality." Software, Telecommunications and Computer Networks (SoftCOM), 2012 20th International Conference on. IEEE, 2012.

[94] P. Počta, R. Hudec, and M. Bojmír. "Performance evaluation of video transmission over IEEE 802.11 b WLANs from the video quality perspective." Mediterranean Journal of Computers and Networks.–UK 5.2 (2009): 53-58.

[95] J. Klaue, B. Rathke, and A. Wolisz. "Evalvid–A framework for video transmission and quality evaluation." Computer Performance Evaluation. Modelling Techniques and Tools. Springer Berlin Heidelberg, 2003. 255-272.

[96] M. Nezveda, S. Buchinger, W. Robitza, E. Hotop, P. Hummelbrunner,and H. Hlavacs, “Test persons for subjective video quality testing:Experts or non-experts?” in QoEMCS workshop at the EuroITV-8th European Conference on Interactive TV, 2010.

[97] “qoe-rnn”, http://code.google.com/p/qoe-rnn/

[98] http://en.wikipedia.org/wiki/Over-the-top_content

[99] J. Jacobs, Five Major Challenges for Telecommunications Providers, June 7, 2013, Retrieved October 14, 2014 from http://www.telco-ott.com/opinions/2013/06/five-major-challenges-for-telecommunications-providers/.

[100] R. A. Bradley and M. E. Terry, "Rank analysis of incomplete block designs: I. The method of paired comparisons,"Biometrika, vol. 39, p. 324U″ 345, 1952.

[101] Tavakoli et al., "Subjective quality assessment of an adaptive video streaming model," in IS&T/SPIE Electronic Imaging. Int. Soc. For Optics and Photonics, 2014.

[102] K.-T. Chen, C.-J. Chang, C.-C. Wu, Y.-C. Chang, and C.-L.Lei, "Quadrant of euphoria: a crowdsourcing platform for QoE assessment," Network, IEEE, vol. 24, no. 2, pp. 28–35, Mar. 2010.

[103] F. Ribeiro, D. Florencio, C. Zhang, and M. Seltzer, "Crowdmos:An approach for crowdsourcing mean opinion score studies,"in Acoustics, Speech and Signal Processing (ICASSP),2011 IEEE International Conference on, May 2011, pp. 2416–2419.

[104] F. Wickelmaier and C. Schmid, "A matlab function to estimate choice model parameters from paired-comparison data,"Behavior Research Methods, Instruments, & Computers, vol. 36, no. 1, pp. 29–40, 2004.

[105] R. Mok, E. Chan, and R. Chang, "Measuring the quality of experience of http video streaming," in IFIP/IEEE Int. Symposium on Integrated Network Management (IM), 2011, pp. 485–492.

[106] S. Van Kester, T. Xiao, R. Kooij, K. Brunnstr¨om, and O. Ahmed, "Estimating the impact of single and multiple freezes on video quality," in IS&T/SPIE Electronic Imaging. Int. Soc. for Optics and Photonics, 2011.

[107] F. Curbera, M. Duftler, R. Khalaf, W. Nagy, N. Mukhi, and S. Weer-awarana, "Unraveling the web services web: an introduction to soap, wsdl, and uddi," IEEE Internet Computing, vol. 6, pp. 88–93, 2002.

[108] F. Lalanne, A. Cavalli, and S. Maag, "Quality of experience as a selection criterion for web services," Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference on, pp. 519–526, 2012.

[109] Y. Badr, A. Abraham, F. Biennier, and C. Grosan, "Enhancing web service selection by user preferences of non-functional features," Next Generation Web Services Practices, 2008. NWESP '08. 4th International Conference on, pp. 60–65, 2008.

[110] P. Wang, K.-M. Chao, C.-C. Lo, C.-L. Huang, and Y. Li, “A fuzzy model for selection of qos-aware web services,” in e-Business Engineering, 2006. ICEBE ’06. IEEE International Conference on, 2006, pp. 585–593.

[111] S. Tiwari and S. Kaushik, “A non functional properties based web service recommender system,” in Computational Intelligence and Software Engineering (CiSE), 2010 International Conference on, 2010, pp. 1–4.

[112] S. Lohr, “For impatient web users, an eye blink is just too long to wait,” New York Times, 2012.

[113] Nokia, “Quality of experience of mobile services Can it be measured and improved?” Nokia white paper, 2004.

[114] M. Pinson and S. Wolf, “A new standardized method for objectively measuring video quality,” Broadcasting, IEEE Transactions on, vol. 50, no. 3, pp. 312–322, 2004.

[115] D. Rossi, M. Mellia, and C. Casetti, “User patience and the web: a hands-on investigation,” in Global Telecommunications Conference,2003. GLOBECOM ’03. IEEE, vol. 7, 2003, pp. 4163–4168 vol.7.

[116] N. Vicari and S. Kohler, “Measuring internet user traffic behavior dependent on access speed,” In proceedings of the 13th ITC specialist seminar on IP Traffic Measurement, Modeling and Management, 2000.

[117] S. Egger, T. Hossfeld, R. Schatz, and M. Fiedler, “Waiting times in quality of experience for web based services,” in Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on, 2012, pp. 86–96.

[118] A. Bouch, A. Kuchinsky, and N. Bhatti, “Quality is in the eye of the beholder: meeting users’ requirements for internet quality of service,” in Proceedings of the SIGCHI conference on Human Factors in Computing Systems, ser. CHI ’00. New York, NY, USA: ACM, 2000, pp. 297–304.

[119] A. Padovitz, S. Krishnaswamy, and S. W. Loke, “Towards efficients election of web services,” Autonomous Agents and Multi-Agent Systems(AAMAS 2003), 2003.

[120] N. Keskes, A. Lehireche, and A. Rahmoun, “Web services selection based on context ontology and quality of services,” International Arab Journal of e-Technology, vol. 1.

[121] Y. Liu, A. Ngu, and L. Zeng, “Qos computation and policing in dynamic web service selection,” Proceedings of the 13th international World WideWeb conference on Alternate track papers and posters, 2004.

[122] I. Egambaram, G. Vadivelou, and S. P. Sivasubramanian, “Qos based web service selection,” The International Conference on Computing, Communications and Information Technology Applications (CCITA-2010), 2010.

[123] J. Hua, “Study on the application of rough sets theory in machine learning,” in Intelligent Information Technology Application, 2008. IITA’08. Second International Symposium on, vol. 1, 2008, pp. 192–196.

[124] K. Laghari, 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.

[125] “Rosetta- a rough set took kit.” [Online]. Available: http://www.lcb.uu.se/tools/rosetta

[126] D. S. Johnson, “Approximation algorithms for combinatorial problems,” Journal of Computer and System Sciences, vol. 9, no. 3, pp. 256 – 278,1974.

[127] ITU-T, “Objective perceptual video quality measurement techniques for digital cable,” ITU-T Recommendation J.144,Geneva, Switzerland, March 2003.

[128] Y. Wang, T. Jiang, S. Ma and W. Gao, "Spatio-temporal ssim index for video quality assessment," Visual Communications and Image Processing (VCIP), 2012, IEEE , vol., no., pp.1,6, 27-30 Nov. 2012.

[129] ITU-T, Objective perceptual multimedia video quality measurement in the presence of a full reference,’ ITU Recommendation J.247, 2008.

[130] A. Gray and S. MacDonell, “A Comparison of Techniques for Developing PredictiveModels of Software Metrics,” Information and Software Technology. Vol. 39, p.425-437, 1997.

[131] A. Adeli, and M. Neshat, “A fuzzy expert system for heart disease diagnosis,” In Proceedings of International Multi Conference of Engineers and Computer Scientist. Hong Kong (Vol. 1), 2010.

[132] E. W. T. Ngai, and F. K. T. Wat, “Design and development of a fuzzy expert system for hotel selection,” Omega. 31(4), 275-286, 2003.

[133] T. Wang, and H. Lee, “Constructing a Fuzzy Decision Tree by Integrating Fuzzy Sets and Entropy”, Proceedings of the 5th WSEAS international conference on Applied computer science (ACOS'06), 2006.

[134] M. Fiedler, T. Hossfeld, and P. Tran-Gia, (2010) “A generic quantitative relationship between quality of experience and quality of service”, IEEE Network, vol. 24, no. 2, pp. 36–41.

[135] M. Siller and J. Woods, (2003) “Improving Quality of Experience for Multimedia Services by QoS arbitration on a QoE Framework,” in IEEE Packet Video, Nantes.

[136] F. Agboma &A. Liotta, (2008) “QoE-aware QoS management,” in Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia, pp. 111-116.

[137] V. Menkovski, A. Oredope, A. Liotta, &A. Cuadra Sánchez, “Predicting Quality of Experience in Multimedia Streaming,” in Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia, pp. 52-59, 2009.

[138] V. Machado, C. Oliveira , A. Marcelino, S. Carlos, N. Vijaykumar, and C. Hirata, “A New Proposal to Provide Estimation of QoS and QoE over WiMAX Networks”, 978-1-4673-0279-1, IEEE , 2011.

[139] H. Du, C. Guo, Y. Liu and Y. Liu,“Research on Relationships between QoE and QoS based on BP Neural Network”, In: Proceedings of IC-NIDC 2009, pp. 312-315, 2009.

[140] H. J. Kim and S. G. Choi, (2010) “A study on a QoS/QoE correlation model for QoE evaluation on IPTV service,” in Proceedings of the 12th International Conference on Advanced Communication Technology (ICACT ’10), pp. 1377–1382.

[141] A. Khan, L. Sun, E. Ifeachor, J. 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, 2010.

[142] T. Malinovski, T. Vasileva-Stojanovska, and V.Trajkovik, “Quality-of-Experience Perception for VideoStreaming Services: Preliminary Subjective andObjective Results, International Journal of Research and Reviews in Next Generation Networks (IJRRNGN) Vol. 1., 2011.

[143] 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 Pers Commun, 2012.

[144] P. Ramos, J. Salmerón, R. Leal, and F. Vidal, “Estimating Perceived Video Quality from Objective Parameters in Video over IP Services”, Copyright (c) IARIA, ISBN: 978-1-61208-193-9, 2012.

[145] H. Koumaras, C-H. Lin, C-K. Shieh, and A. Kourtis, “A Framework for End-to-End Video Quality Prediction of MPEG Video”, J. Vis. Commun. Image R, 2009.

[146] P. Frank and J. Incera, “A neural network based test bed for evaluating the quality of video streams in IP networks”, IEEE, Proceedings of the Electronics, Robotics and Automotive Mechanics Conference (CERMA'06), 2006.

[147] P. Calyam, P. Chandrasekaran, G. Trueb, N. Howes, R. Ramnath, D. Yu, Y. Liu, L. Xiong, and D. Yang, (2012) “Multi-ResolutionMultimedia QoEModels for IPTV Applications”, Volume 2012, Article ID 904072, 2012.

[148] M. Elkotob, D. Granlund, K. Andersson and C. Åhlund, “Multimedia QoE optimized management using prediction and statistical learning”. LCN: 324-327, 2010.

[149] M.S. Mushtaq, B. Augustin, and A. Mellouk, "Empirical study based on machine learning approach to assess the QoS/QoE correlation", Networks and Optical Communications (NOC), 17th European Conference on, on page(s): 1 – 7, 2012.

[150] T. Hoßfeld, R. Schatz, M. Seufert, M. Hirth, T. Zinner, &P. Tran-Gia, “Quantification of YouTube QoE via Crowdsourcing,” in IEEE International Workshop on Multimedia Quality of Experience - Modeling, Evaluation, and Directions (MQoE), Dana Point, CA, USA, 2011.

[151] M. Alreshoodi, and J. Woods, “Survey on QoE\ QoS Correlation Models For Multimedia Services.”, International Journal of Distributed and Parallel Systems (IJDPS) 2013.

[152] Screen test ,Retrieved October 14, 2014, from https://github.com/St1c/screentest/

[153] Microworkers, Retrieved October 14, 2014, from https://microworkers.com/

[154] Mechanical Turk, Retrieved October 14, 2014, from https://www.mturk.com/mturk/welcome

[155] Clickworker, Retrieved October 14, 2014, from http://www.clickworker.com/en/

[156] Alcatel lucent, Video shakes up the IP edge: Extended – A Bell Labs study on rising video demand and its impact on broadband IP networks, whitepaper, 2012

[157] T. C. Thang, H. X. Nguyen, and A. T. Pham, “Perceptual difference evaluation of video alternatives in adaptive streaming,” International Conference on Communications and Electronics, 322–326, 2012.

[158] N. Cranley, and L. Murphy, “Incorporating user perception in adaptive video streaming systems,” Digital Multimedia Perception and Design, Idea Group Inc., 2006

[159] D. Robinson, Y. Jutras, and V. Craciun, “Subjective video quality assessment of HTTP adaptive streaming technologies,” Bell Labs Technical Journal, 5-23, 2012.

[160] L. Yitong, S. Yun, M. Yinian, L. Jing, L. Qi, and Y. Dacheng, "A study on Quality of Experience for adaptive streaming service," Communications Workshops (ICC), 2013 IEEE International Conference on , vol., no., pp.682,686, 9-13 June 2013

[161] M.S. Mushtaq, B. Augustin, and A. Mellouk,, "Crowd-sourcing framework to assess QoE," Communications (ICC), 2014 IEEE International Conference on , vol., no., pp.1705,1710, 10-14 June 2014

[162] CELTIC-Plus of EUREKA ICT Cluster, Project CP07-009, “IPNQSIS”, Retrieved October 14, 2014 http://projects.celtic-initiative.org/ipnqsis/