Quality of experience and video services adaptation

Mamadou Tourad Diallo

Ecole Doctorale EDITE

Thèse présentée pour l'obtention du diplôme de Docteur de Télécom SudParis

Specialité :
Informatique et Télécommunication

Par :

Mamadou Tourad DIALLO

Qualité d'Expérience et Adaptation de services vidéo

Directeur de thèse : **Hossam AFIFI**
Co-directeur de thèse : **Nicolas MARECHAL**

Soutenue le 04/06/2015

Jury

Adlen KSENTINI, Professeur Associé, Univ. de Rennes 1, France Rapporteur
Ken CHEN, Professeur, Univ. de Paris 13, France Rapporteur
Hassnaa MOUSTAFA, Senior Scientist, Intel Corporation, USA Examinateur
Luigi ATZORI, Professeur, Univ. de Cagliari, Italie Examinateur
Laurent RUCKENBUSCH, Team Manager, Orange, France Examinateur
Houda LABIOD, Professeur Associé, Telecom ParisTech, France Examinateur
Hossam AFIFI, Professeur, Télécom SudParis, France Directeur de thèse
Nicolas MARECHAL, Project Leader, Orange, France Co-directeur de thèse

NUMERO DE THESE : 2015TELE0010

Abstract

Video services are experiencing an unprecedented growth for the last few years. According to Cisco's VNI forecasts, it is predicted that 79% of all internet traffic will be video and that the online video will be the most highly adopted among online services, growing from 1.2 billion users in 2013 to 1.9 billion by 2018. As such, market players are fighting to increase Average Revenue per User (ARPU), limit churn and improve their market share. From a marketing point of view, one possible option is to focus on improving end-users' satisfaction, namely Quality of Experience (QoE) for short. QoE is a young research subject with limited considerations regarding contextual information that deserves a deeper understanding and can offer great commercial perspectives.

Semantically speaking, this concept is closely related to the Quality of Service (QoS), even if the former is now associated to the enforcement of purely technical constraints so that to ensure a given level of service expectations. Contrary to this, QoE goes beyond the technical background.

In this PhD thesis, we first provide a technical overview on video services and architectural deployments for IPTV (Internet Protocol TeleVision) and WebTV (Web TeleVision) services. Then, a state-of-the-art about both QoE measurement techniques and Content & Delivery Adaptation is also provided. According to these surveys, two methods can be considered to understand how users interact with services and estimate their QoE. On one hand by monitoring and analyzing the impact of quality metrics on user engagement, in order to understand the effects of technical video metrics (video startup time, average bitrate, buffering ratio) and content popularity on user engagement. The user engagement can be considered as the user's behaviour at any point in time and over time, between a user and the service. On other hand, one can consider subjective approaches such as the Mean Opinion Score (MOS) for evaluating QoE, in which users are required to give their assessment/rating. The MOS presents the exact user perception of the viewed video, which is considered as a better indicator of video quality as it is given by humans.

Our results show that video buffering and content popularity are critical parameters which strongly impacts the end-user's satisfaction and user engagement, while the video startup time appears as less significant. In the third part, we propose to assess QoE in terms of MOS (Mean Opinion Score) through introducing contextual information. We did tests with users to get their feelings while watching video contents under varying conditions (context parameters). A detailed overview and statistical analysis of our study shows the existence of non-trivial parameters impacting MOS (the type of device, the content type for constant video bitrate: football, cartoon etc.). We also propose mathematical models to develop functional relationships between the QoE and the context information which in

turn permits us to estimate the QoE. To assess the performance of our proposal, we compare it with an operational QoE measurement tool. Our results prove that contextual information is an important parameter and one which needs to be taken into account for monitoring and providing an accurate assessment of QoE. Finally, in the last part of this manuscript, we provide general QoE multi-users optimization, that optimizes the network resources by considering the end-user satisfaction in terms of MOS. Our proposals improve the perceived QoE for different video sessions sharing the same local network, while taking QoE fairness among users as a leitmotiv. This approach is validated by simulations and corresponding prototype architecture is proposed. We also propose a utility-based approach in which a global utility function is computed based on different constraints (e.g. target strategies coming from the actors of the delivery chain). In conclusion, the different contributions proposed in this thesis improves the understanding of hidden relationships between quality parameters and user engagement, and about how contextual information may influence the end-user's perceived video quality. Finally, we proved that this work can help improving the network usage, reducing congestion phenomena and in ensuring a level of QoE for connected users.

Résumé

Définitions et motivations du travail de these:

La croissance des vidéos en ligne et la demande croissante des services multimédias et audiovisuels rendent la Qualité d'Expérience (QoE) un facteur déterminant de la réussite ou de l'échec des applications et services. Il est important de comprendre les exigences des usagers en terme de qualité. Ainsi la QoE apparaît comme une mesure de satisfaction des clients d'un service, en fournissant une évaluation de leurs attentes, leurs sentiments, leurs perceptions, l'acceptation par rapport à un service ou une application particulière.

Mesurer la QoE permet aux opérateurs et aux fournisseurs de services à limiter le taux de désabonnement, en augmentent le revenu moyen par utilisateur (ARPU), la part de marché et également veiller à la satisfaction des utilisateurs. Il existe de nombreuses définitions de la QoE. Selon la 3GPP, la QoE indique les mesures de performance en termes du point de vue de l'utilisateur du service. En général, un prestataire de services fixe les exigences de service pour répondre à la QoE attendue par les utilisateurs, qui doit être traduit en paramètres ou mesures que le fournisseur de services peut contrôler ou mesurer.

Pour l'UIT-T (ITU Telecommunication Standardization Sector), la QoE est l'acceptabilité globale d'une application ou d'un service, tel qu'il est perçu subjectivement par l'utilisateur final. Selon l'UIT-T, la QoE prend en compte les informations de bout en bout (client, terminal, réseau, infrastructure de services, etc.) et l'acceptabilité globale peut être influencée par les attentes des utilisateurs et le contexte de l'utilisateur.

L'ETSI (Institut européen des normes de télécommunications) a fourni une définition alternative de la QoE, afin d'étendre la QoE au-delà des mesures subjectives de la perception de l'utilisateur, d'inclure des mesures objectives de processus de communication. La définition d'ETSI prend en compte des paramètres techniques (perte de paquets, délai etc.) et les informations l'utilisateur (son efficacité, sa satisfaction, son envie d'utiliser le service).

Les mesures psychologiques appropriées seront tributaires du contexte de communication. Contrairement aux mesures objectives, les méthodes subjectives sont basées sur l'avis de l'utilisateur (la qualité perçue). Cependant, la QoE, tel qu'elle est définie et mesurée aujourd'hui ne suffit pas pour adapter le contenu et la livraison afin d'améliorer la satisfaction des utilisateurs. Par conséquent, nous définissons dans cette thèse, la notion de la QoE comme une estimation subjective qui reflète le degré de satisfaction de l'utilisateur suivant la définition de l'UIT-T.

En plus de cela, la QoE est impactée par des paramètres contextuels de l'utilisateur et de son environnement (type de terminal utilisé, les caractéristiques

du réseau, type de contenu consommé, les paramètres d'encodage de la vidéo, les préférences des utilisateurs, la localisation des utilisateurs ...).

Ainsi, dans cette thèse, nous proposons des méthodes pour mesurer et analyser la QoE en prenant en compte des informations contextuelles des terminaux, les réseaux et les contenus. Le contexte est un élément fondamental de la communication et elle influence la QoE. Dans la deuxième phase, nous étudierons l'adaptation du contenu et la livraison, qui prennent en compte la mesure de la QoE qui est fonction des informations de contexte. Et enfin, nous définissons dans cette thèse, une nouvelle notion d'adaptation des contenus basée sur la mesure de la QoE.

Contributions de la thèse:

Ce manuscrit présente six principales contributions dans le domaine de la recherche:

- Nous analysons l'impact du temps de démarrage, le bitrate moyen et la popularité du contenu sur l'engagement de l'utilisateur.

- Nous étudions l'impact des informations de contextes sur la QoE par l'expérimentation. Afin d'étudier les effets des paramètres contextuels (type de contenu, le type de terminal, le débit client), nous proposons des tests. Les tests consistent à laisser les utilisateurs à regarder des contenus différents sur des appareils différents et avec des bandes passantes différentes. Après chaque visualisation, les utilisateurs finaux sont invités à donner un avis sur leur satisfaction selon les recommandations de l'UIT-T, soit selon une échelle de qualité de cinq points allant de mauvais (1) à excellent (5). En plus de donner une note globale de perception, les utilisateurs sont aussi invités à fournir une note pour le temps de démarrage de la vidéo et le pourcentage de « buffering », qui correspond au moment où la vidéo est bloquée. Cette dernière note permet de comprendre qui est ce qui a impacté la note globale. Un nombre important de données a été recueilli au cours des expériences.

- Deux modèles mathématiques sont proposés: la première prédit la QoE, fonction du type de terminal utilisé, le type de contenu vidéo et de l'etat de la liaison. La seconde évalue la QoE en fonction du type de terminal utilisé, le type de contenu vidéo et la qualité intrinsèque de la vidéo (profil vidéo).

- Nous proposons MDASH (MOS Dynamic Adaptive streaming sur HTTP), qui améliore la perception de la QoE pour différentes sessions vidéo, partageant le même réseau et qui maximise ainsi la QoE de l'utilisateur qui a la plus faible QoE. La proposition a été validée par des simulations et une architecture a été proposée.

Aperçu de cette thèse:

Ce manuscrit est organisé comme suit:

- Le chapitre 2 présente un aperçu des services vidéo et les architectures de déploiements, en particulier pour l'IPTV (Internet Protocol Television) et des services (WebTV de Web Télévision). En plus de cela, des techniques de streaming vidéo existants sont présentés et une comparaison entre les techniques HTTP et non HTTP techniques adaptatives sont proposées. Un état de l'art des techniques de mesure de QoE, l'adaptation de contenu vidéo et des techniques de livraison de contenus sont également présentés.

- Le chapitre 3 propose de répondre à ces questions: comment l'engagement des utilisateurs varie avec le temps de démarrage, le taux de buffering, le bitrate moyen, la fréquence de buffering et la popularité du contenu.

- Le chapitre 4 met l'accent sur l'effet des paramètres de contextes (type de contenu, le type de terminal, le débit moyen) sur le Mean Opinion Score (MOS). Dans ce chapitre, nous décrivons la plate-forme d'expérimentation développé, les conditions des tests adoptés, analysons également les résultats obtenus et proposons des modèles mathématiques pour évaluer la fonction MOS en fonction des informations de contextes.

- Chapitre 5 propose une approche d'adaptation de contenu vidéo appelé le MDASH (MOS Dynamic Adaptive streaming sur HTTP), ce qui améliore la perception de la QoE pour les sessions de vidéo différents partageant le même réseau et améliore ainsi la QoE de cet utilisateur parmi les autres qui a la plus faible QoE. Nous validons la proposition par des simulations et une architecture correspondante est proposée.

- Enfin, le Chapitre 6, présente les conclusions, les questions ouvertes et les perspectives de ce sujet de recherche.

Résumé des différentes contributions:

Chapitre 2 : Etat de l'art

Ce chapitre présente un état de l'art sur:

- Les services vidéo et les architectures de déploiements, en particulier pour l'IPTV et WebTV.

- Les techniques de mesure de QoE.

- Des techniques de livraison d'adaptation sont également présentées.

Chapitre 3 : l'engagement de l'utilisateur en fonction des indicateurs de performance:

Les fournisseurs de contenu sont de plus en plus intéressés à comprendre la façon dont leur contenus sont consommés (utilisation) et apprécié (perception) par leur clients. Ils comptent habituellement sur les systèmes de livraison pour livrer leurs contenus, tels que les CDN (Content Delivery Network) et Cloud Networks. Le suivi des indicateurs apparaît pour les fournisseurs comme un moyen complémentaire aux méthodes de surveillance de QoS en temps réel pour améliorer la qualité globale de leurs réseaux.

Le suivi des mesures de qualité était réalisé pendant Roland Garros 2013, à Paris, en France. Cet événement est la deuxième plus prestigieuse compétition de tennis dans le monde après Wimbledon et le plus suivi. Cette occasion nous a permis de faire une analyse de l'utilisation de contenu et des mesures de qualité dans le cas d'événements spécifiques. Dans cette étude, nous avons analysé les impacts du taux de buffering, le bitrate vidéo, le temps de démarrage et la popularité du contenu. L'engagement de l'usager étudie les interactions entre les utilisateurs et les services.

L'engagement de l'utilisateur est défini comme le lien émotionnel, cognitif et /ou comportemental qui existe à tout moment entre un utilisateur et une ressources technologiques. Dans notre analyse, le taux de buffering apparaît comme un paramètre critique qui affecte le temps de transmission de la vidéo quel que soit le type de contenu. En outre, le débit de la vidéo est un autre paramètre critique qui affecte l'engagement de l'usager jusqu'à un certain seuil dans le cas de contenus populaires.

Nous constatons également que le temps démarrage de la vidéo est moins important jusqu'à un certain niveau pour les contenus populaires. Ces résultats prouvent également que les mesures de la qualité et de leur impact sur l'engagement des utilisateurs est complexe. Comme nous l'avons remarqué, le temps de la lecture de la vidéo peut dépendre de la qualité (ces paramètres dépendent des conditions réseau), mais aussi de paramètres plus subjectifs tels que le comportement de l'utilisateur et de leurs attentes concernant le contenu.

Comme le montre l'étude, nous avons observé une corrélation entre les paramètres de qualité et l'engagement de l'utilisateur. Cependant, même si des giga-octets d'informations ont été recueillis et analysées sur des milliers de clients, certaines questions restent ouvertes. Une d'entre elles est le facteur qui est le plus prépondérant sur engagement de l'utilisateur. Quel est l'impact des informations de contextes. Il est vrai que la participation des utilisateurs en termes de lecture vidéo peut nous aider à caractériser le comportement des utilisateurs finaux sur la base de l'état du service. Mais ce paramètre ne démontre pas exactement la perception de l'utilisateur final et comment l'optimiser ?

Chapitre 4: Impact des informations de contexte sur la QoE:

La QoE est un facteur déterminant pour l'évaluation « End-to-End » des applications et services. Être capable de comprendre les besoins humains en termes de la qualité et les attentes, est au cœur de toute entreprise. Cependant, le comportement humain est subjectif, et de nature aléatoire et varie en fonction de l'environnement et le contexte. Dans ce chapitre, nous avons etudié l'influence des informations de contexte sur la mesure de QoE. Il est sans doute que les informations de contexte ci-dessous influent sur la perception du service consommé.

- Contenu (encodage et nature)
- Le réseau
- Le type de terminal

Dans notre cas, nous avons fait le choix de traiter la QoE qualitativement à travers le concept de MOS. Le MOS est la plus célèbre métrique utilisée dans la mesure subjective, où l'utilisateur est tenue de donner une note, où (1) est mauvais et (5) excellent. Ces enquêtes ont été menées à la fois expérimentalement et théoriquement:

- Dans un premier temps, une platforme d'experimentation a été developpé, afin que recueillir des notes de satisfaction des utilisateurs en fonction de la variation des informations de contexte.
- Des modèles théoriques pour MOS sont dérivés des données expérimentales. Dans notre étude, d'un point de vue technique, premièrement, nous avons analysé l'impact de la dégradation de réseau sur la QoE et deuxièmement, nous avons traité l'effet de la qualité intrinsèque de la vidéo (bitrate vidéo) sur la QoE. Pour chaque étude, nous avons considéré une variation du type d'appareil et le type de contenu.

Chapitre 5: Utilisation de MOS pour l'adaptation de contenu vidéo:

En raison de la demande croissante en matière d'accès au réseau sans fil, il est important de développer des approches et des mécanismes pour gérer la congestion du réseau.

Dans ce chapitre, nous avons proposé une mesure générale de la QoE de plusieurs flux compétitifs partageant les mêmes ressources.

Nous avons comparé notre proposition DASH (Dynamic Adaptive, le streaming via HTTP), qui est une solution d'adaptation de profil vidéo en fonction du débit disponible et des caractéristiques du terminal. Cette adaptation ne prend pas en compte la QoE des clients.

Nous avons donc proposé une approche d'adaptation de contenu vidéo, qui prend en compte la satisfaction de l'utilisateur final en termes de MOS, appelé MDASH (MOS dynamique Adaptive streaming sur HTTP). La solution proposée maximise la QoE de l'utilisateur, qui a la plus basse QoE entre les utilisateurs. Nous avons aussi proposé une architecture dans le Home Network, basée sur l'architecture fonctionnelle UIT-T pour la réservation de QoS. Nous avons validé avec des tests, qu'en prenant en compte le MOS (qui est fonction des informations de contextes), le MDASH offre de bonnes performances comparés au DASH.

Conclusion générale:

Les services vidéos sont de plus en plus populaires sur internet, l'objectif principal des fournisseurs de services est de promouvoir de nouveaux services et améliorer l'ARPU (Revenu Moyen par utilisateur). La QoE apparaît comme une mesure de la satisfaction des usagers d'un service en fournissant une évaluation de la perception de l'utilisateur. Par conséquent, il est nécessaire de développer des méthodes précises d'évaluation de la QoE, qui sont en mesure de montrer précisément les mesures de qualité qui sont vraiment perçues par les utilisateurs.

Nous avons analysé d'abord l'effet des mesures de qualité (taux de buffering, le temps de démarrage de la vidéo, et le débit moyen) sur l'engagement de l'utilisateur pour les contenus populaires et non populaires durant l'événement Roland Garros 2013.

Nous avons montré que, le taux de buffering est un paramètre critique qui impacte l'engagement de l'utilisateur indépendamment du type de contenu (populaire et non populaire). En plus, le débit vidéo est un paramètre critique qui influe sur l'engagement de l'utilisateur jusqu'à un certain seuil pour le cas des contenus populaires. Nous avons constaté également que le temps de démarrage de la vidéo est moins important jusqu'á un certain niveau pour les contenus populaires.

Afin d'avoir une évaluation précise de la QoE, nous avons étudié dans cette thèse par expérimentations, l'influence des caractéristiques de contenu, le type de terminal, l'état du réseau et le bitrate vidéo sur le MOS. Nous avons aussi fourni des modèles mathématiques qui donnent avec précision la fonction MOS en prenant en compte des informations contextuelles en termes de type de contenu et de terminal.

Comme il y a des besoins pour l'adaptation de contenus vidéo, nous avons proposé dans cette thèse, l'optimisation de trafic de plusieurs flux compétitifs partageant les mêmes ressources. Notre approche améliore la QoE perçue pour différentes sessions vidéo partageant le même réseau. Nous avons validé la proposition par des simulations et avons proposé pour une architecture dans le Home Network, qui est une extension de l'UIT-T, de l'architecture fonctionnelle dans NGN pour la réservation de QoS.

Acknowledgements

This work is the symbol of several years on the school. It's a personal satisfaction, certainly, but above that of all the people who supported me and my school.

I dedicate this thesis to my father Tourad DIALLO and mother Aissata NIANG who spared no effort to put me in the best conditions of study. This is a gift that I make them for all the sacrifices and efforts for the success of their children.

I also dedicate this thesis to all my family, my brothers and sisters, I also encourage perseverance in telling them, that success is at the end of the effort.

It is difficult to overstate my gratitude to my PhD supervisor Hossam AFIFI in Telecom SudParis, with his enthusiasm and his great efforts to explain things clearly and simply.

I would also like to show my gratitude towards Dr. Nicolas MARECHAL, my PhD advisor in Orange labs. He is a fanstastic friend, his guidance and instructions helped me to improve my knowledge in mathematics and IP-network. He helped to make my PhD less painful for me, his feedback on my thesis report assisted me in improving my mistakes.

My sincere thanks also go to to all my colleagues in AVD (Audiovisual Video Delivery) in Orange labs. In particular, the manager Laurent RUCKENBUSCH and Hassnaa MOUSTAFA for giving me the opportunity to work in Orange R&D lab and their availability for support and help. I also want to thank Frédéric FIEAU, Renaud EDEL, Jean-Baptiste HENNEQUIN, Yiping CHEN, Iuniana OPRESCU and Didier GORGES from whom I have picked up much knowledge on QoE measurements techniques, Content Delivery Network (CDN) architecture and various discussions.

I would like to thank my friends, family and colleagues, Zoulfaghar BA, Habib LY, Adama Touradou DIALLO, Coumba BA, Mamoudou DIALLO, Abdoul DIALLO, Ousmane SALL, Zoumana TRAORE, Aminata SEYDI, Yassir KHAN, Emad ABD-ELRAHMAN, Antonio MONTANI JIMENEZ, Xavier ROUARD for bringing in moments of immense joy and laughter.

To my friends and all those names I missed to mention. I would like to thank you all for your moral support and motivation.

I extend my thanks to my committee members. Thank you, Adlen KSENTINI, Ken CHEN, Liugi ATZORI, Hassnaa MOUSTAFA, Houda LABIOD and Laurent RUCKENBUSCH for accepting to review my dissertation and to be member of my PhD defense jury.

Mamadou Tourad DIALLO

Articles and Communications

Journals

1. Mamadou T. DIALLO, H. MOUSTAFA, H. AFIFI, N. MARECHAL, "Adaptation of Audio-Visual Contents and their Delivery",Communications of the ACM magazine, available on November 2013 Vol. 56, No. 11

2. Mamadou T. DIALLO, H. MOUSTAFA, H. AFIFI, N. MARECHAL, "Context Aware Quality of Experience for Audio-Visual Service Groups", IEEE COMSOC MMTC, March 2013 MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE

3. Mamadou T. DIALLO, N. MARECHAL, H. Moustafa, H. Afifi, E. ABD-ELRAHMAN, "Optimization of TV Multicast Delivery", Journal of Electronic Systems, Volume 3, Issue 4, December, 2013, Pages 135-147

4. Mamadou T. DIALLO, F. FIEAU, H. AFIFI, "Big data for Cloud Based Video adaptation", IEEE COMSOC MMTC, Vol. 9, No. 6, November 2014 MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE

Conference proceedings

1. Mamadou T. DIALLO, H. Moustafa, H. Afifi, Khalil Laghari, "Quality of Experience for Audio-Visual Services", EuroITV 2012, 4 to 6 July Berlin, Germany

2. Zied Aouini; Mamadou Tourad DIALLO, Ali Gouta, Anne-Marie Kermarrec, Yannick Lelouedec, "Improving caching efficiency and quality of experience with CF-Dash", ACM workshop on Network and Operating Systems Support for Digital audio and video (NOSSDAV 2014), Singapore March 2014

3. Mamadou Tourad DIALLO, Frédéric FIEAU, Jean-Baptiste HENNEQUIN, "Impacts of Quality of Experience in User Engagement in a Live Event", IEEE International Conference on Multimedia and Expo (ICME), Chengdu (China), July 14-18, 2014

4. Mamadou Tourad DIALLO, Frédéric FIEAU, Emad ABD-ELRAHMAN, Hossam AFIFI, "Utility-based Approach for Video Service Delivery Optimization", International Conference on Systems and Networks Communications (ICSNC), Nice, October 12-16, 2014

Poster

1. Mamadou T. DIALLO, Nicolas MARECHAL, H. Afifi, “A Hybrid Contextual User Perception Model for Streamed Video Quality Assessment”, IEEE International Symposium on Multimedia (ISM 2013), in Anaheim, California, USA

Contents

List of Figures

List of Tables

Nomenclature

3GPP	3rd Generation Partnership Project
AAC	Advanced Audio Coding
AES	Advanced Encryption Standard
ARPU	Average Revenue per User
CC/PP	Composites Capabilities Preferences Profiles
CDF	Cumulative Distribution Functions
CDN	Content Delivery Network
CP	Content Provider
CPE	Customer Promises Equipment
DASH	Dynamic Adaptive Streaming over HTTP
DRM	Digital Right Management
DSCQS	Double Stimulus Continuous Quality Scale
DSIS	Double Stimulus Impairment Scale
DSLAM	Digital Subscriber Line Access Multiplexer
DTT	Digital Terrestrial Television
ETSI	European Telecommunications Standards Institute
EZ	Eligible Zones
HAS	HTTP Adaptive Streaming
HD	High Definition
HDS	HTTP Dynamic Streaming
HLS	HTTP Live Streaming
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
ISP	Internet Service Providers
ITU-T	The ITU Telecommunication Standardization Sector

KMS	Key Management System
KPI	Key Performance Indicator
MMS	Microsoft Media Server
MMS	Microsoft Media Services
MNO	Mobile Network Operator
MOS	Mean Opinion Score
MPEG	Moving Picture Experts Group
MSS	Microsoft Smooth Streaming
NCUPM	Network Contextual Perceptual Model
NEZ	Non Eligible Zones
NGN	Next Generation Network
OMA	Open Mobile Alliance
OTT	Over The Top
PD-FE	Policy Decision Functional Entity
PEVQ	Perceptual Evaluation of Video Quality
PIM	Protocol-Independent Multicast (PIM)
PSNR	Peak Signal-to-Noise Ratio
PSQA	Pseudo Subjective Quality Assessment
PVR	Personal Video Recoder
QoE	Quality of Experience
RACF	Resource and Admission Control Functions
RNN	Random Neural Network
RTCP	Real-time Transport Control Protocol
RTMP	Real Time Messaging Protocol
RTMP	Real-Time Messaging Protocol
RTSP	Real Time Streaming Protocol
RTT	Round-Trip Time

SCF	Service Control Functions
SD	Standard Definition
SP	Service Providers
SSCQE	Single Stimulus Continuous Quality Evaluation
SSIM	Structural Similarity Index
STB	Set Top Box
TF	Transport Functions
TRC-FE	Transport Resource Control Functional Entity
UAProf	User Agent Profiles
VCUPM	Video Contextual Perceptual Model
VoD	Video on Demand
VoIP	Voice Over IP
VQM	Video Quality Metric

Chapter 1

Introduction

1.1 Definitions and work motivations

With the explosion of online video and increasing demand of multimedia and audio-visual services, Quality of Experience (QoE) has become a crucial determinant of success or failure of applications and services. As there are needs to understand human quality requirements, QoE appears as a measure of users' satisfaction from a service through providing an assessment of human expectations, feelings, perceptions, cognition and acceptance with respect to a particular service or application [1]. Measuring QoE helps operators and service providers to limit churn, increase Average Revenue per User (ARPU), market share and also ensure users' satisfaction.

There are many definitions of QoE. According to 3GPP [2], QoE indicates performance metrics as expressed from the service user's point of view. In general, a service provider fixes service requirements to meet end-users' expected QoE, which needs to be translated into parameters or metrics that the service provider can control or measure.

For ITU-T [3], QoE is the overall acceptability of an application or service, as perceived subjectively by the end-user. According to [3], the QoE captures the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.) and the overall acceptability may be influenced by the user's expectations and context.

ETSI (European Telecommunications Standards Institute) [4] has provided an alternative definition of QoE in addition to the ITU-T one, in order to extend the QoE beyond subjective measures of user-perception to include objective measures of communication process. It takes into account technical parameters (packet loss, delay etc.) and usage context variables (e.g. communication task) and measures both the process and outcomes of communication (e.g. user effectiveness, efficiency, satisfaction and enjoyment). The appropriate psychological measures will be dependent on the communication context. Unlike objective psychological measures, subjective ones are based on the opinion of the user (e.g. perceived quality of medium, satisfaction with a service).

However, QoE as defined and measured today is not sufficient to adapt content or delivery for improving the users' satisfaction. Consequently, we define in this thesis, the notion of QoE as a subjective estimation that reflects the degree of the user satisfaction following the ITU-T definition. In addition to that, the QoE is impacted by contextual parameters on the user and his environment (device type, network characteristics, content type, video encoding parameters, users preferences, user location ...).

In order to maintain a good video QoE, in most cases, video service providers and network operators are adapting contents. We may consider two types of adaptations, the content adaptation and the delivery adaptation. The content adaptation is the process of selecting, generating or modifying content (e.g. several video image qualities are proposed, where quality here must be understood as the intrinsic video quality resulting from encoding parameters such as video resolution and bitrate, etc.). This adaptation is performed manually by the end user (through the Service Portal or video player) or automatically depending on the terminal or network conditions, in order to suit the user's preferences, consumption style, computing and communications environment and usage context. The delivery adaptation is how the content is delivered through the network (from which server(s), from a cloud, through which access network)? QoE as defined and measured today is not sufficient to adapt content or delivery for improving the users' satisfaction. Actually, QoE is measured through four main methods: i) objective means based on network parameters (e.g. packet loss rate and congestion notification from routers), ii) subjective means based on the quality assessment by the users giving the exact user perception of the service, iii) hybrid means which consider both objective and subjective methodologies and iv) parametric models based on mathematical formulas.

Some research contributions also introduce methods to evaluate the QoE based on users' behavior, technical parameters, statistical learning ... [5]. We notice also several research contributions on content adaptation based on terminal capacity, user preferences, network congestion, and so on [5]. However there is a lack of methods making the use of contextual information regarding terminals, networks and contents.

So in this thesis, we propose methods to measure and analyze QoE with accuracy by considering contextual information regarding terminals, networks and contents. As described in [6], context is a fundamental part of communication ecosystem and it influences QoE. In the second phase , we investigate content and delivery adaptation, that consider the measured QoE, function of context information. Consequently, we define in this thesis, a new notion for QoE introducing more contextual parameters on the user and his environment to accurately predict the QoE and propose content and delivery adaptation techniques based on measured QoE.

1.2 Contributions of the thesis

In this thesis, we focus on QoE measurement techniques and adaptation for video services. This manuscript introduces six main contributions to the research field:

- We analyze the impacts of video startup time, buffering ratio, average bitrate and content popularity on user engagement. This contribution was published in [7].

- We investigate the impact of contextual information on QoE by experimentation. In order to investigate the effects of contextual parameters (content type, device type, user throughput), we set up experiments. Experiments consist in letting users to watch different contents on different devices and with different network bandwidths. After each session, end-users are asked to give an opinion about their satisfaction according to ITU-T recommendations, i.e. according to a five point quality scale ranging from bad (1) to excellent (5). Another set of important perceptual parameters were collected during experiments, users are asked to provide a rating for the video startup time and the percentage of buffering, in order to understand how these metrics reduce the QoE. These contributions were published in [1] [8].

- Two mathematical models are proposed: the first predicts the QoE, function of used device, video content type and the quality of the link. The second assesses the QoE function of used device, video content and video quality. These contributions were also published in [8].

- We propose MDASH (MOS Dynamic Adaptive Streaming over HTTP), which improve perceived QoE for different video sessions sharing the same network, and maximizes the QoE of user which has the lowest QoE among others, the proposal was validated by simulations and an architecture is proposed.

- We define a Utility-based approach for Video Service Delivery Optimization. Through this optimization, a global utility function is calculated based on different constraints. However, each actor has in the delivery chain has a global score for his vision, the overall optimization aims to satisfy the three actors. In this phase, a GUI (Graphical User Interface) is developed to simulate and study this utility approach. This contribution was published in [9].

1.3 Outline of this thesis

This manuscript is organized as follows:

- Chapter 2 presents an overview on video services and architectures deployments, especially for IPTV (Internet Protocol TeleVision) and WebTV (WebTeleVision) services. After that, existing video streaming techniques are presented and a comparison between HTTP based and non-HTTP based adaptive techniques is proposed. A state of art of QoE measurement techniques, video content adaptation and delivery adaptation techniques are also presented.

- Chapter 3 proposes to answer such questions: how user engagement varies with startup time, buffering ratio, average bitrate, buffering events and content popularity for live popular and unpopular contents events.

- Chapter 4 focuses on the effect of context parameters (content type, device type, user throughput) on Mean Opinion Score (MOS). In this chapter, we describe the developed experimentation platform, the adopted and related tests conditions, analyze also the obtained results from experiments and propose mathematical models to assess MOS function of contexts parameters, based on experiments results.

- Chapter 5 proposes some multi-user QoE metrics and provide optimization based on multi-users metrics in the case of DASH (Dynamic Adaptive Streaming Over HTTP), that optimizes the network resources by considering the end-user satisfaction in terms of MOS. Our approach improves perceived QoE for different video sessions sharing the same network and then improves the QoE of that user among the others which has the lowest QoE. We validate the proposal through simulations and corresponding architecture is proposed.

- And finally Chapter 6, presents the conclusions, open questions and perspectives of the research.

Chapter 2

Background and the state of art

2.1 Introduction

As communication technologies evolve, content providers and operators are facing a strong heterogeneity of devices, access technologies, protocols, network architectures, in order to maintain the gain and improve the QoE of their customers.

In this chapter, we first give an overview on video services and architectures deployments, especially for IPTV[1] and WebTV[2] services. After that, we explore existing video streaming techniques and compare recent HTTP streaming protocols with (legacy) and non-HTTP streaming protocols (in order to clarify the difference between them). We present also different methods measuring QoE. A state of art in video content and delivery is also discussed.

2.2 Video services and architectures

During the last decade, the influence of broadband Internet accesses for retail customers has driven a wide transformation of how medias are delivered. Now, a great amount of the information we consume is delivered through the IP networking protocol (telephony, TV & video, music & radio, newspapers...). In this paragraph, we introduce different services and architectures in IPTV and WebTV platforms.

2.2.1 Overview of IPTV services

IPTV is a principle of transmitting television programs through IP networks. IPTV works on a TV, with set-top-box that accesses channels and subscription services in a secure and managed IP network. Contents are diffused in managed

[1](Internet Protocol TeleVision)
[2](Web TeleVision)

mode (ie. controlled by networks operators), where they can define desired QoS for a given service.

IPTV may also include web services such as Internet and Voice Over IP (VoIP), where it may be called Triple Play and all these services are supplied by the same broadband operator using the same infrastructure [10]. In IPTV, we can consider two types of clients: those in Eligible Zones (EZ) and those in Non Eligible Zones (NEZ). The Eligibility is based on the client's distance to the nearest Digital Subscriber Line Access Multiplexer (DSLAM) and whether this DSLAM is able to provide TV services to this client. Clients in EZ (i.e. near the DSLAM) correspond to customers who benefits from sufficient connectivity conditions (value thresholds regarding bandwidth, latency and packet losses to define their $\ll$ Eligible $\gg$ client) to access to live streams using IPTV network. These thresholds are variable across network operators and mainly depend on quality acceptance levels that operators and service providers define (under those level, quality is supposed to be bad enough not to offer the service). Eligibility are improved with the emergence of new codec like H265 [11] that offers the same quality than H264 [11] only half of its bandwidth requirements. For NEZ clients, two ways are used by operators to provide IPTV services: i) Live streams are accessed by satellite or Digital Terrestrial Television (DTT) and ii) Video on Demand (VoD) are received via Internet using Progressive Download, where the player starts to playback the video before the download is complete. We are going to present in details, the principle of this technique in section 2.3.1.1.

In traditional TV programs, services are pushed to users. In IPTV, operators and content providers can deliver interactive services. Most of them propose the following services:

- Broadcast Television: This service corresponds to live diffusion. This type of transmission is common to the traditional TV delivery, in most cases, operators implement IP broadcasting to stream the most popular channel while achieving bandwidth savings in comparison to a unicast mode.

- Video On Demand: It corresponds to the case where the user can select a video content to be played anytime by the end-user. In other words, VoD is a dematerialized equivalent to brick-and-mortar video rentals shop. The content is stored in operator or content provider side. Unicast delivery is naturally used for service diffusion.

- Personal Video Recorder: PVR concept, is associated to the case where equipments (set-top box, cloud storage ...) can store video or audio streams for later playback.

- Time shifting: This function suspends the broadcast of a live program to be able to resume it later. An example is to pause to whatever purpose and restart it where we left.

- Personalization: Some operators offering IPTV services to propose some customization of TV streams so that to match end-users' preferences or habits (e.g. through content recommendation).

2.2.2 IPTV architecture

For Live TV, streams are injected in the network from the broadcaster head-end, and then a requested channel is transmitted once and replicated down to the network: generally channel switching is done through the Internet Group Management Protocol (IGMP) and the Protocol-Independent Multicast (PIM) is used to build multicast trees.

VoD contents are generally sent in unicast (i.e. two end-users watching the same content on different devices will generate two separate data streams on the network), which allows some kind on the top of initial drivers (content rights management, billing & charging...). The Real Time Streaming Protocol (RTSP) has been commonly used for video playback control (play, pause, forward, rewind, stop, change resolution).

ETSI covers some elements related to IPTV ecosystem: the customer network, the service provider network and the media content distribution [12]. In particular, ETSI provides standard use cases, functions and interfaces on standardizing use-cases, functions, and interfaces to allow interoperability and inter networking between equipment vendors, network service providers and media content distributors. Fig 2.1 presents an overview of IPTV ecosystem components.

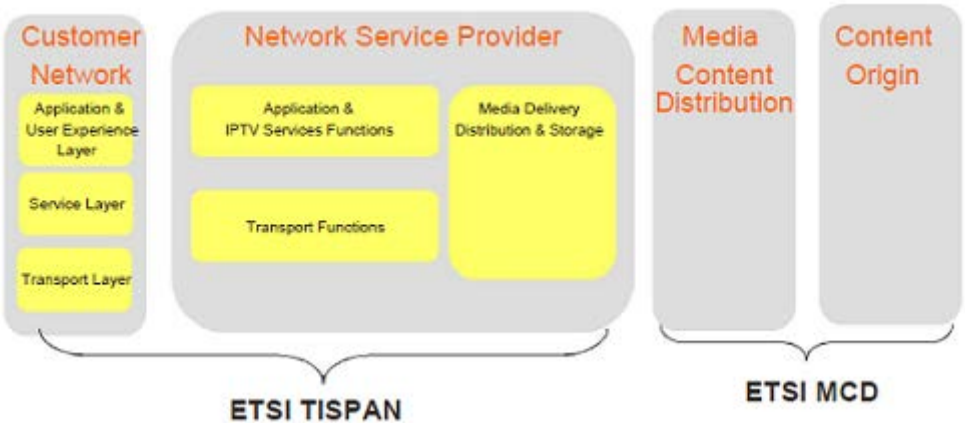

Figure 2.1: The ETSI's IPTV ecosystem

2.2.2.1 Customer Network

The customer network is in charge of providing functions for user connection to the networks as well as control over the services [13]. It is composed by the following components:

- Application and User Experience Layer: This layer comprises IPTV applications that exposes a user interface, and communicate with the Service Layer and use provided services by the Service layer to measure end-user Quality of Experience. Example of applications are: Video on Demand, Live TV environment etc.

- Service Layer: This layer provides relevant IPTV functionality to applications that are used for service management and control.

- Transport Layer [12]: This entity provides transportation capabilities for bringing IPTV streams down to the end-user's screen. Examples are: streaming functions, network attachments etc.

2.2.2.2 Network Service Provider

The network service provider provides information to access to IPTV platforms, media preparation/distribution and resource management. It is composed by the following entities:

- The application & IPTV services functions: This layer is composed by the customer facing and operator facing. The customer facing provides authorization and service provisioning of IPTV services. The operator facing IPTV applications provide operator control over IPTV system, content preparation, subscriber management, media management, etc.

- The media delivery distribution and storage: The key functionality of this layer is to provide media distribution & selection, allocation of media delivery and content storage.

- The transport functions: This layer is in charge of transport control which provides policy control, admission control, resource reservation, IP address provisioning. It includes also transport processing control in order to ensure networks data links and transmission.

2.2.2.3 Media Content Distribution

This component addresses the issues of content fragmentation and lack of interoperability of solutions for contents distribution across platforms.

2.2.2.4 Content origin

This function is the source of the content. It serves all content that is available in the network.

2.2.3 Overview of WebTV services

WebTV is mostly independent to the IPTV services offers and is available through Internet channel of many multiplay provider. The WebTV or Internet TV does not offer managed delivery and then does not support reservation of QoS (Quality of Service) [5]. Compared to managed networks (IPTV), resources reservation are based on best effort principle, where the network does not provide any guarantees that the data is delivered in a given guaranteed Quality of Service (QoS). Table 2.1 gives a comparison between IPTV and WebTV. As in IPTV, in WebTV content providers propose at least live service and in same cases, the Video on Demand may be proposed.

2.2.4 WebTV architecture

In this section we describe the WebTV architecture. Figure 2.2 gives an overview of the architecture supporting LiveTV and VoD services for Web TV.

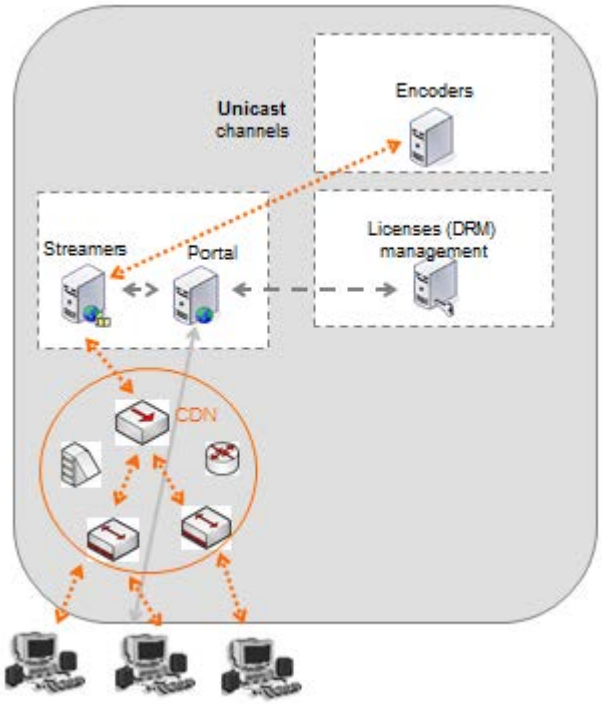

Figure 2.2: Web TV architecture

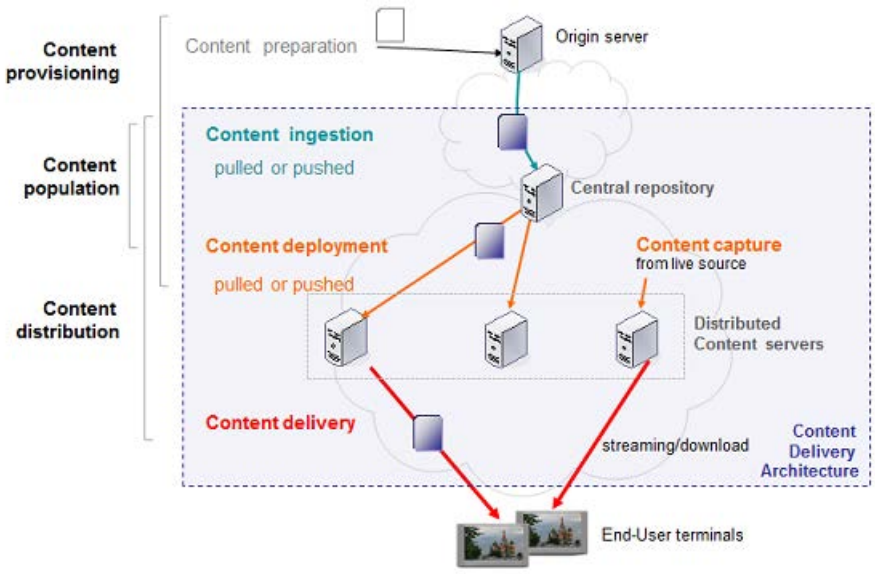

Figure 2.3: Content Delivery phases

The architecture is composed of:

- ENCODER: Convert video content and/or streams from one format (in most case raw signal, high resolution masters) to one or several secondary formats. Here the term format refers to the codec, video resolution, information flow rate (bitrate) ...

- THE PORTAL PLAN: In which all web portals are hosted, with specific service logic (open to users, geolocalization tests, etc...) and relying on an open interface (APIs collection accessible to other services).

- STREAMERS: It is a server that is in charge of streaming content to the network.

- CONTENT DELIVERY NETWORK (CDN): Generally, WebTV services providers stream the content by using a CDN. As said in [14], the CDN provides mechanisms to adapt the internet content delivery performance to the usages and to fulfill consumers' expectations. They optimize networks' and servers' resources by replicating content in the network closer to the end-user. Akamai is one of the most famous CDN multi providers over Internet as it handles almost 30% of global Internet traffic all over the world [15]. The architecture is composed of: i) Content provisioning: The first content copy is made available within the CDN network. ii) Content population: it corresponds to the content ingestion and replication within the cache nodes that store content based on specific criteria (e.g., popularity) and delivers it to the end user. Distribution to the caches nodes can be either pulled

Parameters	IPTV	Web TV
Market	Telco broadband subscribers	Internet users
Subcription	Yes	Depends on the web site
Target devices	TV screen behind a Set Top Box	PC, mobile, connected TV, Tablet
Revenue sources/Business Model	Subscription fees and premium offers, pay per-view	Advertising revenues, pay per-view
Network Type	Managed network (Private VC)	WebTV(public), unmanaged network
Delivery Mode	Streaming	Progressive Download or Streaming
Quality of Service	QoS is guaranteed	Best effort

Table 2.1: IPTV vs WebTV

(ingestion triggered by end-user request) or pushed (upload content to a storage area. iii) Content delivery: It corresponds to the content delivery through cache nodes.

- The delivery plan: We have two delivery architectures which are independent from the portal. One delivery plan is for on-demand content (with several load-balanced centralized platforms in most deployment and are currently consider a CDN architecture), that is an overlay network that replicates content closer to the end-users in order to reduce delay, save bandwidth and generate new revenue, see Figure 2.3. Another delivery plan is for live TV content where streams are provided by a dedicated head-end host by a broadcaster.

2.2.5 License Digital Right Management

Digital Right Management (DRM), is a way used by content providers, operators, publishers, individuals in order to control and protect the use of content or devices. We distinguish two aspects of content protection*: i) The scrambling of the content:* it defines what encryption algorithm is used (e.g. AES), how it is applied to the content, what is the underlying format of the content etc. Two main families

of scrambling exist: *i-1) Scrambling at content level*: this type of scrambling is independent of the transport and therefore have several advantages: Content can easily be scrambled in advance (before delivery), preferably on the encoder. Another advantage is the content can easily be recorded after delivery *and i-2) scrambling at protocol level:* this type of scrambling has to be managed on the delivery server or via a gateway handling the protocol. If this scrambling is handled individually for each delivery, it improves the security, but implies huge CPU capacities on the server side for scrambling for each client connection. *ii) The Key Management System (KMS)* : used to control the usage of content and in charge of key delivery for descrambling. Many of Key Management Systems exist: OMA (Open Mobile Alliance) DRM version 2, Microsoft DRM, Playready, Marlin... On PC, the most deployed DRMs or Key Management Systems are the Microsoft one: the Playready [16]. The main drawback of Microsoft DRM was that it imposed Windows media codecs and formats. This is not the case anymore with Playready. Today, most webTV on PC use Microsoft DRM, but none of these DRMs allows parental control. Figure 2.4 presents the WebTV Portal access security process [17].

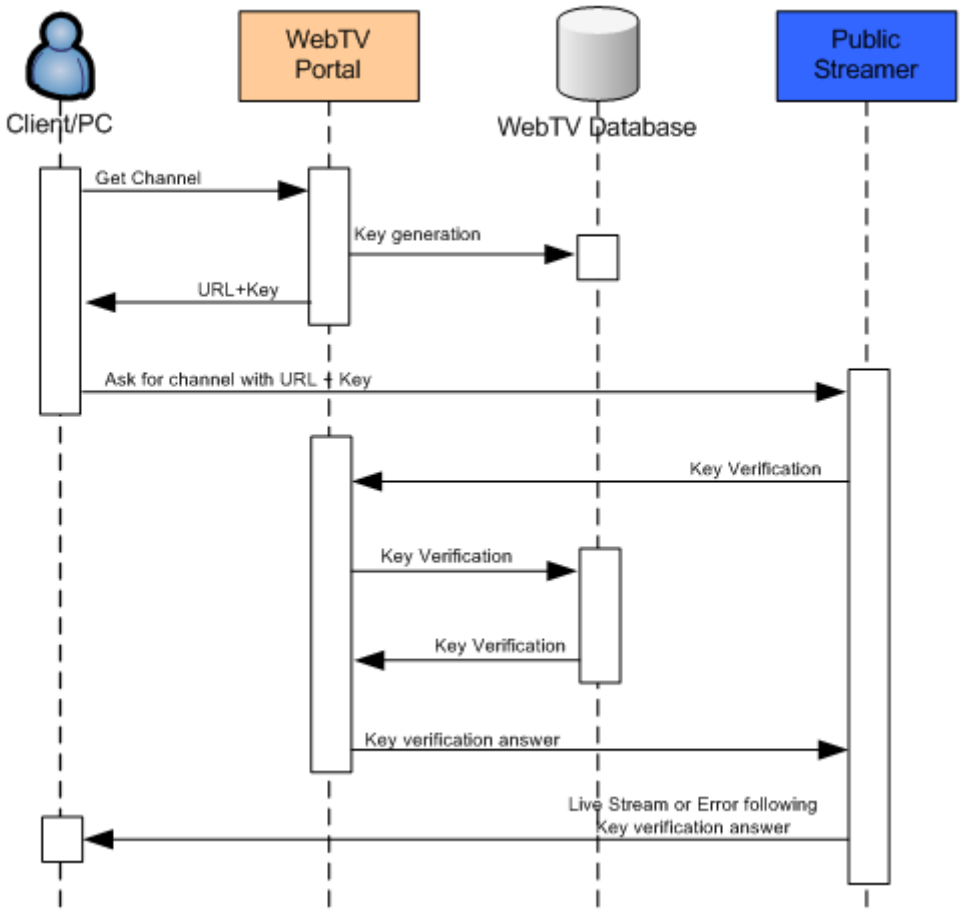

Figure 2.4: WebTV Channel Access Security Check

2.3 Streaming techniques

Nowadays there are several streaming methods to deliver audiovisual content, including: HTTP streaming, RTP Streaming (developed by IETF), MMS (Microsoft Media Services), and RTMP (Real Time Messaging Protocol) for Adobe Systems. This section reviews these different techniques.

2.3.1 HTTP-Based Streaming Techniques

HTTP-based techniques carry out a dynamic content adaptation before and/or during the session following a client-based approach and the adaptation is managed by the player. This section describes the different HTTP-based streaming techniques:

2.3.1.1 HTTP Progressive Download

This method is based on HTTP and allows the player to begin playback of media before the download is complete. The key difference between media streaming and the Progressive Download is, how the digital media content is received and stored by the end-user's device that is accessing the digital media.

The media player for the Progressive Download playback makes use the meta data located in the header of the file and a buffer in the user device. When a specified amount of content becomes available in the buffer, the playback is started.

In Figure 2.5, we represent HTTP network communication layers. Each layer in the communication stack is responsible for a number of responsibilities. HTTP is an application layer protocol, it allows applications to communicate. TCP (Transport Control Protocol) corresponds to the transport layer protocol. The TCP gets the data and ensures that, it is delivered through the network. IP (Internet Protocol represents the network layer protocol, it is responsible for taking data and moving them through the networks (routers, gateways, ...). Ethernet is a data link layer technology that transfers data between network entities.

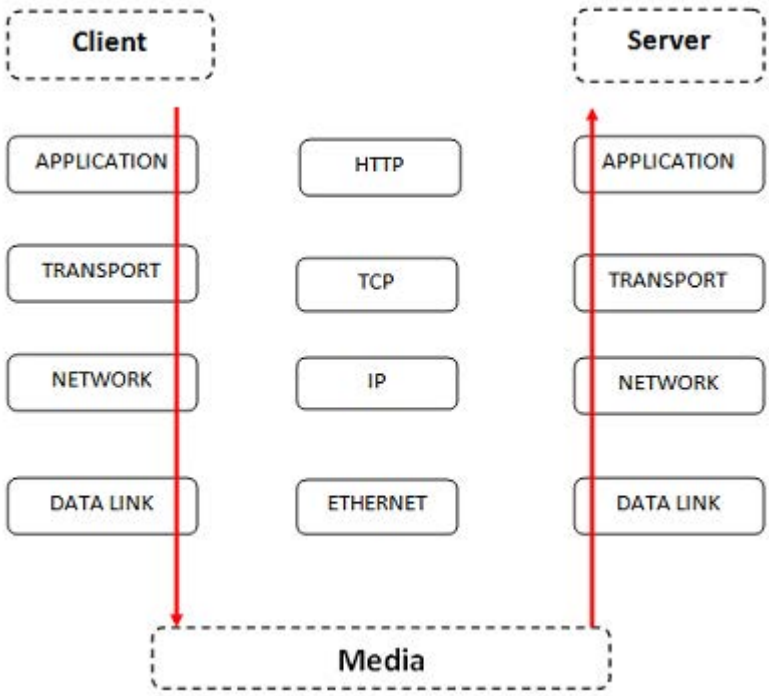

Figure 2.5: HTTP connections layer

2.3.1.2 Adaptive Streaming based on HTTP

HTTP Adaptive Streaming (HAS) [18][19] allows a multi-rate client-based streaming for multimedia content, where different bit-rates and resolutions are available for the same content and the streaming is managed by the client.

As explained in [20], in comparison with traditional adaptive streaming techniques, deployment of HAS presents opportunities for services and content providers. The server sends a manifest file containing the description of content pieces namely chunks (supported codec, minimum bandwidth, resolution, bit rates, URL...). Once the client receives the manifest, it is able to request some indexed fragments according to its environment (available bandwidth, screen resolution, supported codec...). The following are examples of Adaptive Streaming methods based on HTTP.

2.3.1.2.1 Apple HLS (HTTP Live Streaming) [21]: This solution was introduced by Apple in 2009. It was very quickly adopted by OTT (Over The Top players) and is now available on all Apple devices (iPhone, iPad, iPod...) as well as some STB (AirTies, Netgem, Amino...) and most of players and video embedding frameworks (VLC media player release 1.2.0, QuickTime X Player...).

The native codecs chosen for HLS are MPEG H.264 for video and AAC (Advanced Audio Coding) for audio. In order to implement video streaming over HLS, the following steps are required: i) Encoding video in H.264/TS format at different bitrates ii) For each encoding profile, a Stream Segmenter cuts each version of the content into short pieces named chunks, typically 10 seconds each, and generates a playlist m3u or m3u8 format containing URL for each chunk of this encoding profile. iii) Generating a general index file (manifest) indicating each available encoding profiles (bitrate, codec ...) and the URL of the corresponding playlist files. iv) Distributing content chunks, playlists and manifest to the HTTP server (origin or cache). v) Measuring playback conditions on the user device (bandwidth, CPU, device capabilities ...) and selecting the most suitable chunk accordingly.

2.3.1.2.2 Google WebM [22]: This method is the Google's royalty free approach for video adaptive streaming proposed in 2010. It uses VP8 video codec for video and Vorbis for audio and doesn't require segmentation of the media into chunks. However one media stream is seen as one file. To stream a video through WebM, the following steps takes places: i) Encoding the video and audio content in VP8 and Vorbis respectively, in different bitrates (i.e. quality profiles). ii) Multiplexing them into a single WebM file. iii) Using a Web server (origin or cache) to deliver the WebM files. The adaptive bitrate process mainly relies on the server, which selects the audio/video streaming bitrates before multiplexing and pushes the video content in an output buffer. While sending the content to

the network, the server detects if there is enough bandwidth towards the client, otherwise it scales down to a lower quality profile (lower bitrate).

2.3.1.2.3 Microsoft Smooth Streaming (MSS) [23]: Smooth Streaming is a streaming protocol released by Microsoft in 2009 as an extension of Silverlight 3.0 [24] which is an application framework for writing and running rich Internet applications. MSS specifications only allows for H264 and AAC codecs. Smooth Streaming general principle is quite similar to HLS streaming, as depicted by the following implementation steps: i) Encoding video and audio in different bitrates (i.e. quality profiles). ii) Using a Stream Segmenter to generate content fragments(chunks) and multiplexing them into a container iii) Distributing video content through HTTP server (origin or cache). iv) Generating and distributing a manifest file that lists the available profiles (bitrates, resolution...), languages, corresponding URLs for chunks.

2.3.1.2.4 Adobe HTTP Dynamic Streaming (HDS) [25]: Adobe's solution for streaming media over HTTP is a comprehensive open source video delivery. The principle of Adobe is not very different from Microsoft Smooth Streaming. The HDS principle follows these different steps: i) Creation of manifest files (.f4m) ii) Creation of segmented files (.f4f) which correspond to chunks (fragments) iii) Creation of index files (.f4x) containing specific information about the fragments inside the segmented files (available bitrates, codec's, URL's to stream content ...).

All these files are multiplexed into a single stream and sent to the client device. The supported codecs for HDS are H.264 and VP6 (video), and AAC or MP3 (audio). The terminal in the manifest file has many quality choices and selects the most suitable.

2.3.1.2.5 Dynamic Adaptive Streaming over HTTP (DASH) [26]: MPEG DASH is a promising ISO Standard for video streaming services over HTTP published in April 2012 and which is gaining popularity. DASH has the potential to replace existing proprietary technologies like Microsoft Smooth Streaming, Adobe Dynamic Streaming, and Apple HTTP Live Streaming (HLS). A unified standard is needed because it will help for rationalization of cost storage, development, maintenance, support and evolution for all DASH devices.

All HTTP-based adaptive streaming technologies have two components: the pure encoded audiovisuals streams, and manifest files that indicate to the player which streams are available (bitrates, codecs, resolutions...) and how to access them (e.g. chunk URL). For DASH, the AV streams are called the Media Presentation, while the manifest file is called the Media Presentation Description which is encoded in an XML format. Like other adaptive streaming techniques,

this manifest identifies alternative streams, their respective URLs, network bandwidth and CPU utilization. On this basis, the player chooses the most adapted stream. Two types of file segment types are allowed in DASH: MPEG2 TS (currently used by HLS), and ISO Base media file format (ISO BMFF, currently used by Smooth Streaming and HDS). This simplifies potential migration of existing adaptive streaming platforms to MPEG DASH, as the media segments can often remain the same, and only the index files need to be migrated to the MPD (Media Presentation Description) format.

In Figure 2.6 from [27], there is the principle of adaptive streaming over HTTP. The device selects a representation (2000 kbps, 1000 kbps and 500 kbps) based on the available throughput.

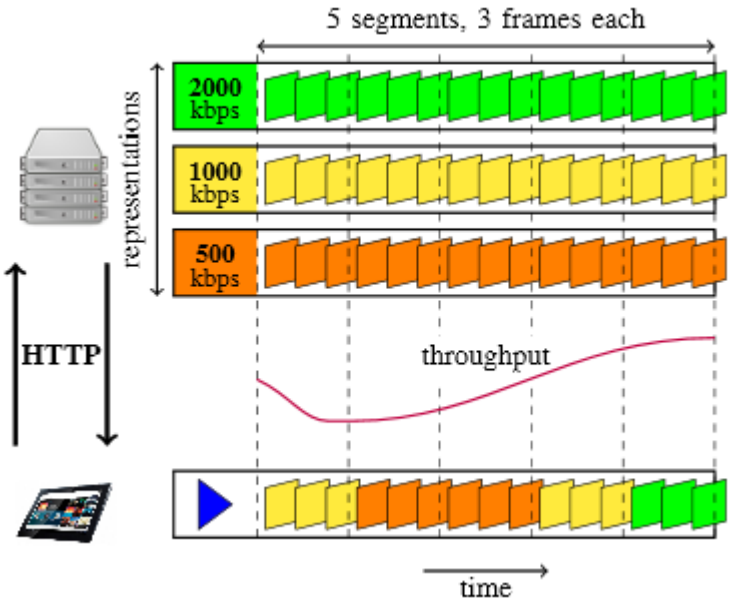

Figure 2.6: Video representations principle in adaptive HTTP streaming

2.3.2 Adaptive Streaming Techniques not based on HTTP

Real-Time Protocol (RTP) [28], Real-Time Messaging Protocol (RTMP) [29] and Microsoft Media Server (MMS) [30] streaming techniques are not HTTP based and the adaptation (if it is enabled) is managed by the server following a server-centric approach. In RTP, the content adaptation can take place making use of the RTCP (Real Time Control Protocol) reports sent between the clients to the server. These reports contain information such as packet loss, jitter, RTT (measured/estimated at the client side) that can help the server in adapting the content to network conditions. For example, if the connection deteriorates and the transfer rate decreases, the content is streamed with a lower quality so that playback interruptions are avoided, and stream quality is increased if the connection

becomes more fluid.

An advantage of this type of streaming technique is the fast start ability, that is, to start content streaming without delay. On the other hand, the drawback mainly stands in the need of a dedicated server with non-negligible license cost (examples are Xiph, Icecast, Real Helix Streaming Server, Windows Media Services, Adobe Flash Media Server, QuickTime Streaming Server ...). Most of these streaming techniques use UDP (User Datagram Protocol) and this transport protocol does not retransmit lost data and has difficulties of passing a proxy caching. Consequently, most of these streaming techniques are particularly suited for streaming video services over a fully controlled end-to-end architecture (managed networks) for which QoS and operability can be assessed and mastered.

The following subsections discuss these streaming techniques:

2.3.2.1 Audiovisual delivery based on RTP

The Real-time Transport Protocol provides end-to-end network transport functions suitable for applications transmitting real time data. RTP does not guarantee reservation of QoS for real-time applications [28].

- The RTCP (Real-time Transport Control Protocol) is used for monitoring information about the service, for example information about delay, jitter, packet loss ... [28].
- The RTSP (Real Time Streaming Protocol) is an application level protocol designed for use in entertainment and communications systems to control media streaming [31].

2.3.2.2 Audiovisual delivery based on RTMP

Real Time Messaging Protocol (RTMP) was initially a proprietary protocol developed by Macromedia which is now Adobe property and free to use. This protocol is used for streaming audio, video and data over the Internet, between Adobe player (Flash) and a server [29].

2.3.2.3 Audiovisuals delivery based on MMS

Microsoft Media Server (MMS) is the name of Microsoft's proprietary network streaming protocol used to stream content in Windows Media Services. MMS can be transported via UDP/TCP [23].

2.3.3 Comparison

In this section, we compare the two families of streaming protocols described on the previous sections. The comparison is based on criteria like standard player,

origin server, chunk duration, proprietary or not, we enumerate advantages and drawbacks.

PARAMETERS	NON HTTP ADAPTIVE STREAMING		
Methods	RTSP	MMS	RTMP
User or server centric	Server centric	Server centric	Server centric
Standard Player	Depend on server that implemented the solution	Windows media player	Adobe flash player
Origin Server	Streaming server	Windows Media Server	Adobe Flash server
Recommended chunk duration	Not available	No chunk	No chunk
Standard content protection	Conditional Access System (CAS)	Conditional Access System (CAS)	Conditional Access System (CAS)
Proprietor or public	Standardized by IETF	Microsoft property	Public use
Advantages	Adapted to real time, it is server based then easy control by the operator		It is in public use
Disadvantages	Packet losses cause artifacts. Dedicated server is required		

Table 2.3: Comparison of Non HTTP Adaptive Streaming techniques

Parameters	HTTP Adaptive Streaming			
Methods	HLS	SS	HDS	WebM
User or server centric	User centric	User centric	User centric	Server centric
Standard Player	iOS for mobile and Quick time for (PC)	Silverlight for PC and Win Mobile 7	Flash Player10.1	Chrome navigator
Origin server	Web server HTTP 1.1	IISv7	Flash Media Server 3.5	Web server HTTP 1.1
Recommended chunk duration	10 seconds	2 seconds	2- 4 seconds	No chunk
Standard content protection	Advanced Encryption Standard (AES)	PlayReady	Adobe Flash Access 2	No protection
Proprietor or Public	Apple property	Microsoft property	Adobe property	Google property
Advantages	1. Adapted to bandwidth variation 2. The user application manages the client bitrate 3. Fast content switching The required resource is a HTTP server, Firewalls/ NATs traversal			Royalty free and open solution
Drawbacks	1. The operator has no control over its bandwidth 2. The client switches from one flow to other if the network conditions allow it. 3. It is client centric then lack of operator control, start with delay, packet loss retransmitted.			Client's expectations are not taken into account

Table 2.2: Comparison of HTTP Adaptive Streaming techniques

2.4 Assessment of video Quality of Experience

Service Providers (SPs) use Quality of Service (QoS) parameters such as bandwidth, delay or jitter to guarantee good service quality. QoS is achieved if a good QoE is also achieved for the end users in addition to the classical networking configuration parameters [32]. The challenging question is how to quantify the QoE measure. In this section, we present different methods measuring the QoE.

2.4.1 Objective measures

Objective QoE measuring techniques are based on network related parameters that need to be gathered to predict the users' satisfaction. These techniques work without human intervention. Objective measuring methods follow either an intrusive approach, that requires reference image/video/audio content or a non-intrusive approach that does not require reference information to predict the Quality of Experience.

2.4.1.1 Intrusive methods

Several objectives QoE measurement solutions follow an intrusive approach. They need both the original and degraded signal (audio, video, and image) to measure QoE. Although intrusive methods are very accurate and give good results. The following subsections present some objective intrusive techniques.

- PSNR (Peak Signal-to-Noise Ratio)

This objective method uses the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. It is defined via the Mean Squared Error (MSE) between an original frame "o" and the distorted frame "d"as follows [33].

$$MSE = \frac{1}{M.N}. \sum_{m=1}^{M} \sum_{n=1}^{N} |o(m,n) - d(m,n)|^2 \tag{2.1}$$

$$PSNR = 10.log(\frac{255^2}{MSE}) \tag{2.2}$$

Where each frame has M × N pixels, and o (m, n) and d (m, n) are the luminance pixels in position (m, n) in the frame. Then, PSNR is the logarithmic ratio between the maximum value of a signal and the background noise (MSE). If the maximum luminance value in the frame is L (when the pixels are represented using 8 bits per sample, L = 255) then:

- Perceptual Evaluation of Video Quality (PEVQ)

PEVQ is an accurate, reliable and fast video quality measure. It provides the Mean Opinion Score (MOS) estimates of the video quality degradation occurring through a network, e.g. in mobile and IP-based networks. PEVQ can be ideally applied to test video telephony, video conferencing, video streaming, and IPTV (Television over IP) applications.

The degraded video signal output from a network is analyzed by comparison to the undistorted reference video signal on a perceptual basis. The idea is to consider the difference between the luminance and the chrominance domains and calculates quality indicators from them.

Furthermore the activity of the motion in the reference signals provides another indicator representing the temporal information. This indicator is important as it takes into account that in frames series with low activity the perception of details is much higher than in frames series with quick motions. After detecting the types of distortions, the distorted detected information is aggregated to form the Mean Opinion Score (MOS) [34].

- Video Quality Metric (VQM)

VQM is a software tool developed by the Institute for Telecommunication Science (ITS) to objectively measure the perceived video quality. It measures the perceptual effects of video impairments including blurring, jerky/unnatural motion, global noise, block [35].

- Structural Similarity Index (SSIM)

SSIM uses a structural distortion based measurement approach. Structure and similarity in this context refer to samples of the signals having strong dependencies between each other, especially when they are close in space. The rational is that the human vision system is highly specialized in extracting structural information from the viewing field and it is not specialized in extracting the errors. The difference with respect to other techniques mentioned previously such as PEVQ or PSNR, is that these approaches estimate perceived errors on the other hand SSIM considers image degradation as perceived change in structural information. The resultant SSIM index is a decimal value between -1 and 1, where the value 1 indicates a good score and the value -1 indicates a bad score [36]. It is easy to collect network parameters and to have reference video for no-real time traffic.

The main problem is these techniques do not consider user's opinion, gathering network parameters require more signaling, monitoring sensors and algorithms. The main drawback is, these techniques are not suitable to real-time applications as it is not always easy to have the original signal.

2.4.1.2 Non-Intrusive methods

The objective non-intrusive approach presents methods that can predict the quality of the viewed content based on the received frames without requiring the reference signal but using information that exist in the receiver side. The following are some methods that predict the user perception based on the received signals. The method presented in [37] is based on the blur metric. This metric is based on the analysis of the spread of the edges in an image which is an estimated value to predict the QoE. The idea is to measure the blur along the vertical edges by applying edge detector (e.g. vertical Sobel filter which is an operator used in image processing for edge detection.).

Another method is presented in [38] based on analyzing the received signal from the bit stream by calculating the number of intra blocks, number of inter blocks, and number of skipped blocks. The idea proposed in this work is to predict the video quality using these parameters. The predictor is built by setting up a model and adapts its coefficients using a number of training sequences. The parameters used are available at the decoder (client side). The E-model proposed in [39] uses the packet loss and delay jitter to quantify the user perception of service. The E-model is a transmission rating factor R.

$$R = R_o - I_s - I_d - I_e + A \tag{2.3}$$

Where R_o represents the basic signal-to-noise ratio, I_s represents the impairments occurring simultaneously with the voice signal, I_d represents the impairments caused by delay, and I_e represents the impairments caused by low bit rate codecs. The advantage factor A is used for compensation when there are other advantages of access to the user. It is easy to collect network parameters in the client side. The main problem is, these techniques do not consider user's opinion as the non-intrusive methods. Unlike intrusive techniques, these methods are suitable for real time traffic, because, there is no need for reference information; everything is done in the receiver side.

The proposed model in [40], the Packet-E-Model (P-E-model), is a subjective and dynamic quality evaluation for voice over IP. It extends and adapts the E-model to the particular context of IP networks, characterized by a high variability and complexity. It computes and provides a MOS score, reflecting the subjective quality of the voice communication. P-E-Model takes into account parameters such as delay or packet losses, observed on the IP path, for MOS calculation.

2.4.2 Subjective Techniques

Subjective QoE measurement is the most fundamental methodology for evaluating QoE. The subjective measuring techniques are based on surveys, interviews

MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Table 2.4: Mean Opinion Score Rating

and statistical sampling of customers to analyze their perceptions and needs with respect to the service and network quality. Several subjective assessment methods suitable for video application have been recommended by ITU-T and ITU-R. The subjective measures present the exact user perception of the viewed content (audio, video, image...) which is considered as a better indicator of video quality as it is given by humans. The most famous metric used in subjective measurement is the MOS (Mean Opinion Score), where subjects are required to give a rating using the rating scheme indicated in Table 2.4.

In order to analyze subjective data, quantitative techniques (e.g., statistics, data mining, etc.) and qualitative techniques (e.g., grounding theory and CCA framework) could also be used [41]. Once subjective user study is complete, data are to be analyzed using some statistical or data mining approaches. Conventionally, non-parametric statistics is used for ordinal and nominal data, while parametric statistic or descriptive statistics is used for interval or ratio data. Among subjective techniques proposed in ITU-R Rec BT.500-11, we can mention:

- Double Stimulus Impairment Scale (DSIS)

In this method, the reference sequence is always displayed before the test sequence. Observers are asked to judge the level of impairment for each test sequence, using a five-point scale.

- Double Stimulus Continuous Quality Scale (DSCQS)

Pairs of multiples sequences (containing degraded and references contents) are presented to users in this approach. End-users are asked to give their perception, after watching degraded and original contents.

- Single Stimulus Continuous Quality Evaluation (SSCQE)

In this method observers are asked to watch the degraded video, the reference is not presented. The end-user gives continually the quality since, it changes during the streaming.

Subjective measures are very accurate (based on user's perception) and are relevant for any multimedia traffic. The main drawback is, it is not realistic to ask all viewers their perception about the service, because, it require manpower and engaged users.

2.4.3 Hybrid Techniques

Hybrid QoE measurement merges both objective and subjective means. The objective measuring part consists of identifying the parameters which have an impact on the perceived quality for a sample video database. Then the subjective measurement takes place through asking a panel of humans to subjectively evaluate the QoE while varying the objective parameters values.

After statistical processing of the answers each video sequence receives a QoE value (often, this is a Mean Opinion Score, or MOS) corresponding to certain values for the objective parameters. To automate the process, some of the objective parameters values associated with their equivalent MOS are used for training an RNN (Random Neural Network) and other values of these parameters and their associated MOS are used for the RNN validation.

To validate the RNN, a comparison is done between the MOS values given by the trained RNN and their actual values. If these values are close enough (having low mean square error), the training is validated. Otherwise, the validation fails and a review of the chosen architecture and its configurations is needed [42].

In this approach, training the RNN system is done by subjective scores in real-time usage. The system maps the objective values to obtain the Mean Opinion Score (MOS). Advantages of this method are minimizing the drawbacks of both approaches and it does not require manpower (except in the subjective quality assessment preliminary step). The disadvantage of this method, in order to have accurate measures, training is time consuming.

Engaged users (for subjective tests) and neural networks are required in order to build hybrid models. It is easy to have subjective values (only once for learning). Once the tool has been trained, nonlinear function can map any possible combination (corresponding to selected parameters) into MOS (Mean Opinion Score). This method is very accurate, because it considers both subjective and objective parameters and also can be used for any multimedia traffic.

The method presented in [43], the PSQA (Pseudo-subjective Quality Assessment (PSQA) is a hybrid technique for QoE measurement. In this approach, training the RNN system is done by subjective scores in real-time usage. The system learns the nonlinear relation between the objective values and the perceived quality.

2.4.4 Parametric models

In this section, different parametric models are presented; these models are based on mathematical formulas, for QoE estimation. Each of them, uses different parameters as input (packet loss, frame rate, encoding bitrate ...), we review some parametric models in this paragraph and describe considered parameters in each model.

2.4.4.1 ITU-T G 1070 model

This model is standardized by ITU-T Recommendation, it takes into account the encoding bitrate (bits/seconds), frame rate (fps), and is expressed by the formula below:

$$I_c = I_{cmax}.e^{-\frac{ln(f)-ln(v1+v2.b)}{2.(v6+v7*b)^2}} \tag{2.4}$$

The parameter is b is the encoding bitrate, f is the frame rate, $v1$, $v2$, $v6$ and $v7$, are the coefficients of the model. I_{cmax} is the maximum value.

2.4.4.2 M. Ries & al model

These authors proposed a method for estimating QoE depending on the encoding rate (bits/seconds) of the video, the frame rate (fps) and the type of content. They proposed to classify video classes differ depending on the type of content, so spatial information in each video. The developed model is:

$$I_c = A + B.b + \frac{C}{b} + D.f + \frac{E}{f} \tag{2.5}$$

b is the video encoding bitrate, frame rate is f, A, B, C, D and E, the model coefficients that were calculated in the case of a VGA screen.

2.4.4.3 A. Khan & al model

These authors proposed a method for measuring video quality function of the encoding rate, the frame rate (number of frames per second), distortion due to packet loss and the type of content. The contents were classified into three types: fast movements, slow movements and means motion. Models are shown in (3), 4 and 5. Parameter f is the number of frames per second (fps), b the encoding rate, the parameters a_1, a_2 and a_3 are the coefficients of the model.

$$Vq = I_c.I_t \tag{2.6}$$

$$I_c = a_1 + a_2.f + a_3.ln(b) \tag{2.7}$$

$$I_t = \frac{1}{1 + a_4.p + a_5.p^2} \tag{2.8}$$

V_q, is the estimation of video quality, I_c predicts the video quality due to the encoding process. This metric depends on the frame rate (fps), the encoding rate (b) and constant a_1,a_2 and a_3 (models parameters). The metric I_t, estimates the

quality due to the transmission process, it depends on the packet loss (p) and the parameters a_1, a_2 and a_3, the coefficients of the model. The model was tested on a QCIF screen with frame rate between 10 fps to 30 fps, encoding bitrate between 18 kbps to 512 kbps, packet loss between 1 % to 20 %. The results were validated using the PSNR metric without doing user testing.

2.4.4.4 Yen-Fu & al model

Yen-Fu & al have presented a model of QoE. The proposal takes as input parameters the frame rate (fps), the type of terminal and the content type.

$$V_q = V_{qmax}.(\frac{1 - e^{-c\frac{f}{fmax}}}{1 - e^{-c}}) \tag{2.9}$$

V_{qmax}, is the quality is obtained for the frame rate $fmax$ (30fps), f is the frame rate and c is a coefficient of the model which is calculated for considered terminals (CIF and QCIF) and type of content. The authors did not specify an analytical formula to calculate this parameter.

2.4.4.5 IQX Hypothesis

IQX Hypothesis describes the QoE function of QoS (Quality of Service parameters). QoE appears as a solution to some differential equations and the expressions are functions of QoS parameters. In particular packet losses and packet reordering are studied in [44] and [45]. The solution of differentials equation is shown in equation 2.10. The parameters α, β and μ are the coefficients of the model.

$$IQX = \alpha + \beta.e^{-\mu.QoS} \tag{2.10}$$

Parametric models are easy to implement, since there is no need to access to the original video. They may be applied to network design, network assessments and/or to real time monitoring. The quality estimation is easily computed as the result of a direct mathematical formula.

In this paragraph we make the state of art of different parametric models, that estimate the QoE function of some parameters. We decided in this thesis to build parametric models that assess the QoE more precisely by considering context information.

2.5 Video Content Adaptation techniques

This section describes some means of performing Content Adaptation for IPTV and WebTV services.

2.5.1 Content delivery Adaptation in WebTV and IPTV

For IPTV Live and VoD (Video on Demand) service, the user has generally two choices: Standard Definition (SD) streams and High Definition (HD) streams. Nevertheless, the operator may enforce a given content quality based on the available network bandwidth and on the end-user's subscription type. WebTV contains TV and/or VoD services offered by a 3rd party available from any Internet access. This method is thus by default available on unmanaged networks, where the best-effort is the unique possible QoS traffic class.

Some Content Providers adapt content to the network conditions by using HTTP Adaptive Streaming for their live TV channels. With this technology, users don't care about whatever video quality to choose in order to match their available bandwidth: the video player will automatically request content which are the most adapted to network status and the device capacity. On the contrary, some Content Providers for the VoD let the client chooses the type of delivery as follows:

- Streamed mode available in SD (e.g. 620 kbps stream) for "Instant Viewing"
- Progressive Download mode available for both HD (e.g. 1500 kbps) and in SD (e.g. 620 kbps), with a possible non-negligible start-up delay for HD which depends on the client's bandwidth.

The current Content Adaptation in IPTV and Web TV doesn't not consider a sufficiently large set of parameters to fully enable optimal QoS and QoE. The user context is considered in a limited manner through mainly considering the characteristics of the used device and network. In addition to that, network context is considered only in terms of bandwidth availability while ignoring the cost of using this bandwidth instead of allocating it to monetized services. Neither is considered the matching degree of the content to the users' preferences.

Consequently, Context Awareness need more consideration in Content Adaptation through considering context information (network context, user context, terminal context, content context) in a dynamic manner during the session.

2.5.2 Related research contributions

The classification of existing research contributions on content adaptation, show three main categories: i) Content Adaptation: Which version of a given content shall we transmit? (codec, bitrate, video resolution ...). This aspect is related to the encoding of the content information. ii) Delivery Adaptation: How the content is delivered through the network (unicast, multicast, from which server(s), from a Cloud, through which Access Network)? This aspect is related to the service & network aspect of the content transmission iii) Adaptation of Content and Delivery: Integrating adaptation of both content and its delivery.

2.5.2.1 Video Content Adaptation

Several research contributions exist about adapting content based on terminal capacity, network congestion, user profile and service requirements. For instance, the method in [46] provides a QoE-guaranteed service that maximizes the visual expectation of the viewer by considering the screen size on his device.

In [47] users are allowed to define their preferences (user profile) during service subscription, according to some categories based on QoS requirements (Streaming, Conversational, Interactive, Background). For example, the Streaming traffic class is sensitive to packets losses. The user can also select different types of subscription (Bronze, Silver, Gold, Platinum) for each profile and traffic class. There are maximum and minimum QoS parameters where an adaptation is needed for each type of service and user profile.

The method introduced in [48] adjusts the quality level and transmission rate of video streaming on the basis of the wireless channel status (Modulation Coding Scheme, Signal to Interference-Ratio level),the user location and client buffer status. The transmission rate is determined as a function of the network context (packet loss, jitter ...) and some player buffer ratios.

In [49], the adaptation of the transmission rate is done on the basis of the pre-buffering time and the available bandwidth (network status/context), so the QoE is maximized even in case of network congestion.

In [50], authors propose a concept of reactive control of video adaptation. In this work authors use the technology of active network, in order to conceive an approach of reactive control, in order to adapt the video flow to the variations of network resources. The network supervises the transmission of video packets and reacts to flow variations by sending to the encoder a recommendation of the available bandwidth in the network for its flow without requiring any any feedback from the receiver.

The work in [51] adapts content using two parameters: the "congestion" (C) and the "degradation" (D). The congestion is defined as the fraction of the number of video blocks lost (BL) divided by the total number of video blocks sent (BS) within an interval of time. After predicting the estimated QoE (denoted MOSt), the degradation is defined as the difference between the maximum achievable Mean Opinion Score (MOS) and the estimated MOS (MOSt). The Sender Bit Rate (SBR) is computed by on an algorithm using congestion and degradation.

W3C proposes in [52] Content Adaptation techniques within the Composites Capabilities Preferences Profiles (CC/PP) for web content and User Agent Profiles (UAProf) for mobile phones. These frameworks can be used to deliver devices contexts (screen size, audio/video capabilities...) and users' preferences (language, type of content ....) and allow devices to communicate their capabilities and preferences to servers. The server can then accurately adapt content according to this information.

2.5.2.2 Video Delivery Adaptation

We can divide contributions for video delivery into three main methods: i) the network-centric approach, in which decisions are made at the network side (mainly by network operators), ii) the user-centric approach making the decision based on the user's benefit, and iii) the context-centric approach, where the switching decision is made by considering different context information.

2.5.2.2.1 The network-centric: In this approach, decisions are made by the operators and they are principally based on their benefits. Authors in [53] propose a distributed strategy to get network topology information, and use Internet Control Message Protocol (ICMP) ping method to measure Round-Trip Time (RTT), in order to switch to a network which has the lowest RTT.

The work in [54] proposes the load balancing algorithm which automatically selects network candidate based on local resource conditions. The main advantage of this method is the network resources optimization. But all these techniques do not consider content provider expectations and users QoE.

2.5.2.2.2 The user-centric: Network switching is made in order to satisfy user's benefits, without considering network load and content provider expectations. In [55] authors consider the QoE measurements over different access types. After predicting a MOS with Pseudo Subjective Quality Assessment (PSQA), a vertical handover (change in access network) is carried out towards the network offering the best MOS. It can be noticed that the user-centric approach has the main drawback from a load balancing perspective, since users generally consider only their own benefits while making decisions and letting the Operator and Content Provider benefits.

2.5.2.2.3 The context-centric approach: In this approach, the delivery decision optimization is made by considering different contexts (Content Provider, Operator, and Client). In [56], an algorithm, called Smooth Adaptive Soft-Handover Algorithm (SASHA). Its goal is to improve the user perceived quality while roaming through heterogeneous wireless network environments. The score of each connection is evaluated based on a comprehensive Quality of Multimedia Streaming (QMS) including the following metrics: QoS, QoE, Cost, Power efficiency and user preferences. The idea is to adapt delivery in the network that has the best (QMS) score. The disadvantage is the no consideration of content provider expectations in the adaptation process.

In [57], Hierarchical and Distributed Handover (HDHO) method is proposed, a distributed handover decision framework which takes into account the objective of Content Provider by considering the content requirements in terms of resources, Operator in terms of network load and user preferences by considering cost sen-

sibility. Even if, this proposal takes into account the aim of each actor on the delivery chain, some relevant parameters are omitted. In content provider side the cost of transmitting the content in a network is missed, in network side the cost and hardware status are absent, in client side the perceived QoE is not taken into account. In order to maximize a perceived QoE in users' side, respect conditions of content providers and the operators' benefits, we need to define a new video delivery optimization which takes into account the objective of each actor.

In [58, 59], we proposed a solution that adapts the multicast delivery for Mobile TV service through optimizing the tree structure of multicast nodes in a dynamic manner according to the different context of the user and the network.

2.5.2.3 Adaptation of Content and Delivery

The solution proposed in [60] chooses the most suitable content to be delivered to the user and selects the best delivery mean. Two decision entities are considered, namely the Service Manager responsible for the service delivery, and the Mobility Manager responsible for the network connectivity.

For Content Adaptation, the service management entity will be notified when a terminal request the streaming of a new video (contents encoding is done with SVC), and decides which version should be sent according to the user rights, to his preferences, to his terminal capabilities and to the network congestion. The Service Management entity then provides its decision to the Service Execution entity which sends the corresponding signalization.

For Delivery Adaptation, the Mobility Manager gets notified about the network-related events and service requirements and retrieves network-related information and decides which possible network connection(s) must be used for every service based on information such as cost, network load, and user preferences.

2.5.2.4 Comparison of content and delivery adaptation techniques

This section provides a comparative study regarding different issues that should be addressed in content adaptation techniques, mainly considering user context, user satisfaction, network congestion and required resources. There are a lot of research contributions on content adaptation and its delivery. We have therefore classified them into several categories based on:

2.5.2.4.1 Terminal capacity: This technique has a lot of advantages among which we can mention the consideration of user context by using the terminal capacity. The disadvantage of this method is not considering the dynamic variation of user's needs, network resources optimization, user satisfaction etc.

2.5.2.4.2 Network congestion: The main advantage of this method is the network resources optimization. The lack is the no consideration of others context

information like user context (his location, his preferences, his profile...), QoE,...

2.5.2.4.3 User profile and service requirements: The advantage is the consideration of user context by considering his profile. This adaptation technique is easy to implement because the user profile and service needs are known by the deliver. The disadvantage is the no consideration of user context, network congestion in the adaptation process.

2.5.2.4.4 Network congestion & terminal capacity: It considers some context information like terminal capacity and helps on resources optimization. The needed resource is a centralized server for gathering network status and terminal feedback. The measured QoE, is not considered in the adaptation technique.

2.5.2.4.5 Network congestion & measure QoE: This type of adaptation considers the user satisfaction (QoE) and the state of the network. Some context information are missed in the adaptation technique like user location, terminal capacity...

In the literature we notice some limitations in the existing work as follows: The method in [46] doesn't consider the dynamic variation of users' needs and the network resources optimization. The solution presented in [47] could not adequately enhance the user's experience since the media source is not aware of the context information. The presented method in [48] can be difficult to implement, because it is not easy to ask each user to implement his profile when he subscribes to a service. Some important context information is missed, for example in the proposed method in [49], the user localization and terminal capacity are not taken into account.

2.6 Parameters affecting QoE for video services

There are many parameters which impact the QoE for videos services. These factors depend on the quality of video source, type of device, characteristics and state of network, user's preferences ... We can generally classify them in the following groups:

2.6.1 The quality of video at the source

Parameters in the source can strongly influences the perceived quality:

- Frame rate: is how fast the content is moving. It is equal to the number of images/seconds. Example for fast moving (football, music video...), this value influences the observed quality. Even for slow moving (news...), the frame rate has to be greater than a minimum value.

- Type of service: The customer will have more expectations, if he paid service than if the considered service is free.

- Video resolution: Before sending data, the provider defines the video resolution (240p, 380p, 720p, 1280p ...) which corresponds to the number of pixels in the image. The greater this value be, the better the quality will be.

- Encoding bitrate: The encoding bitrate adversely affects the overall user experience because the encoding bit rate is proportional to the video quality.

2.6.2 The delivery of content over the network

Network conditions affects strongly the QoE:

- Packet loss: This parameter corresponds to packet loss during transmission. Depending on used transport protocol (TCP/UDP). For example in TCP case, packet loss may result in blocking of the video because the TCP protocol need acknowledgment to send next packet. For UDP, the consequences of packet loss are freezes.

- Jitter: It corresponds to packets arrival time. If this period is variable, it may influence the perceived quality.

- Delay: The delay is an very important parameter which can influence the perceived quality. The start up time (delay between clicking and the time when the video is played). Example for video delay superior to 10 s is not accepted and for conversational the delay should not exceed 150 ms.

- Bandwidth: The available bandwidth influences the user throughput. More this value is elevated more the perceived quality can be better.

2.6.3 User Context

It includes information about the user:

- User preferences: User interests are crucial parameters which can influence the QoE.

- Location: The user location is a parameter which influences the perceived quality. It is logical that watch video in house is more comfortable than in train or bus.

2.6.4 Device context

User can watch video services through different devices, which have different characteristics. The following devices parameters may influence the QoE of end-user's.

- Display size: User can access the multimedia content through various devices which may have different screen size. It is obvious that watch a movie in Ipad (resolution 1280x720) is better than in Iphone (resolution 340x240).
- CPU capacity: devices don't have the same CPU capacity, this parameter can influence the QoE.

2.6.5 Perceived parameters in the terminal

QoE is correlated to some parameters perceived at terminal level, and especially at video player side. These parameters include for instance:

- Video Start up Time: Delay before the beginning of the playback.
- Re buffering events: Video buffer starvation causes video freezing events.
- Video bitrate: It is the current video encoding profile.
- Buffering ratio: Defined as the ratio of the cumulative time, i.e. buffering duration, during which the video was buffering over the total video session duration.
- Fluctuant Video Quality events and distribution: Case of adaptive streaming delivery, where the video quality change according to available bandwidth.

2.7 Conclusion

In this chapter, we first give an overview on video services and architectures deployments, in particular for WebTV and IPTV services. After that, we explore existing video streaming techniques and compare them. A survey on content adaptation techniques considering the content adaptation and the adaptation of the delivery methods is presented. The content delivery means are reviewed and compared considering both operational solutions and research contributions. A state of art of methods measuring the QoE are also discussed.

However in general, there are lacks of systems, that are able to show precisely the quality metrics that are really perceived on user-side and also fully contextual methods making use of context information on the terminals, networks and contents for QoE assessments. In addition to that, there are needs to develop solutions for content and delivery adaptation based on measured QoE that is function of contextual parameters.

Consequently, we propose in the next chapter, to analyze, how quality metrics impact the user engagement in terms of user engagement (that can be considered as a metric of satisfaction).

Chapter 3

Understanding the impact of quality metrics on user engagement

3.1 Introduction

Content Providers are more and more interested in understanding the way their contents are consumed (usage) and appreciated (perception) by their audience. They usually rely on simultaneous delivery systems to deliver their contents, such as CDN, and Cloud Networks. To their purpose, a wide panel of methods can be implemented to catch useful information about User Engagement and playback information.

When relying on several heterogeneous delivery systems, it becomes difficult to get an aggregated view on the overall quality metrics. Nevertheless, quality is at the heart of Content Providers (CP's) revenues either under the form of direct revenues (e.g. subscriptions, content purchases...) or indirect ones (e.g. placement of advertisement banners and clip ...). Thus, assessing and improving both quality and usages appear naturally as both technical and business challenges to address. In most cases, they are handled via metrics and Key Performance Indicators (KPIs) which feeds monitoring (are our users satisfied? how long did they watch contents on our platform?...) and/or decision tools (which content shall i propose to this user? under which format? ...).

On their side, Internet Service Providers (ISPs) and Mobile Network Operators (MNOs) are also interested in getting the same information regarding content services (their own services, 3rd party or OTT services...) which are delivered to their clients through their access networks. By doing this, an operator tries to know how his networks are performing in from of his competitors' ones, and possibly to gain visibility through benchmarks (e.g. regulatory or private rankings...).

Monitoring these metrics appears for them as a complementary means to the real-time QoS monitoring methods to improve the overall quality of their networks.

In this context, many companies (see for instance [61][62]) are now selling products for monitoring quality metrics (APIs, network probes...). Their services require slight modifications on the video player.

In [63], authors investigated the impacts of video startup time on user's perceived QoE by dealing with subjective tests on the user satisfaction of different applications (Youtube video streaming, wireless Internet connection setup and social networks authentication). They showed that the perception of video startup time depends on the considered application. For Youtube video streaming service, up to 30 seconds, end-users were satisfied. For the setup in the wireless Internet connection, that corresponds to the delay from pressing the button connection and the successful connection establishment, authors found that, the end-users perception was acceptable up to 15 seconds. However for social networks authentication, users were more demanding, a video startup time of 8 seconds led to a MOS value of 2, which was considered as poor.

In [64], authors measured quality metrics such as the video startup time, buffering ratio, average bitrate, rendering quality and rate of buffering events. They mainly focused on correlation between the user engagement and those metrics. They indeed demonstrated that the buffering ratio is the most critical parameter in the case of live and VoD services. In particular, authors found that a 1% increase in the buffering ratio can reduce user engagement by more than 3 minutes for a 90 minutes video live event. But, they did not find any correlation between the average bitrate and the user engagement.

The method proposed in [63] measured the end-user's perception by using client-side log data captured directly from the video player in user's terminal, which enabled understanding the impact of QoS parameters and user engagement.

There are many studies related for monitoring and reporting technical quality metrics, but the impact of these parameters on user engagement is not well investigated. In this chapter we propose to answer the questions: how the user engagement in terms of video play time varies with startup time, buffering ratio, average bitrate and buffering events for live popular and unpopular contents events?

In our case, we tested quality measurements with a developed prototype, that is able to retrieve metrics on client side. The quality metrics monitoring was realized on a single event, Roland Garros 2013, that is an international tennis tournament occurring every year since 1928, in Paris, France [65]. This event is the second prestigious tennis competition in the world after Wimbledon and the most watched. This opportunity allowed us to begin an analysis of contents usage and quality metrics in the case of specific events.

3.2 User Engagement

The user engagement study the interactions between users and services. In [66], the user engagement is defined as "The emotional, cognitive and/or behavioral connection that exists, at any point in time and over time, between a user and a technological resource". QoE is a subjective measure, the user engagement can be measured objectively. In [67] authors propose the following indicators:

- Click-Depth Index : It corresponds to the page and event views.
- Duration Index : Time spent by users in the video platform or website.
- Recency Index : As the number of time the user has visited the platform.
- Loyalty Index : Level of long-term interactions the user has with the site or product (frequency).
- Brand Index : Apparent user's awareness of the brand, site, or product (search terms).
- Feedback Index : Qualitative information including propensity to solicit additional information or supply direct feedback.
- Interaction Index : User interaction with site or product (click, upload, transaction).

These are simple and effective measures that can easily assess user engagement. These measures allow to characterize the quality of video streaming. In this thesis, as we analyze the quality metrics of video streaming contents, the duration index or video playtime appears to be a good indicator to represent the user engagement and can be measured objectively. From a business point of view, this metric is particularly interesting for Content Providers and ISPs who want to measure the audience and analyze the quality metrics and its impacts.

3.3 Event description and data collection

In this section, we briefly introduce the Roland Garros event, the developed prototype and how data were collected on the client side.

3.3.1 The Roland Garros Event

Roland Garros is a broadcast live event. Among the available audiovisual sources, the Roland Garros 2013 event is aired on TV, but also on the Internet. Actually, Roland Garros channels were accessible for all Francophone Internet users including France, Dom-Tom (French overseas departments and territories) and Monaco for this trial. The 2013 edition spread over 2 weeks (from the 05.27.2013 to 06.09.2013). Each tennis court is displayed on a dedicated streaming channel.

3.3.2 Channel delivery

Two solutions were used for channels delivery:

- An origin service platform located in Paris composed by two centralized servers.
- A CDN, with distributed servers in 15 regional locations in France. Those two solutions are able to deliver the Roland Garros channels for the Internet broadcast. The channels were available HAS techniques and derived in SD and HD profiles (SD profiles ranging from 342 to 1340 kbps , HD profiles ranging from 1910 up to 2860 kbps) [7].

3.3.3 Prototype and Collected Data

In this section, we describe the platform/system used to measure in real-time the quality of video streams (or channels) received by the end-user on the Internet, and to analyze later the aggregated data. The developed solution is depicted in Figure 3.1 and the collection process is presented in Figure 3.2.

On the client side, a java script code is inserted to the video player loaded by the browser to retrieve metrics during the users' video session. On the server side, a centralized module periodically collects and stores data received from the client-side component. Furthermore, data analytics are performed by another standalone module.

The prototype gathers a number of metrics, the most important being:

- Video bitrate, i.e. the current video encoding profile,
- Startup time, i.e. the delay between users click on playback button (command) and start (response).
- Number of buffering events, i.e. number of times when the player's buffer lacked of data forcing the player to freeze,
- Buffering duration, i.e. the cumulative time during which the video was buffering,
- Video play time, i.e. the total playback duration for a given end-user (i.e. the total session duration restricted to video playback). At peak, the prototype was able to monitor and analyze 50k sessions.

The whole system has been developed with open-source web technologies and is shown in Figure 3.1.

- Agent : on client side, sending quality metrics.
- Collector : entity in charge of receiving and collecting events reports and measures.

- Storage : persistent oriented document database (noSQL), mongo DB, providing a secure and scalable storage.

- Evaluator : this module is in charge of data and sessions aggregation & accounting and make analytics reports (dashboards).

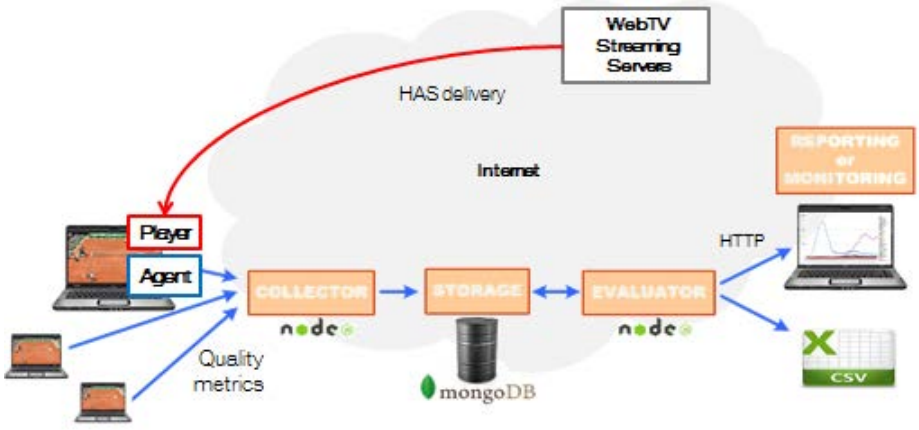

Figure 3.1: Solution Overview

Figure 3.2: Overview of the collected process

3.4 Results and analysis

In this section, we present results regarding the effects of video quality metrics on user engagement, during Roland Garros. In our analysis we focus on the impact of startup time, buffering ratio and average bitrate on the video play time. However, regarding buffering events, we decided not to report the impact of this metric on user engagement, because nearly 75% of all sessions had less than 2 buffering events, meaning it was not significant enough to perceive a possible correlation between both metrics (see Fig.3.6).

3.4.1 Impacts of quality metrics on user engagement for popular contents

In this section, we take the example of one of the most popular games. We observed a peak in service load during this game likewise other matches occurring on the Chatrier court. The popularity of this game can be characterized by:

- 40% higher compared to other contents.
- Channel hopping (switching between channels) was less frequent and also sessions duration was 30% higher in average than for less popular contents.

In following paragraphs, we show the Cumulative Distribution Functions (CDF) and the correlation between quality metrics and the User Engagement, associated to quality metrics for popular contents.

3.4.1.1 Distribution of quality metrics

In this section, we analyze the distribution of quality metrics in terms of CDF, that describes the probability of observing a value less than "probability of observing a value less than a given threshold". In other words, the CDF captures the statistical/probabilistic distribution of observations/measures [68].

3.4.1.1.1 Distribution of video startup time: The CDF of the video startup time shows that 64% of sessions had of sessions started in less than 1 second, meaning that more than half of the sessions started almost instantly (see Fig. 3.3).

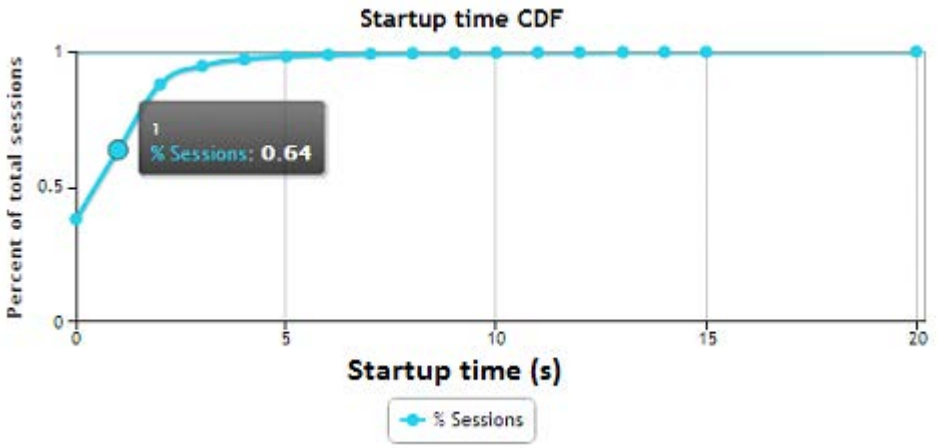

Figure 3.3: Cumulative Distribution Function for startup time

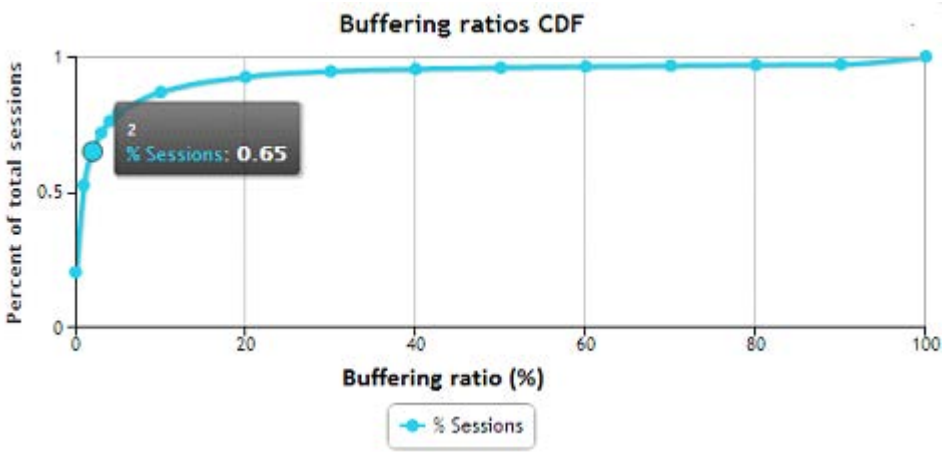

Figure 3.4: Cumulative Distribution Function for buffering ratio

3.4.1.1.2 Distribution of buffering ratio: The buffering ratio defined as the ratio of the cumulative time, i.e. buffering duration, during which the video was buffering over the total video session duration. It is commonly expressed as a percentage. According to [64], are considered as "impacted by buffering" when the buffering ratio exceeds 5 seconds or 2% of the total session duration. Following this latter threshold, we saw therefore that 45% of sessions were affected by buffering (see Figure 3.4).

This result was quite surprising at first because, we observed that buffering occurred on all sessions, whatever the average bitrate range and whatever the networks (Autonomous Systems) involved. This told us that neither the network (nodes and links) nor the servers delivering content (content delivery network and central platform) couldn't be the cause of buffering, but instead the player on

client side. Indeed, as many sessions lasted a few seconds (mainly due to channel hopping), if a short buffering event happened on these sessions, the number of impacted sessions would have quickly increased.

3.4.1.1.3 Distribution of Video bitrates: In this section, we focus on the distribution of the average bitrate, where averaging is performed over the life of each single session, Figure 3.5 shows that 80% of sessions had an average bitrate above 1100 kbps. We also noticed that a high number of sessions are impacted by buffering ratios and only a small number of sessions had higher average bitrates (greater than 1100 kbps).

Indeed, the most requested bitrate profile was 977 kbps. This also confirmed that network connectivity was acceptable for a wide range of users, as it offered more than 1 Mbps and low startup time (less than 1 second).

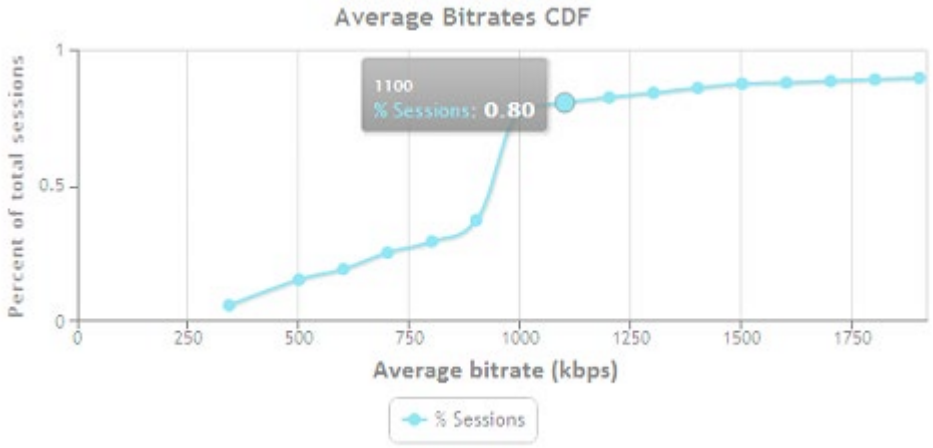

Figure 3.5: Cumulative Distribution Function for average bitrate

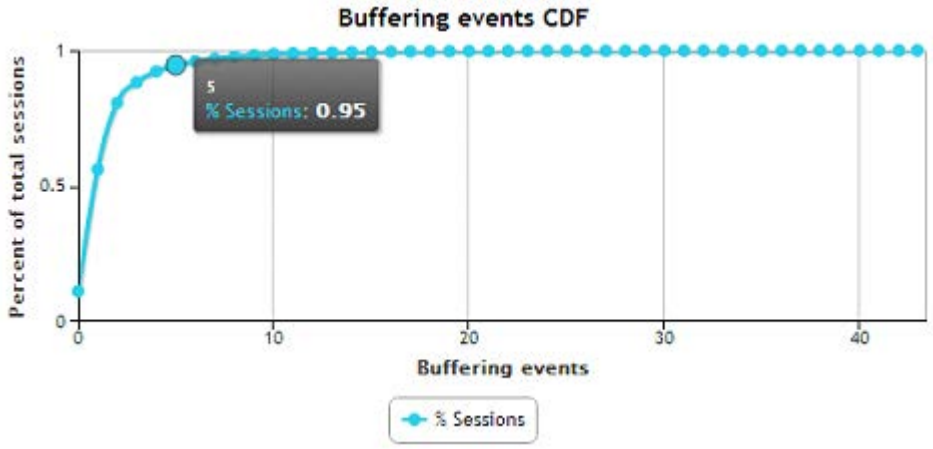

Figure 3.6: Cumulative Distribution Function for buffering events

3.4.1.2 Correlation between quality metrics and user engagement

In this section, we analyze how quality metrics are correlated with the user engagement (here materialized through the mean playing time). The mean playing time corresponds to an average of end-users playing time.

3.4.1.2.1 Correlation between startup time and user engagement:

From Figure 3.7, no clear trend has been observed between the video startup time and the mean video play time. We could therefore infer that up to 10 seconds, the video startup time did not disturb the viewers. As a matter of facts, users were ready to wait for their session to start. A reason for this could be that live sport events like Roland Garros have the singularity to be broadcast once a year. That's why they are likely to catch the attention of more users than casual sports programs like weekly football matches. Another reason could be, in general, users are engaged for long streams, then they can wait few seconds before the beginning of the show.

As previously said in [69], the viewer's sensibility to video startup time depends on the considered application, in our case (live sport event), it was found that the sensibility was indeed different than the considered applications in this work (see Fig. 3.7).

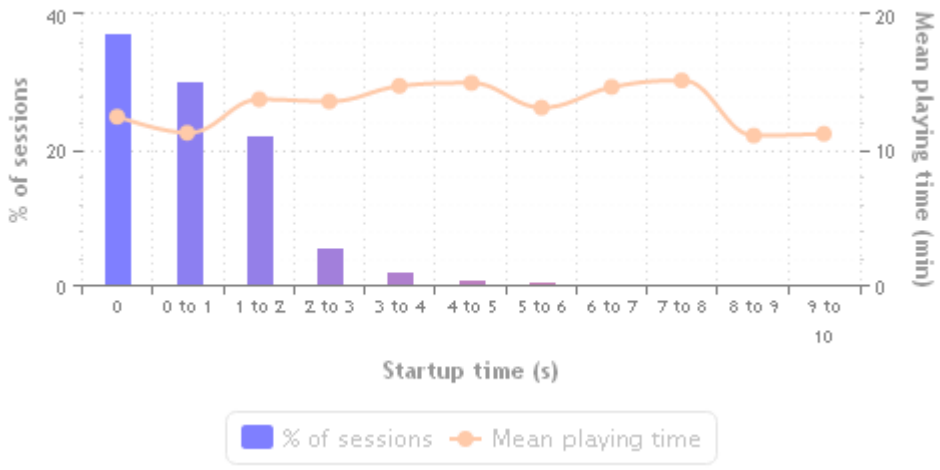

Figure 3.7: User engagement versus startup time

3.4.1.2.2 Correlation between buffering ratio and user engagement: Regarding buffering ratio, we observed a correlation between the buffering ratio and the user engagement (see Fig.3.8). In this case a jump of 1% buffering ratio induced a drop of the user engagement from 37 min to 12 min. In other words, a 1% gain in buffering ratio induced a difference of 25 min.

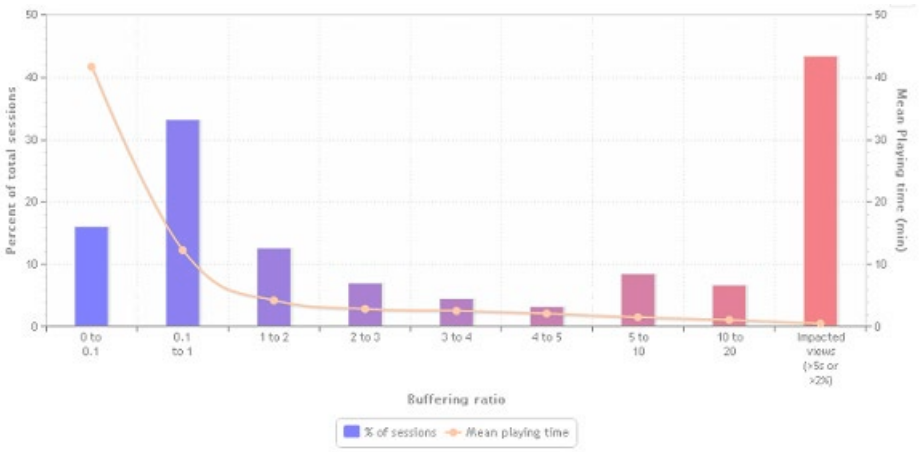

Figure 3.8: User engagement versus buffering ratio

3.4.1.2.3 Correlation between Average bitrate and user engagement: According to Fig.3.5, we observed a correlation between the video play time and average bitrate. We did notice that the user engagement increased when the session average bitrate reached a threshold near [1100-1100] kbps. Beyond this threshold, the correlation between average bitrate and user engagement remains almost stable. The user's behavior after the threshold could indicate that users' were not able to differentiate the video quality after a certain level. Increasing the video quality after this point did not increase the user engagement.

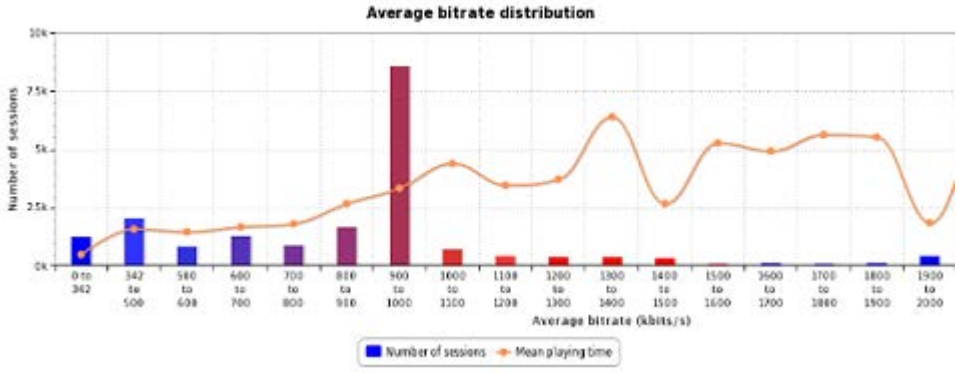

Figure 3.9: User engagement versus Video bitrate

3.4.1.2.4 Mathematical correlation between metrics and user engagement: Fig. 3.10, 3.11 and 3.12 depicts a scatter plot of the video play time as a function of several application metrics for each session. Table 3.1 shows values of the Kendall tau rank which is a statistic measure to quantify the relation between two variables [70] . Unlike the Pearson correlation coefficient, this coefficient does not make the assumption that the relationship between parameters is linear [71].

Applied to our context, Kendall correlation coefficient confirms that buffering ratio is a critical metric that may negatively impact the User Engagement, but for video startup time and average bitrate Kendall coefficient reveal less correlation.

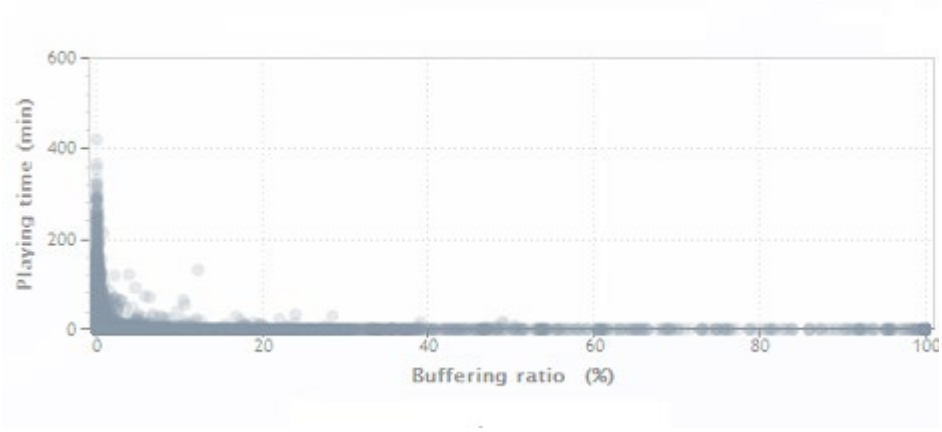

Figure 3.10: Scatter plots for playing time versus buffering ratio

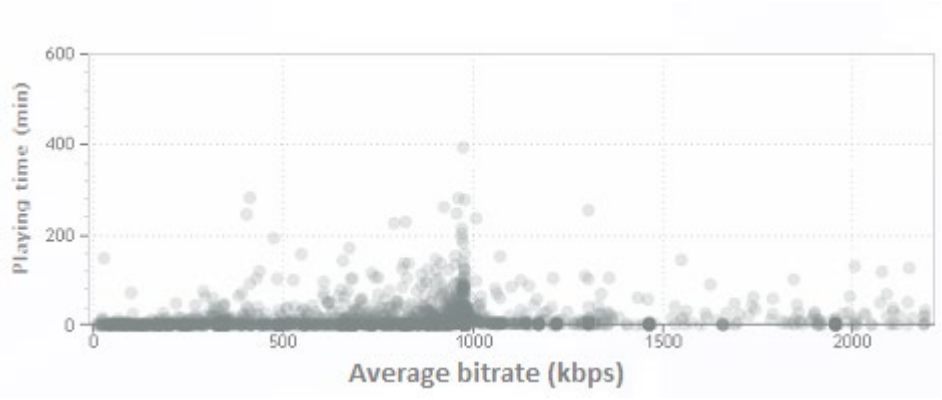

Figure 3.11: Scatter plots for playing time versus Video bitrate

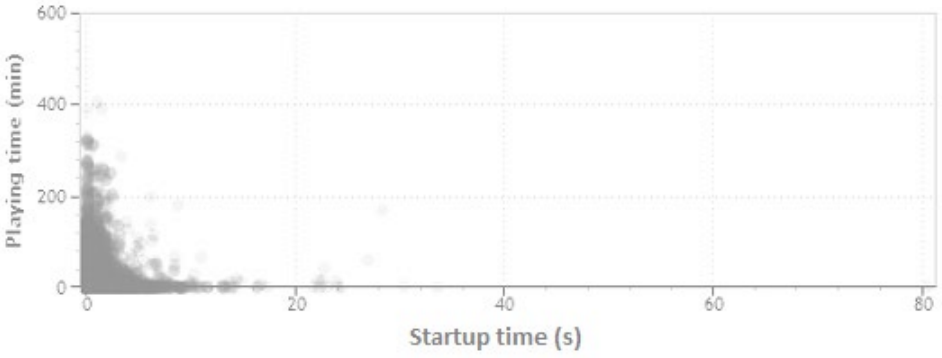

Figure 3.12: Scatter plots for playing time versus startup time

Metrics	Kendall Coefficient correlation w.r.t user engagement
startup time	0.04
buffering ratio	-0.57
Average bitrate	0.23

Table 3.1: Kendall coefficient shows a correlation between two variables

3.4.2 Impacts of quality metrics on user engagement for unpopular content

This section shows in brief the impact of QoE metrics for unpopular content. We consider as an example of unpopular content, the ladies'3rd round (Sharapova-Zhang) on Philippe Chatrier court. This match attracted far less people compared to the other matches, and from the technical point of view, we observed more frequent channel hopping phenomena together with a shorter session duration (on average 30% less) than for popular games.

We still note that even if the content is not popular, the buffering ratio remains a critical parameter that may negatively influence the video play time. However, there is no trend between the video play time and others metrics (average bitrate and video startup time). As a matter of fact, the content popularity changes the users' expectations. as such, we conclude that the content popularity is a crucial parameter to take into account in the quality impacts analysis on user engagement, mainly because the user behavior might not be as predictable as for popular content.

3.4.3 Correlation between quality metrics

Quality metrics mutually independent between them. In this paragraph, we analyze the interaction between some of these metrics.

3.4.3.1 Video bitrate and buffering ratio

In order to reduce buffering , HAS algorithms adapt the Video bitrate based on the current network status as evaluated by the player on the user's device. As shown in Figure 3.8 , users are very sensitive to buffering, and then reducing this parameter can improve the user engagement. In the following Figure 3.13, we can see that the average bitrate is correlated with buffering ratio. The Kendall coefficient is $\sim$ - 0.20 which indicates without much surprise that both metrics are negatively correlated.

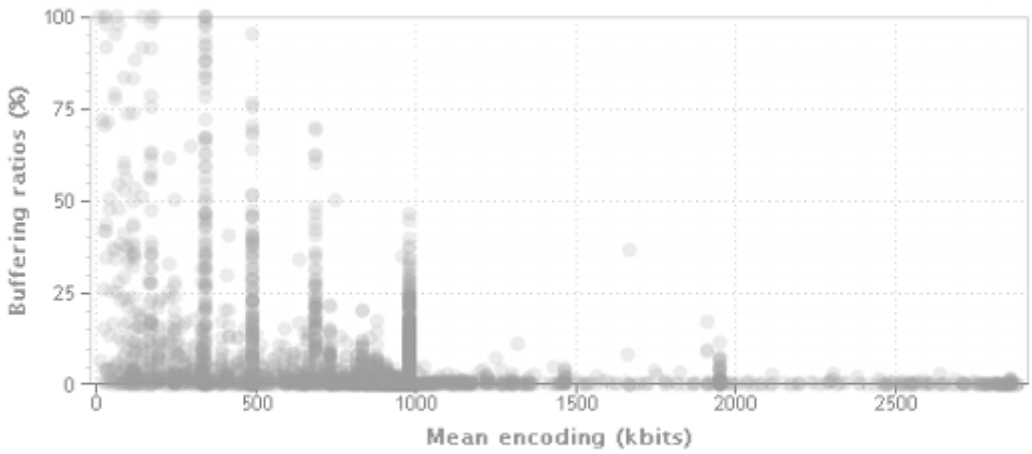

Figure 3.13: Average bitrate versus buffering ratio

3.4.3.2 Video bitrate and video startup time

According to [72], higher video quality would mean higher video profile, because the video player would take longer to buffer at session start. However, this fact was not confirmed in our tests as we did not notice any correlation between the Video bitrate requested by the player at start and the startup time. It can mean that the requested bitrate profile by the player at session start is always adapted to the network bandwidth capacity between the end-user and the server.

3.5 Conclusions

In this chapter, we analyzed the effect of quality metrics (buffering ratio, video startup time, and average bitrate) on user engagement for popular and unpopular contents during Roland Garros tennis championship event.

In our analysis, the buffering ratio appears as a critical parameter which impacts the video play time whatever the content type.

Additionally, the Video bitrate is another critical parameter which impacts the user engagement up to a certain threshold in the case of popular contents. We find also that video startup time is less significant until a certain level for popular contents.

This results prove also that analyzing quality metrics and their impact on user engagement is complex. As we noticed, the video playtime may depend on quality metrics (these metrics depend on network conditions), but also on more subjective parameters like the user behavior and their expectations regarding the contents.

As the study shows, a correlation between quality parameters and the user engagement has been observed. However, even if gigabytes of information were

gathered and analyzed over thousands of clients, some major questions remain open. One among them is which factor is the most preponderant on the user engagement? what is the impact of device context, user context, content context on end-user video perception?

It is true that user engagement in terms of video playtime can help us to characterize the end-users behavior based on the state of the service. But this parameter does not reflect exactly the end-user perception and how to optimize it. Some investigations are proposed in the next section to address this problem.

Chapter 4

Experimental evaluation of QoE and proposed QoE models

4.1 Introduction

Quality of Experience is a determinant factor for the end-to-end evaluation of applications and services. Being able to understand human requirements in terms of quality and expectations is at the very heart of any business. However, the human behavior is subjective, random in nature and and varies as a function of the environment and context.

In this chapter, we investigate the influence of the most trivial contextual characteristics for video playback on QoE :

- Content (encoding and nature)
- The network
- The device type

In our case, we made the choice of addressing QoE qualitatively through the concept of MOS. The MOS is the most famous metric used in subjective measurement, where subjects are required to give a rating according to some predefined scheme (e.g. see Table 2.4).

These investigations were conducted both experimentally and theoretically:

- An experimental framework is introduced at first so that to collect reference information and rating coming from real users
- Theoretical models for MOS are derived on the top of experimental data

From a technical scope point of view, in our study, we first analyze the impact of network degradation on QoE and in the second one, the effect of video bitrate on QoE. For each study, we consider a variation of device type and content type.

4.2 Effect of network conditions over QoE

In this section, we investigate experimentally the effect of network conditions (mainly bandwidth) on end-users' QoE. In order to put the stress on the need of accounting for contextual information, we included secondary inputs such device type and content type in the experimental design.

4.2.1 Experimental framework

We reproduce a typical user sphere with a desktop PC, a smartphone and a server (see Fig. 4.1). While the server is used for streaming video content, the desktop PC and the smartphone are used for video playback (control and display).

The mobile phone is Samsung Galaxy S3, 1280x720pixels, 4.8" (12cm). The desktop PC is equipped with a 17" LCD screen with a 1280x960 pixels resolution.

The server runs a fictive video portal developed in HTML5 [9] (implemented over the Apache 2 HTTP server) through which contents are selected and streamed. Network emulation (here, bandwidth limitation) is directly performed on the HTTP server by means of the mod_bw module.

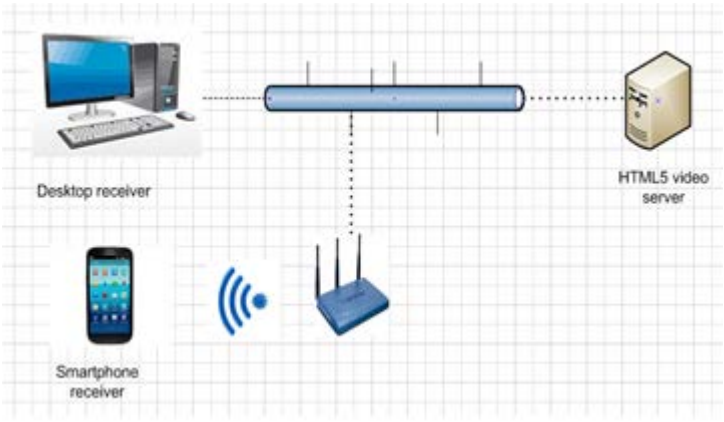

Figure 4.1: Experimental platform

4.2.2 ITU-T recommendations for testing video services

The ITU-T recommends using a five point quality scale for subjective evaluation of perceived video quality. On this scale, a bad quality is associated with the mark 1, while a good quality is associated with the mark 5. In addition to using this standard scale, we also made the choice of using the Single Stimulus Continuous Quality Evaluation (SSCQE) (see [73]), which consist in evaluating quality after each single content playback. Users were asked to give their opinion for each test as described in Table 4.3.

The ratings for each configuration are then averaged over all subjects to obtain a Mean Opinion Score (MOS). After choosing their quality rating, users had to confirm their perception by answering some qualitative questions to understand what had impacted their choice.

The test equipment selected is representative of the type of device used to watch audiovisual content (desktop and smartphone).

As requested in ITU-T Recommendation the used videos for testing must have different characteristics (spatial and temporal information), we took this into account when dealing with different content types. The video duration is representative of the typical scenario of watching short video in WebTV or OTT platforms (YouTube, Dailymotion). The video content type and description are summarized in Table 4.1 below.

Mutimedia Content	Genre	Description
Sport	Football	Football Match Barcelona FC - Real Madrid
Music	Video Clip	Music video clip (Psy - 'Gangnam Style')
News	TF1 news	TF1 journalist reading news story and some sequences reports
Animation	Cartoon	3D animation movie - Big Buck Bunny (Peach Open Movie Project)

Table 4.1: Description of test sequences

4.2.2.1 Preparation of tests sequences of video contents

The duration of tests sequences is important to capture user's perception [74]. In previous studies, the duration varied between 8 to 30 seconds [75]. These times are not sufficient to get a significant feedback from users. According to studies in [76] and [77], the video duration length in mobile phone is between two to five minutes.

In our tests the video duration is four minutes which is sufficient to have user's feedback about their perception for viewing videos in mobile and desktop environment. Each video sequence was encoded by using FreewebM converter, for each terminal (Desktop PC and Smartphone). Considered bitrates are defined by a video resolution, the video encoding bitrate, video codec, audio codec and frame rate. Table 4.2 below summarizes used bitrate characteristics.

4.2.2.2 Description of experimental procedures

The idea is to gradually decrease the values of network throughput to determine their perception function of used device with different content types as described in Table 4.2. The considered model will answer these questions :

bitrate Characteristics	Smartphone	Desktop
Video encoding bitrate(kbps)	1100	2100
Audio encoding bitrate (kbps)	128	128
Frame rate (fps)	25	25
Video resolution	640 x 360	854 x 480
Video Codec	VP8	VP8
Audio Codec	Orbis	Orbis
Video duration (mn)	4	4

Table 4.2: Used bitrate characteristics

- Human perception of quality changes with respect to change in throughput parameters.

- Influence of content types (fast moving, slow moving ...) in human perception.

- Influence of device types on human perception.

After watching videos in mobile and desktop terminals, subjects were asked to evaluate their Quality of Experience according to the ITU-T scale (cf. section 2.4.2). Each test corresponds to a different throughput upper limit for both device types (see Table 4.3). For instance, Test 1 corresponds to a throughput limit defined as 200% of the video bitrate (i.e. twice the minimal network bandwidth). Test 1 is the highest throughput, while Test 6 the lowest setting which is equivalent to 20% of the video bitrate. Table 4.3 summarizes combinations for different tests.

Tests	% of video bitrate	Throughput (kbps)/Smartphone	Throughput (kbps) Desktop PC
Test 1	200	2200	4200
Test 2	120	1500	2500
Test 3	100	1100	2100
Test 4	70	700	1400
Test 5	40	400	800
Test 6	20	220	420

Table 4.3: Considered tests combinations

The experimentation was conducted with 79 subjects: 20 females and 59 males. Tests were performed with different kinds of audience, by undergraduate students in Paris, experts engineers in video services and the rest of subjects are naive users (they are not in the telecom field). Before video watching, users could express their profile (name, occupation, gender, age).

4.2.3 Results and analysis

In this section we present the experimental results, provide a first-stage analysis and discuss some findings.

The results obtained for each experiment condition are then consolidated across testers through statistical indicators: an average gives us a Mean Opinion Score (MOS for short) for each content type, device and network conditions, while the standard deviation captures the spread around the MOS due to the subjectivity and irrationality of human perception and expectations.

4.2.3.1 The case of desktop environment

In this section, we analyze user satisfaction as a function of network throughput, content type and device type. Fig.4.2 depicts the subjective scores for the different content types and throughputs (see Table 4.3). We plotted the average of MOS marks for each category, together with their standard deviations.

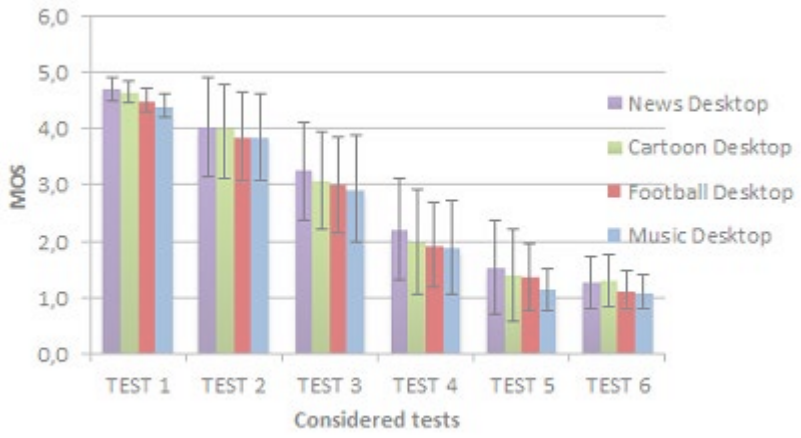

Figure 4.2: QoE for desktop with standard deviation

We can first observe from Figure 4.2 that users were more tolerant to phenomena driven by network shortage (increase in video playback start, increase in the number of video stalls ...) for contents like news & cartoon, than for football or music. We explain that by the fact that the level of temporal motion is higher for football and music clips than for news and cartoons.

Figure 4.3 shows the distribution of scores for all tests combined without considering the content type. In legend of this figure, PNx stands for "Percentage of Notes equal to x", where x ranges from 1 (i.e. "bad") to 5 (i.e. "excellent"). The proportion of unsatisfied users (bad marks) increases as network throughput is decreasing, i.e. as startup time and buffering events are decreasing with through-

put. For high throughput, good scores have good proportions; anyway we still have non-zero bad scores due to some users having high expectations. For low network throughput, users are less tolerant due to increase in video playback start and an increase of the number of buffering events: they don't have the opportunity to enjoy the video even if the quality of video is good.

We represent also in Figure 4.4 and Figure 4.9, the box plots respectively for music and cartoon from our subjective tests for different contents/devices types.

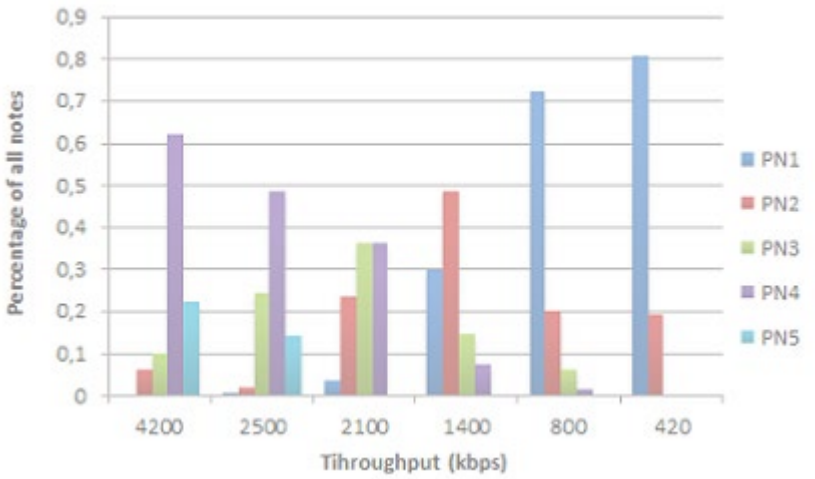

Figure 4.3: Scores repartition for desktop

Figure 4.4: Music box plot in the case of desktop

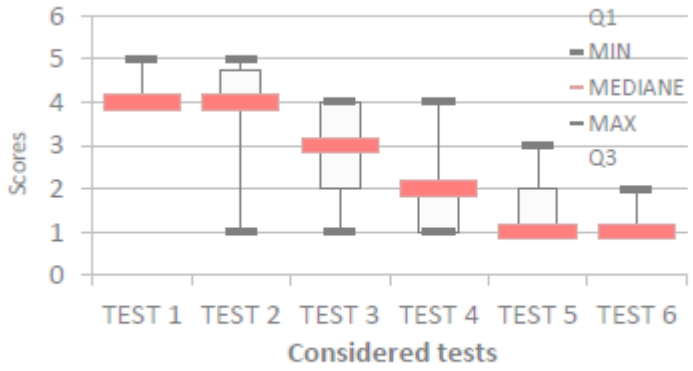

Figure 4.5: Cartoon box plot in the case of desktop

4.2.3.2 The case of smartphone environment

In this section, we perform the same analysis than the precedent, but in the smartphone environment. Figure 4.6 depicts the average and standard deviations of MOS marks for each test category (content type and throughput).

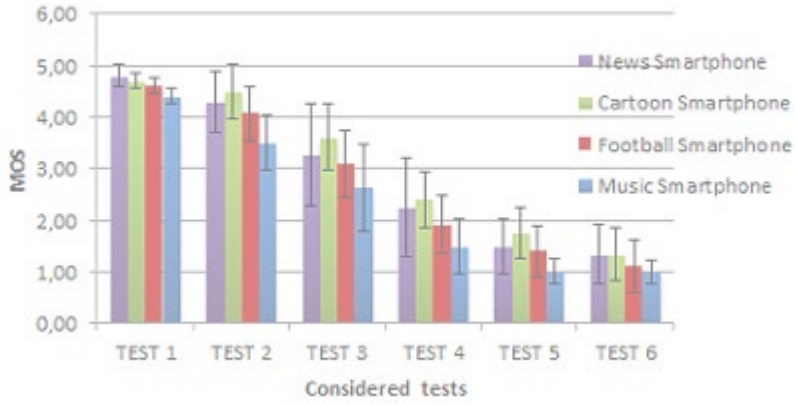

Figure 4.6: QoE for smartphone with standard deviation

As we can see from this Figure, users' behavior in terms of satisfaction is the same than for the desktop environment: end-users are more demanding for video having high temporal motion like football & music clips than cartoons & news. However, we also notice that end-users expectations in terms of quality are less important on the smartphone than the desktop for news, football and sports. Conversely, end-users are more demanding for music clips on the smartphone

than the desktop. We explain that by the lower quality of audio playback on smartphone speakers.

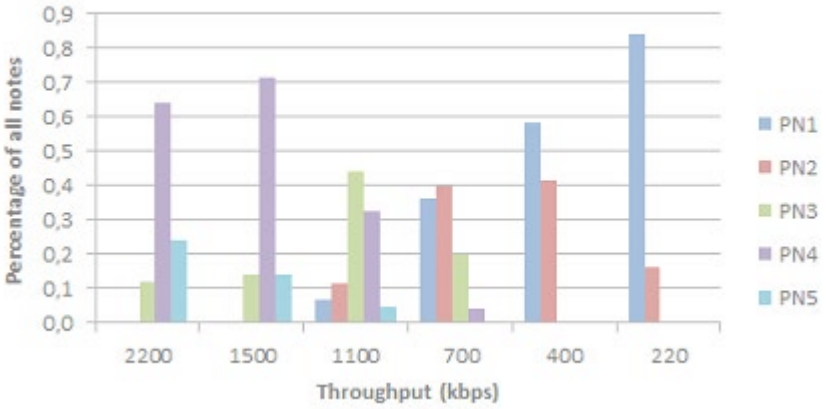

Figure 4.7: Scores repartition for smartphone

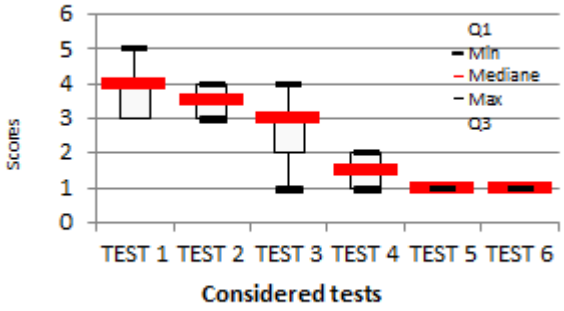

Figure 4.8: Music box plot in the case of smartphone

Figure 4.2 shows also that content type is more impacting in the case of smartphone (significant ramp-up phenomena across content types inside a single throughput condition). Figure 4.7 shows the distribution of scores for all tests combined without considering the content type. As in the desktop environment, the proportion of unsatisfied users (bad marks) increases as network throughput is decreasing. In the case of smartphone, we represent in Figure 4.8 and Figure 4.9, box plots respectively for music and cartoon as in the precedent section from our subjective tests. These figures allow us to see the distribution of scores.

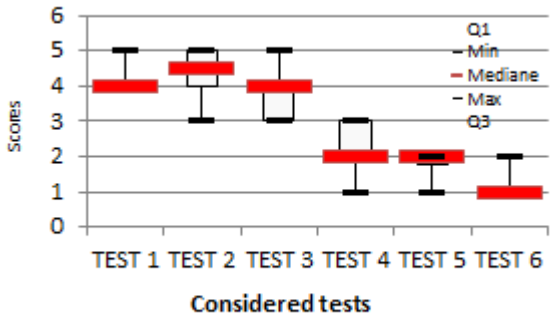

Figure 4.9: Cartoon box plot in the case of smartphone

4.2.4 Proposed Network Contextual User Perception Model

Now that we have the subjective test results, we can tailor a mathematical model to fit with our context, so as to reflect the results obtained from our extensive field test evaluations. Different relationships (linear, exponential and logarithmic) between the used device, the video content type and the quality of the link will be analyzed in terms of regression. The chosen model will be evaluated through the following coefficient of correlation r.

$$r = \frac{cov(x,y)}{\sigma_x \sigma_y} \tag{4.1}$$

where :

$$cov(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x_i})(y_i - \bar{y_i}) \tag{4.2}$$

$$\sigma_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x_i})^2} \tag{4.3}$$

$$\sigma_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y_i})^2} \tag{4.4}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4.5}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{4.6}$$

where

- $cov(x, y)$ corresponds to the covariance of x and y, where x and y correspond to data observations.
- $\bar{x}$ and $\bar{y}$ are respectively mean of variables $x_i$and y_i

Content Type	Parametric Models	Pearson Coefficient of Correlation (PCC)
Football	$1.08 + 6.8\, e^{-1.23 \frac{Dv}{Dri}}$	0.984
	$3.24\text{-}1.61 \ln(\frac{Dv}{Dri})$	0.912
	$4.05\text{-}0.70 \frac{Dv}{Dri}$	0.646
Music	$1 + 7.88\, e^{-1.67 \frac{Dv}{Dri}}$	0.991
	$2.76\text{-}1.46 \ln(\frac{Dv}{Dri})$	0.857
	$3.43\text{-}0.61 \frac{Dv}{Dri}$	0.600
Animation	$1.25 + 5.87\, e^{-0.94 \frac{Dv}{Dri}}$	0.975
	$3.60\text{-}1.62 \ln(\frac{Dv}{Dri})$	0.937
	$4.45\text{-}0.73 \frac{Dv}{Dri}$	0.703
News	$1.25 + 6.65\, e^{-1.17 \frac{Dv}{Dri}}$	0.988
	$3.46\text{-}1.63 \ln(\frac{Dv}{Dri})$	0.919
	$4.28\text{-}0.71 \frac{Dv}{Dri}$	0.704

Table 4.4: Network parametric models case of smartphone

Each model is adjusted by means of least squares regression, which consists in minimizing the average of squared by using the Mean Square Error (MSE) measures. MSE measures the average of squares of the error, that is the difference between the estimator (using the associated model) and experiments values:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\tilde{X}_i - X_i)^2 \tag{4.7}$$

where:

- $\tilde{X}_i$ is the prediction of X_i for user i
- X is the vector of true values
- n is the total number of samples

Content Type	Parametric Models	Pearson Coefficient of Correlation (PCC)
Football	$1.10 + 6.54\, e^{-1.25 \frac{Dv}{Dri}}$	**0.991**
	$3.15\text{-}1.54 \ln(\frac{Dv}{Dri})$	**0.918**
	$3.92\text{-}0.67 \frac{Dv}{Dri}$	**0.698**
Music	$1.00 + 6.54\, e^{-1.23 \frac{Dv}{Dri}}$	**0.989**
	$3.08\text{-}1.55 \ln(\frac{Dv}{Dri})$	**0.908**
	$3.84\text{-}0.67 \frac{Dv}{Dri}$	**0.681**
Animation	$1.23 + 6.95\, e^{-1.34 \frac{Dv}{Dri}}$	**0.969**
	$3.26\text{-}1.55 \ln(\frac{Dv}{Dri})$	**0.895**
	$4.00\text{-}0.66 \frac{Dv}{Dri}$	**0.661**
News	$1.24 + 6.34\, e^{-1.16 \frac{Dv}{Dri}}$	**0.995**
	$3.36\text{-}1.57 \ln(\frac{Dv}{Dri})$	**0.935**
	$4.15\text{-}0.69 \frac{Dv}{Dri}$	**0.724**

Table 4.5: Network parametric models case of desktop

In Table 4.5 and Table 4.4, we find regressions between MOS and available resource characterized by (D_{ri}), by taking into account the type of device (desktop or smartphone), content type (news, cartoon, music and football) and the considered video bitrate (D_v) that is constant, in order to let end-users to focus on the degradation caused by the network.

In the case of desktop and smartphone, the exponential model has higher Pearson Coefficient of Correlation, (PK) ≈ 0.9. We propose to model the degradation caused by the network, by the parametric model below, that we will call, the Network Contextual User Perception Model (NCUPM).

$$\varphi_{N_i}(D_{ri}) = \alpha + \beta e^{-\delta \frac{D_v}{D_{ri}}} \tag{4.8}$$

Where :

- $\varphi_{N_i}(D_{ri})$ is the score caused by network N characterized by D_{ri}
- α , β and δ are the model parameters calculated by using subjective test data from different experiments and are presented in Table 4.5and 4.4.

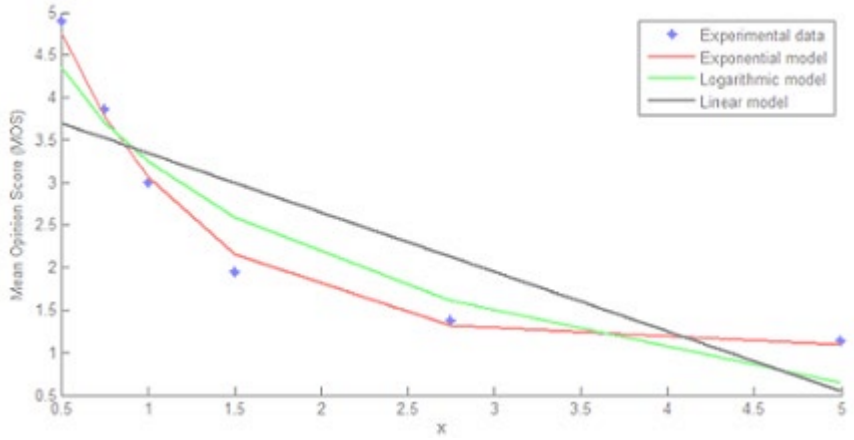

Figure 4.10: Different regressions for network in the desktop for football

In Figure 4.10 and Figure 4.11, in order to show the higher correlation for exponential model, we represent test data as an average for each test and different regressions model (exponential, logarithmic and linear) in the case of football. The experimental data corresponds to the average of MOS marks for each test. We remark that there is a higher correlation between experimental data and the exponential data for different desktop and smartphone.

The parameter x in Figures 4.10 and 4.11 is defined as :

$$x = \frac{D_v}{D_{ri}} \tag{4.9}$$

where : D_v=Considered video bitrate and D_{ri} the network throughput for user i

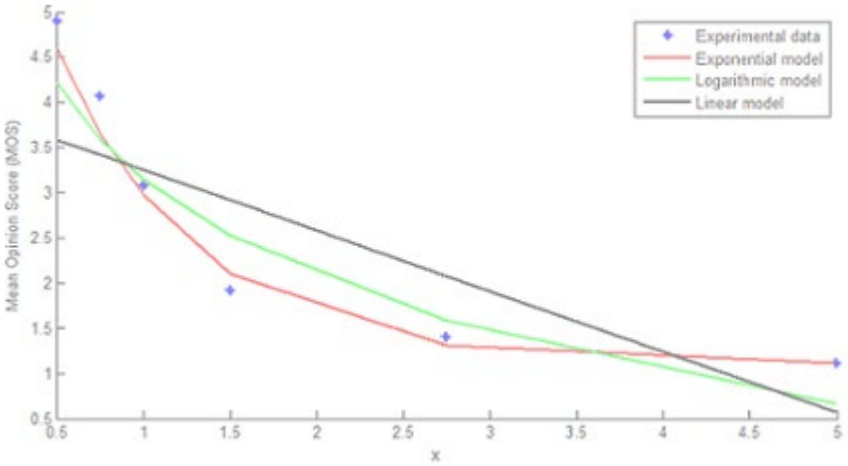

Figure 4.11: Different regressions for network in the smartphone for football

4.3 Impact of startup time and buffering on QoE

In this section, we analyze the impact of startup time and buffering on Quality of Experience. However another set of important perceptual parameters (video startup time and the buffering ratio) were collected during experimentation. In parallel, after each watch, users were asked to rate the video startup time and buffering ratio during the experimentation.

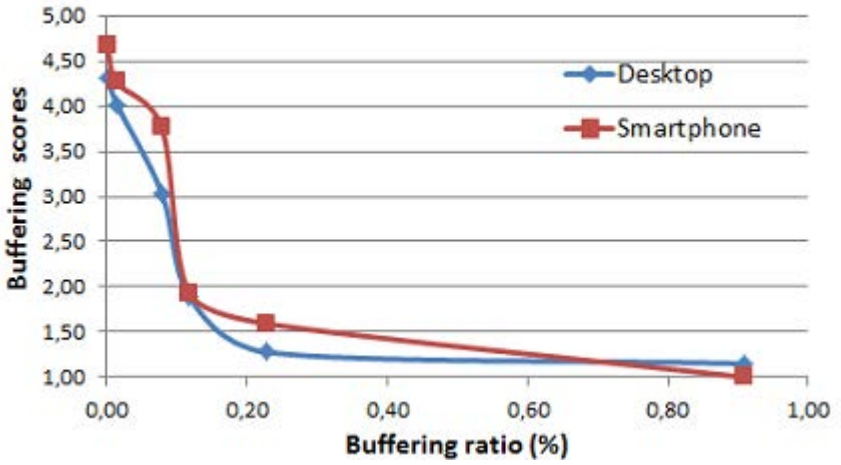

Figure 4.12: General buffering ratio rating

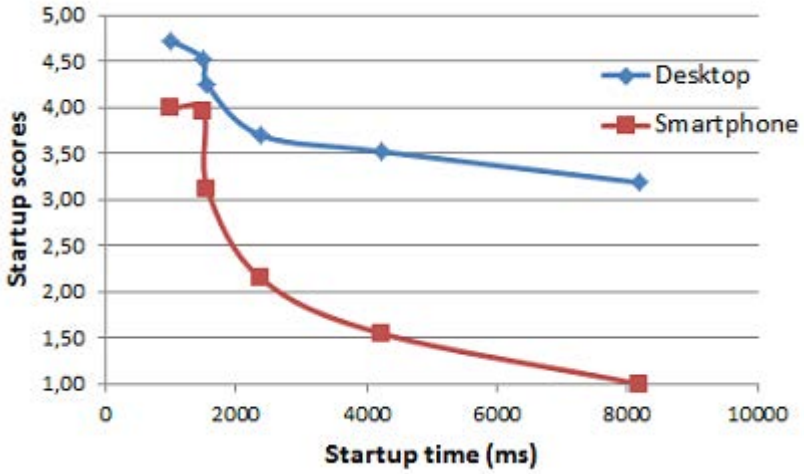

Figure 4.13: General video startup time rating

We discuss the relationship between these metrics and end-users expectations in terms of video startup time and buffering ratio and this for different devices

(smartphone and desktop). For each test we combined all evaluations without considering the type of content.

Figure 4.12 shows in the case of desktop and smartphone that for all content type, the buffering ratio is a critical parameter that impact end-users rating. With a buffering ratio equal to 12% (for desktop and smartphone), we remark that, the end-users rated tends to 2, that corresponds to a poor score according to Table 2.4.

In Figure 4.13, in spite of low throughput, in general, the video startup time scores metric is less critical for users compared to buffering scores. However in the smartphone, we notice that users were more sensitive to the video startup time metric in the case of smartphone than in the desktop.

In the smartphone, for a video startup time approximately equal to 2 seconds, the ends-users scores tend to 2, while in the desktop for the same video startup time, the score in average is superior to 3 (that is an acceptable score) according to Table 2.4.

In general, for both desktop and smartphone whatever the content type, buffering ratio is a critical metric. The video startup time is more negatively correlated to QoE in the case of smartphone than the desktop.

4.4 Evaluation of video bitrate on QoE

In addition to the network imperfections study, the video quality impacts also the QoE. The video bitrate is the considered video quality parameter in our thesis. Therefore, in this section, in order to understand the effect of video bitrate on QoE, we considered various video bitrates (Table 4.6 and Table 4.7) and users are asked to rate each video bitrate for considered device and content type.

Tests	Video bitrate (kbps)/ Smartphone
Test 1	226
Test 2	320
Test 3	680
Test 4	1100
Test 5	3500

Table 4.6: Considered video bitrates for the Smartphone

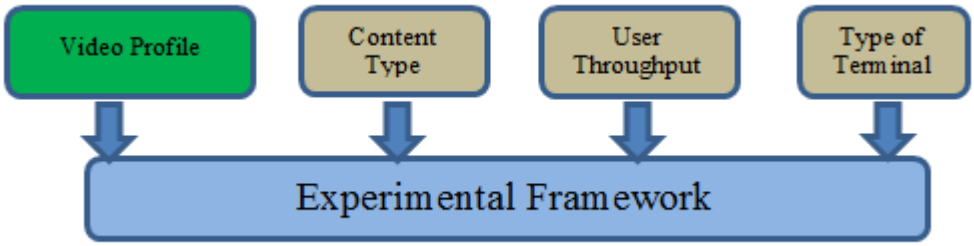

Figure 4.14: Framework to analyze the impact of video bitrate on QoE

Tests	Video bitrate (kbps)/Desktop PC
Test 1	320
Test 2	680
Test 3	1100
Test 4	1600
Test 5	2100
Test 6	3500
Test 7	6000

Table 4.7: Considered video bitrates for Desktop PC

4.4.1 Experimentation

In order to answer such questions: different video sequences (Film, TV news, Animation and Sport) are considered, theses sequences have different characteristics in terms of movement, textures, details, transitions.

In this point, we keep the user throughput constant, but enough in order to prevent the video buffering, which correspond to instants when the player's buffer lacked of data forcing the player to freeze. The effect of video bitrates on QoE is analyzed by considering context information :

- Type of terminal : smartphone and desktop
- Content Type : High motion, low motion ...
- Videos bitrates : Various video bitrates are considered for each user
- Constant User throughput : The considered throughput considered as constant

4.4.2 Results and analysis

As in the study of network degradation effect on QoE, in this section we provide the results and analyze assessment of user study and discuss our findings.

The results obtained for each experiment condition are then consolidated across testers through statistical indicators: an average gives us a Mean Opinion Score (MOS) for each content type, device and network conditions, while the standard deviation captures the spread around the Mean Opinion Score.

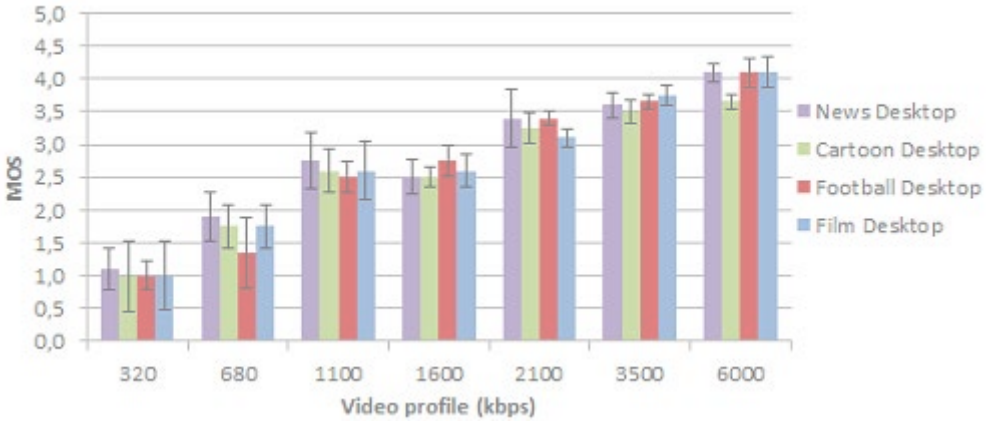

Figure 4.15: Impact of video bitrate on MOS in the case of desktop

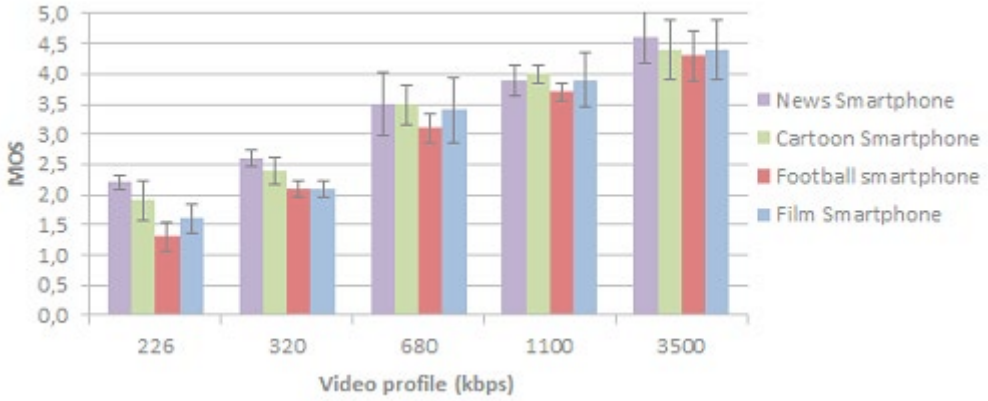

Figure 4.16: Impact of video bitrate on MOS in the case of smartphone

We can observe from Figure 4.15 and Figure 4.16, that users were more tolerant to video bitrate degradation for contents like news & cartoon, than for football or music. We explain that by the fact that the level of temporal motion is higher for football and music clips than for news and cartoons as in the precedent section. In the case of desktop, we remark that there is a small difference in terms of scores between 3500 kbps and 6000 kbps. The average of MOS marks is very close between 1100 kbps and 1600 kbps, and this for all content types.

In the case of smartphone, as in the case of desktop PC, more tolerant con-

tents (news and cartoon) have higher scores compared to less demanding contents (football and music) and for higher bitrates than 680 kbps, for different contents types, the MOS is acceptable.

4.4.3 Video Contextual User Perception Model

As in the study of network impact, we propose also to model the QoE function of video bitrate with variations of context information (device type, content type).

Different relationships (linear, exponential and logarithmic) will be analyzed in this section in terms of regression. The chosen model will be evaluated as in the case of study through the coefficient of correlation expressed in equation 4.1.

In Table 4.8 and 4.9, we find regressions between MOS and the video bitrate (D_{vi}), by taking into account the type of device (desktop or smartphone), content type (news, cartoon, music and football). In order to let end-users to focus on the degradation caused by the video bitrate, in our tests, we considered a high value of D_{ri}, to avoid buffering during the video visualization.

Content Type	Parametric models	Pearson Coefficient of Correlation (PCC)
Football	$1.89 * e^{0.0001*Dv_i}$	**0.65**
	$1.08*\ln(Dv_i) - 5.23$	**0.93**
	$1.57+0.0005*Dv_i$	**0.74**
Film	$1.92 * e^{0.0001*Dv_i}$	**0.70**
	$1.08*\ln(Dv_i) - 5.23$	**0.98**
	$1.63+0.0004*Dv_i$	**0.74**
Animation	$1.95 * e^{0.0001*Dv_i}$	**0.59**
	$1.14*\ln(Dv_i)-5.72$	**0.95**
	$1.57+0.0005*Dv_i$	**0.74**
News	$2.2 * e^{0.0001*Dv_i}$	**0.67**
	$0.95*\ln(Dv_i)-5.72$	**0.95**
	$1.72+0.0004*Dv_i$	**0.67**

Table 4.8: Video parametric models case of desktop PC

In the case of desktop and smartphone, the logarithmic model has higher Pearson Coefficient of Correlation, PCC ≈ 0.9. We propose to model the degradation caused by video bitrate by the parametric model below, the Video Contextual User Perception Model (VCUPM) :

Content Type	Parametric models	Pearson Coefficient of Correlation (PCC)
Football	$2.23 * e^{0.0003*Dv_i}$	**0.60**
	$1.12*\ln(Dv_i) - 4.54$	**0.93**
	$2.05+0.0007*Dv_i$	**0.66**
Film	$2.4 * e^{0.0001*Dv_i}$	**0.57**
	$1.07*\ln(Dv_i) - 4.18$	**0.91**
	$2.26+0.0007*Dv_i$	**0.62**
Animation	$2.64 * e^{0.0001*Dv_i}$	**0.57**
	$0.93*\ln(Dv_i)-2.9$	**0.91**
	$2.52+0.0006*Dv_i$	**0.61**
News	$2.75 * e^{0.0001*Dv_i}$	**0.70**
	$0.88*\ln(Dv_i)-2.45$	**0.97**
	$2.60+0.0006*Dv_i$	**0.74**

Table 4.9: Video parametric models case of smartphone

$$\varphi_V(D_{vi}) = \alpha_1 \log(D_{vi}) + \beta_1 \tag{4.10}$$

Where :

- $\varphi_{V_i}(D_{vi})$ is the score caused by video V characterized D_{vi}
- α_1 and β_1 are the model parameters calculated by using subjective test data from different experiments and are presented in Table 4.8 and Table 4.9.

We represent in Figure 4.17 and Figure 4.18 test data as an average for each test and different regressions model (exponential, logarithmic and linear) in the case of football , in order to show that the logarithmic model has higher correlation with the experimental data, where the experimental data corresponds to the average of MOS marks for each test.

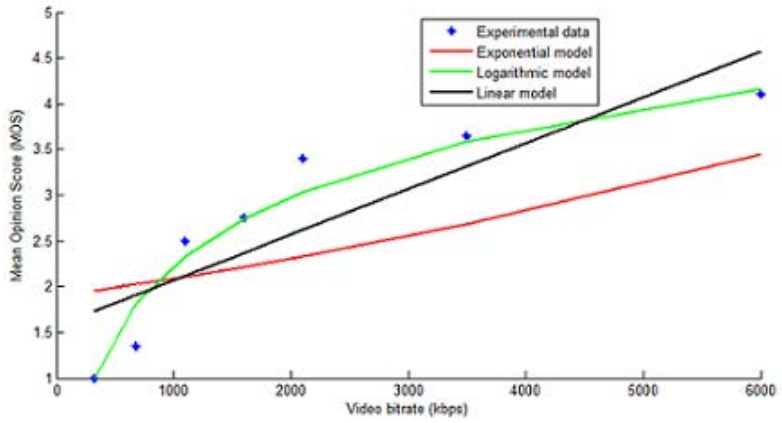

Figure 4.17: Different regressions for video in the desktop for football

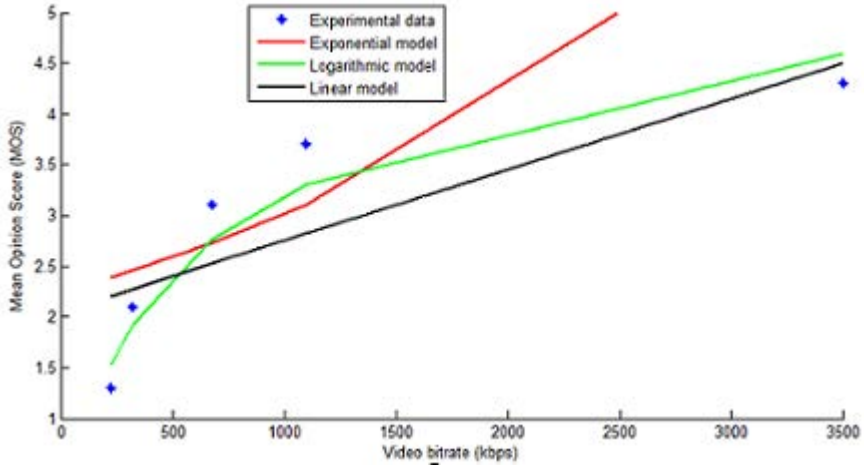

Figure 4.18: Different regressions for video in the smartphone for football

4.5 General QoE modeling

After having worked on the impact of network degradation and video bitrate separately, in this paragraph we present the general QoE modeling.

The idea is when there are enough network resources to absorb the video ($D_{ri} > D_{vi}$), the parameters which influence the perceived quality are the characteristics of video bitrate. In another side, when there is lack of resources, the Quality of Experience is function of available resources. The general QoE model is represented as:

$$\varphi_i(D_{ri}, D_{vi}) = \begin{cases} \varphi_{V_i}(D_{vi}) & if \quad D_{ri} > D_{vi} \\ \varphi_{N_i}(D_{ri}) & otherwise \end{cases} \tag{4.11}$$

Where : $\varphi_i(D_{ri}, D_{vi})$ corresponds to the general QoE model.

4.6 Comparisons

In this section we compare different models. As can be seen from the previous section, we presented many parametric models. Each of the models was designed at different conditions, taking into account specific parameters (content type, type of terminal, video bitrate, frame rate...). A comparison of models in Table 4.10.

Authors	Video bitrate	Network status	Content type	Type of terminal	Buffering
ITU-T G 1070	Yes	No	No	No	No
M. Ries et al	Yes	No	Yes	No	No
A. Khan et al	Yes	Yes	Yes	Yes	No
Yen-Fu et al	No	No	Yes	Yes	No
PSQA	No	Yes	Yes	No	No
CUPM	Yes	Yes	Yes	Yes	Yes

Table 4.10: QoE measurement comparisons

In addition to the qualitative comparison represented in Table 4.10, we compare the performance of NCUPM to aquatool [78], a professional Orange Labs R&D tool. It is based on digital filters, each filter being capable to monitor video quality metrics such as the detection of blocking effect frequently due to lack of throughput, and many others.

In our comparison the used filter triggers a 0/1 signal on a video problem event (such as video buffering), since human eyes are very sensitive to video freeze. This signal is evaluated every 50ms, and equals to :

- "Bad" if and only if a buffering of a duration greater than 500 ms is detected
- "Excellent" otherwise

After the video capture and treatment with aquatool, the generated file gives use the obtained results. With theses data, we define a score :

$$aqua = \frac{5.E + B}{E + B} \tag{4.12}$$

where E is the number of occurrence of the "Excellent" signal, and B is the number of occurrence of "Bad" signal. In other words, aqua maps Aquatool-based signals to the 1-5 MOS scale, but is only based on buffering events.

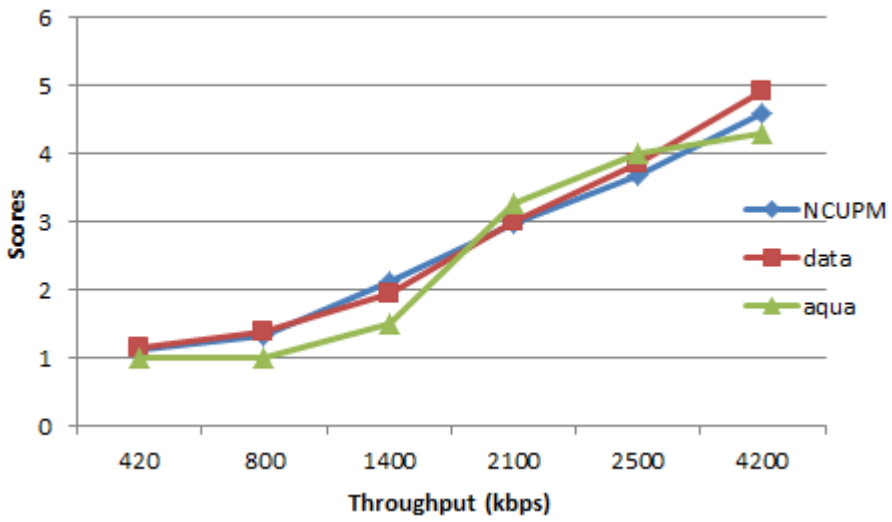

Figure 4.19: Performance comparison

We plotted in 4.19 the obtained scores for the desktop environment independently of the content type. The graph shows the score obtained in terms of detection of buffering, subjective scores and NCUPM. It allows comparing performance of : NCUPM, data (corresponds to the average of MOS marks for each test as described in Table 4.3) and aqua the obtained score by using Aquatool.

As it can be seen, the three approaches are comparable and fits well one to each other. The consistency of the comparison is ensured by the fact that buffering are due to lack of video data in the player buffer, due to a lack of bandwidth on the network to feed this buffer. However, the reader shall keep in mind the differences:

- data score obtained by human-based assessments.
- aqua is obtained by a production tool used on observation of streamed video.
- NCUPM, our model is a proactive approach taking overall and general parameters.

Moreover, it provides an easy parametric model to characterize end-user satisfaction, and can take into account context information for an accurate prediction, as long.

4.7 Conclusion

In conclusion, the overall behavior of end-users' perception while watching video is insufficient when considering only network throughput: MOS marks increase with network bandwidth. However, we bring new phenomena to the light in this analysis. For instance, we show that content type and device type are impacting

factors for MOS marks at equivalent network conditions. As a consequence, content adaptation & recommendations systems can strongly benefit from knowledge about these two parameters.

Another result that we can mention from this study is, in general when we increase the video bitrate, we can increase MOS. But in particular, end-users were more tolerant to video bitrate degradation for contents like news & cartoon, than for football or music.

We also provided mathematical models that accurately captures MOS function of network and video degradation by considering context information in terms of content type and device type. We compared our proposal with an professional tool that provide QoE measurement. The results prove that our model is very well correlated to experimental data.

The impact of quality parameters on end-users satisfaction were also analyzed. In general, for both desktop and smartphone, buffering ratio is a critical metric. The effect of video startup time is more negatively visible in the case of smartphone than the desktop.

In next chapter, in order to manage congestion in network due to increasing demand of ends-users and heterogeneity of used devices, we propose video content and delivery adaptation. Therefore, the proposed solutions will improve the end-users QoE.

Chapter 5

MOS based approach for video adaptation

Video applications are expected to be ever more popular in near future. There are strong needs to develop approaches and mechanisms to manage congestion in network, due to increasing demand in wireless access.

Most of the QoE optimization are based on an individual measure reflecting the satisfaction for a particular user. There are needs to develop approach for assessing the QoE metric over the whole network, taking into account the fairness between users.

In order to achieve efficient resource usage in a network and to provide high QoE to users, it is necessary to optimize the overall network using multi-user QoE metrics, which characterize the overall network performance.

In this chapter, we study some multi-user QoE metrics and provide general QoE multi-users QoE optimization in the case of DASH (Dynamic Adaptive Streaming Over HTTP), that optimizes the network resources by considering the end-user satisfaction in terms of MOS, called the MDASH (MOS Dynamic Adaptive Streaming Over HTTP).

Typically, in principle the DASH allows a multi-rate delivery, where different video quality are available for the same content and the adaptation is based on available bandwidth. This means that DASH adapt the video quality individually for each user without considering the end-user satisfaction. Also, in DASH, clients are not aware of each other and resource sharing may be unfair when different implementations of DASH are competing in the same network.

In principle, our approach maximizes the satisfaction of users, where they are streaming different videos content characterized by their context information (type of device, content type, and video bitrate), in order to make efficient use of available network, and therefore improves the perceived QoE for video sessions sharing the same local network, while taking QoE fairness among users. As an application, we consider the case of domestic network, where there are limited network resources due to increasing demand for high video quality.

In Figure 5.1, we represent the architecture of home network, where multiples users are sharing the same network. In Figure 5.12, there is a video content deliv-

ery chain, composed by the core network (a central part of delivery, where various services to customers can be found), access network (that connects subscribers to their core network) and the CPE (Customer Promises Equipment), a telecommunication device located at the home network, managing different running services. In this work as we focus on the video delivery adaptation in the home network, then all our optimization techniques will be integrated in the CPE.

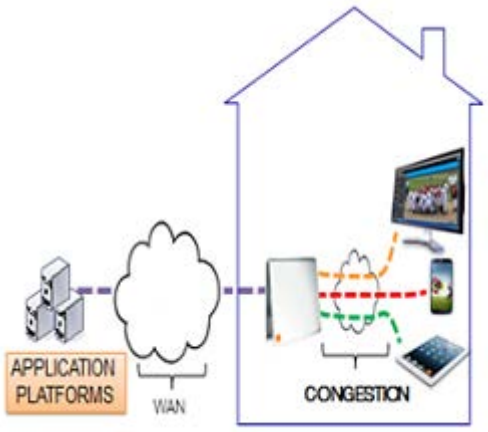

Figure 5.1: Home network architecture

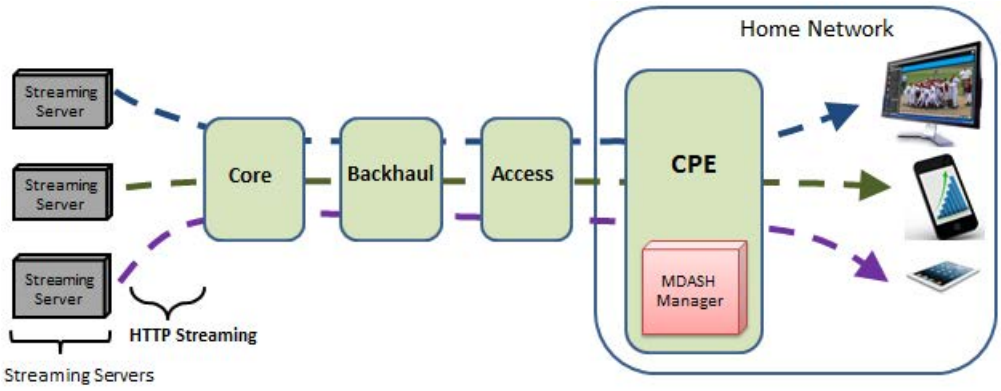

Figure 5.2: Home network delivery chain

5.1 Multi-user QoE optimization

The developed QoE metric expressed in equation 4.8 or 4.10, are individual measures reflecting the satisfaction of a particular user. There are needs to develop approaches for assessing the QoE metric over the entire network (multi-user QoE metrics), taking into account the fairness in the QoE perceived by different users.

The multi-user QoE corresponds to the overall performance for different users sharing the same network. In order to achieve efficient resource usage in a network and to provide high QoE to the users, it is necessary to optimize the overall network by using multi-user QoE metrics, which characterize the overall network performance.

We notice that, our proposal falls in the case of DASH. Following the principle of DASH, in order to limit buffering, the media player selects the most adequate video fragment based on the available bandwidth. We suppose that, the parameter differentiating the end-users perception is the streamed video bitrate. Therefore, the most appropriate single user QoE model is the VCUPM (Video Contextual User Perception Model), expressed as φ_{V_k}, in section 4.4.3 that estimates the QoE function of video bitrate with respect to contextual information (device and content type).

5.1.1 Multi-user QoE metrics

Most of the QoE investigations, consider the QoE perceived by a single user. In this section, we introduce some QoE metrics reflecting the general performance for all connected users in the same network.

5.1.1.1 The average QoE metric

This metric is the most natural that comes to mind and represents the average QoE for$K$ connected users in the network. It is given by the formula 5.1:

$$f_{avg}(D_{vk}) = \frac{1}{K}\sum_{k=1}^{K}(\varphi_{V_k}(D_{vk})) \tag{5.1}$$

where:

- φ_{V_k} is the function for assessing QoE for single user$k$ function of video bitrate and with different context information (device and content type), expressed in equation 4.10.
- D_{vk}, is the selected video bitrate for user k at time t.

This metric is extremely dependent on the characteristic of the single QoE model. In addition, in some case, f_{avg} could be high, when some users have very high satisfaction, while others have a low satisfaction, this could mask unfairness toward users with low satisfaction.

5.1.1.2 The min QoE metric

This metric represents the lower Quality of Experience (QoE) among K users. It is given by the formula 5.2:

$$f_{min}(D_{vk}) = min_{1 \leq k \leq K}(\varphi_{V_k}(D_{vk})) \tag{5.2}$$

Certainly, this metric depends on the characteristic of the single user QoE model. But whatever the type of that single user model, the metric f_{min} is fair since, it takes resources from users that have high QoE and gives it to users that are in the worst case. The fairness between users is guaranteed and finally they have the same QoE.

5.1.2 Multi-user Optimization approach

In the multi-user case, we have different clients sharing the same network, in this situation, searching optimization techniques for managing concurrent video networks is required.

MDASH is an approach that optimizes the network resources by considering the end-user satisfaction in terms of MOS in the video quality adaptation, by resolving the max-min approach that ensures fairness between users. In principle the considered method:

- Finds optimal resource allocation, in order to maximize the satisfaction of user which has the lowest MOS.
- Clients react to the throughput changes and select a video bitrate which matches the available client resource.

In this section, we will present a theoretical and practical approach of the MDASH.

5.1.2.1 MDASH theoretical approach

This principle corresponds to the situation where for each allocated network resource, we have the corresponding video bitrate. We assume that, it is a theoretical approach, because, in reality video bitrates are in a finite interval (finite values).

In DASH principle, in order to limit buffering, the media player selects the most adequate video bitrate (D_v) based on the available network resource (D_r). The relation between network resource and video bitrate is given by the formula 5.3.

$$D_r \geq D_v.\delta_{HAS} \tag{5.3}$$

The value of δ_{HAS} defines the moment where clients switch among different video bitrates according to their available bandwidths and depends on the implementation of algorithm.

In this approach, the value of δ_{HAS} is 1 and for each allocated network resource (D_r), we have the associated video bitrate (D_v). The relation 5.3 between network resource and video bitrate becomes :

$$D_r = D_v \tag{5.4}$$

The objective of the optimization problem is given by the equation 5.5:

$$S^*_{opt} = argmax_{(D_{r1}.....D_{rk})} \{min_{1 \leq k \leq K}(\varphi_{V_k}(D_{rk}))\} \tag{5.5}$$

Subject to:

$$\sum_{k=1}^{K} D_{rk} \leq C \tag{5.6}$$

where:

- S^*_{opt}, are optimal values for network resources allocation.
- φ_{V_k}, is the function for assessing QoE for single client k, function of video bitrate and with different context information (device and content type), expressed in equation 4.10.
- D_{rk}, is the available throughput for user k.
- C is the the available resource in the network.

This optimization problem is fair since, it takes resources from users that have high QoE and gives it to users that are in the worst case.

Resolution of the optimization problem: In this part, we are going to solve the optimization problem proposed in equation 5.5. We will show that, this problem is equivalent to proving intuitively that:

$$\forall i, j \Rightarrow \varphi_{V_i}(D_{ri}) = \varphi_{V_j}(D_{rj}) = Z \tag{5.7}$$

In Figure 5.3, the principle of the optimization problem is represented, it graphically shows the equation 5.7. The following are the steps in the optimization process:

1. Users that have high QoE (satisfied users) will give a little network resource to unsatisfied ones.
2. The given network resources by satisfied users (in the first step) are used by unsatisfied ones in order to increase their MOS.
3. The network resources are allocated for all users in order to have the same score.

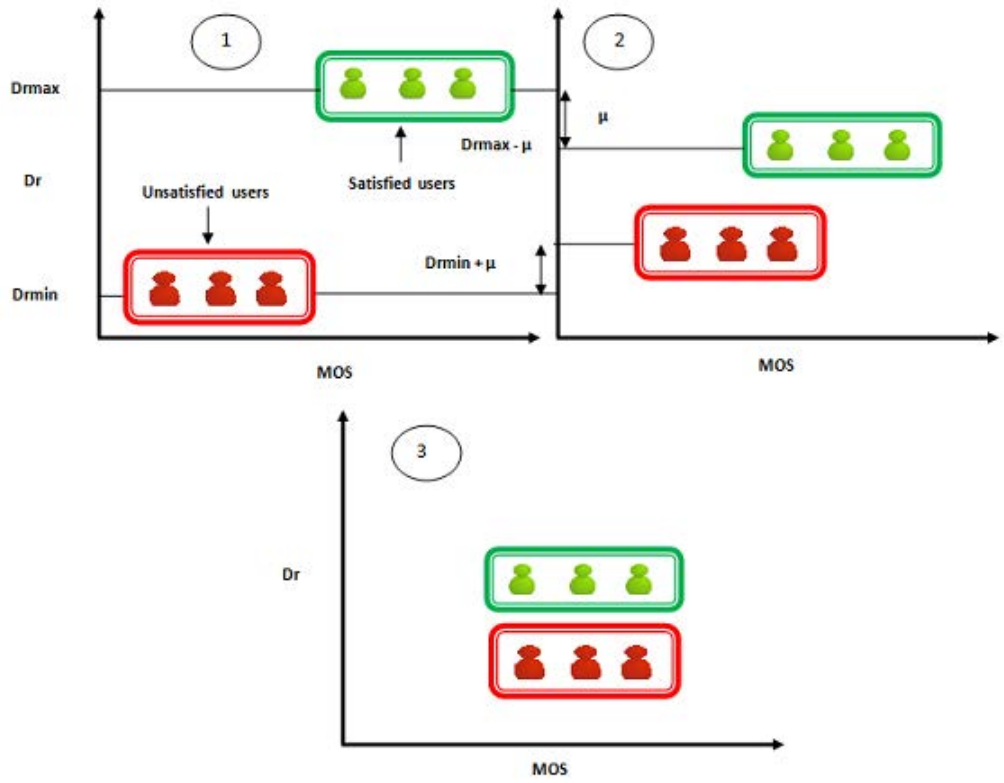

Figure 5.3: Max_min optimization principle

We suppose that, the common score is called Z and is given in equation below:

$$Z = \varphi_{V_k}(D_{rk}) \Leftrightarrow D_{rk} = \varphi_{V_k}^{-1}(Z) \tag{5.8}$$

The allocated network resource C, in order to have the same score equal to Z for all users is given by:

$$C = \sum_{k=1}^{K} D_{rk} = \sum_{k=1}^{K} \varphi_{V_k}^{-1}(Z) \tag{5.9}$$

We suppose that $\omega(Z)$ is to total capacity and is given in equation below:

$$\omega(Z) = \sum_{k=1}^{K} \varphi_{V_k}^{-1}(Z) \tag{5.10}$$

We have from equation 4.10 that:

$$\varphi_{V_k}(D_{rk}) = \alpha_1 \log(D_{rk}) + \beta_1 \Rightarrow D_{rk} = e^{\frac{\varphi_{V_k}(D_{rk}) - \beta_1}{\alpha_1}}$$

$$\varphi_{V_k}^{-1}(Z) = e^{\frac{\varphi_{V_k}(Z) - \beta_1}{\alpha_1}} \tag{5.11}$$

Then equation 5.10 becomes :

$$\omega(Z) = \sum_{k=1}^{K} e^{\frac{\varphi_{V_k}(Z)-\beta_1}{\alpha_1}} \tag{5.12}$$

The function $\varphi_{V_k}^{-1}(Z)$ is an increasing function, then $\omega(Z)$, that is the sum of increasing functions is an increasing one. The question is to find Z_{opt}, where $\omega(Z_{opt}) = C$.

As the $\omega(Z)$ is an increasing and continuous function in R^{*+}, then in order to solve the optimization problem, that will allocate network resources, in the aim to have the same score Z_{opt} for all users, we can use the dichotomy method [79].

5.1.2.2 MDASH practical approach

The MDASH approach corresponds to the case, where the video bitrate takes values from continuous interval $\{D_{v1}, ...D_{vk}\}$. The relation between network resource and video bitrate is equivalent to the given equation in 5.3. In this case, the objective of the optimization problem is:

$$S^*_{opt} = argmax_{(D_{v1}.....D_{vk})} \{min_{1\leq k\leq K}(\varphi_{V_k}(D_{vk}))\} \tag{5.13}$$

$$S^*_{opt} = argmax_{(D_{v1}.....D_{vk})} \left\{min_{1\leq k\leq K}(\varphi_{V_k}(\frac{D_{rk}}{\delta_{HAS}}))\right\} \tag{5.14}$$

Subject to:

$$\sum_{k=1}^{K} D_{rk} \leq C \tag{5.15}$$

$$D_{vk} \in [D_{v1}, ...D_{vk}] \tag{5.16}$$

This approach finds optimal resources allocation in the first step, in order to improve the QoE of user that has the lowest satisfaction, and after that, clients react to the new resources allocation and choose a video bitrate which matches their network resources.

As D_{vk} takes values from a continuous interval, it means that, there are a finite values of D_{vk}. When the available throughput is greater or equal to any one of the many available video bitrates, the client can stream that video. In this case, we consider also that $\delta_{HAS} = 1$, then the equation 5.3 becomes:

$$D_r \geq D_v \tag{5.17}$$

We consider (Res_k) as the difference between allocated network resource D_{rk} for user k and D_{vk} , the associated video bitrate.

We suppose that *sumRes* is the sum of residues. Relations are presented in equations below:

$$Res_k = D_{rk} - D_{vk} \tag{5.18}$$

$$sumRes = \sum_{k=1}^{K} Res_k \tag{5.19}$$

The question is, how we can use the sum of residues, in order to permit users to jump, in terms of network resources and therefore select a higher video bitrate. Initially, we have to check, if the sum of residues allows to select a higher quality by checking the relationship:

$$sumRes > min_k |D_{vk+1} - D_{vk}| \tag{5.20}$$

where, D_{vk}, is the selected video bitrate by user k and D_{vk+1}, the video bitrate that directly follows the selected one.

This condition means, if the relationship 5.20 holds true, the residue could be used in order to allow a customer to select a higher video quality than chosen and thus enables it to improve its QoE.

5.2 Experimental results

In this section, we propose to validate the MDASH algorithm by simulations and compare performances with standard DASH.

5.2.1 Reference algorithm

Almost all the commercial DASH products use their proprietary rate adaptation algorithms. The resource allocation for the considered reference algorithm relies on standard TCP principle, where the available resource is divided by the number of streams without considering the characteristics of streams and used devices.

5.2.2 Experimentation simulations

In home network, we consider clients are requesting different DASH video with different devices. We propose to validate our proposal and to compare performances with the reference algorithm. The reference algorithm is called the Non-Optimization scheme (Non-Opt).

In Table 5.1, we present, the simulation parameters and in Figure 5.2, the considered streaming environment in terms of content and device type. We compute the MOS associated to each context (video bitrate, device and content type), by using the developed model the VCUPM.

Simulations parameters	Values
Total resource (MHz)	8
Number of users	4
Considered video bitrate (Kbps)	226 - 320 - 680 - 1100 1600 - 2100 - 2400 - 2800 3200 - 6000
Application Type	HTTP Adaptive Streaming

Table 5.1: Simulations parameters

Users	Type of content	Type of device
User 1	Sport	Desktop PC
User 2	Movie	Desktop PC
User 3	News	Smartphone
User 4	Animation	Smartphone

Table 5.2: Considered users for simulations

In Figure 5.4, we represent the resource allocation and the associated video bitrate for the DASH that corresponds to the non-optimized scheme (Non-opt). Users have the same resource equal to 2000 KHz (there are four users and the available resource is 8000 KHz), then media player's chooses the video bitrate less or equal to their resource, in our case, the corresponding video bitrate is 1600 Kbps according to the available video bitrate (see Table 5.2).

In Figure 5.5, we represent the resource allocation and the associated video bitrate in the case of MDASH (optimized scheme). As user 1 and user 2 consume (sport and movies) in desktop PC, that have high expectations in terms of resources, the MDASH algorithm allocates more resources to these users than user 3 and user 4, viewing news and animation respectively.

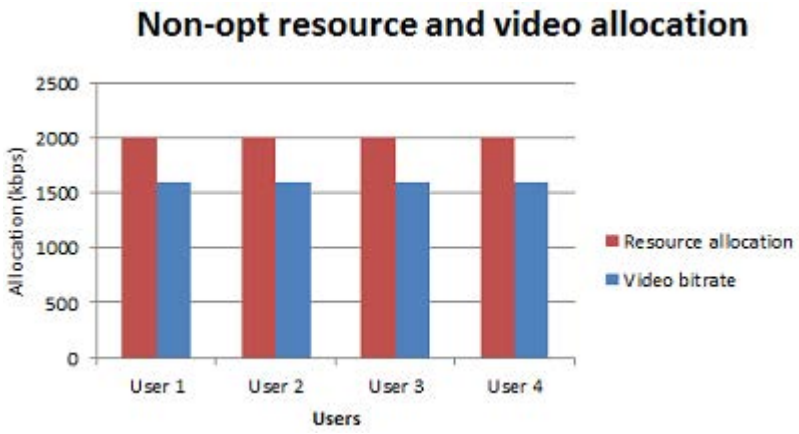

Figure 5.4: Non-opt resource and video allocation

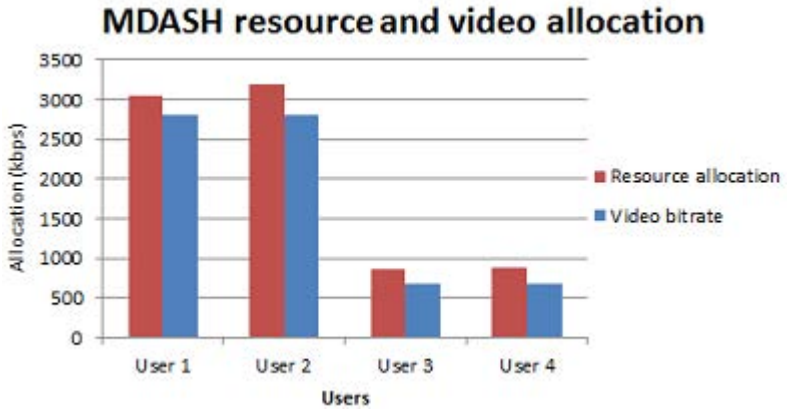

Figure 5.5: MDASH resource and video allocation

Furthermore, in order to explain and highlight the benefits of MDASH based optimization that considers the device type and content characteristics, we evaluate the Mean Opinion Score (MOS), for each users in Figure 5.6, DASH provides good results for less demanding contents and devices (user 3 and user 4), but fails for more demanding ones (user 1 and user 2).

The MDASH allocates resource among the users such that the minimum of users is maximized. Results show the importance of MOS based video adaptation that maintains quality for less demanding users and improve the perceived quality for more demanding ones.

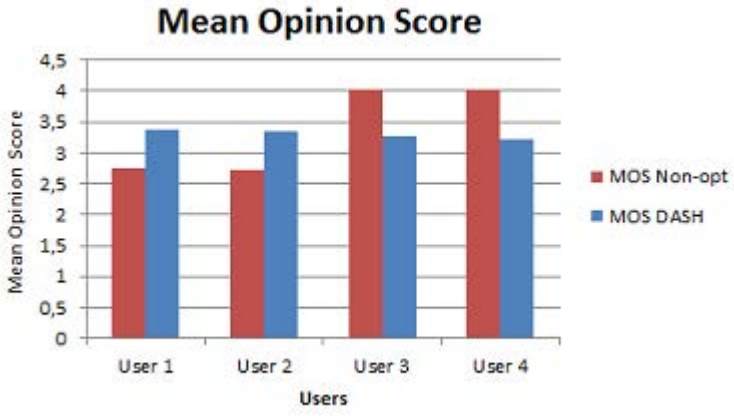

Figure 5.6: Mean Opinion Score before and after optimization

5.2.2.1 Needed capacity for fixed QoE

In order to show the importance of context information in video optimization. In Table 5.3, we consider different configurations. For each configuration, we consider four use cases, in this paragraph we define the needed network capacity for fixed MOS. As said in [80], a MOS > 3 guarantee an acceptable quality for all users. In our work, we decided to choose the MOS threshold equal to 3.2, in order to be a little higher than the minimum fixed in [80].

Use cases	Configuration 1	Configuration 2	Configuration 3	Fixed MOS
1	1 user viewing Football in PC	1 user viewing Football in PC	1 user viewing Football in PC	3.2
2	2 users viewing Football in PC	2 users viewing Football in PC	2 users viewing Football in PC	3.2
3	3 users viewing Football in PC	2 users viewing Football in PC and 1 viewing Football in Smartphone	3 users viewing Football in PC and 1 viewing News in Smartphone	3.2
4	4 users viewing Football in PC	2 users viewing Football in PC and 2 viewing Football in Smartphone	2 users viewing Football in PC and 2 viewing News in Smartphone	3.2

Table 5.3: Considered use cases and configurations

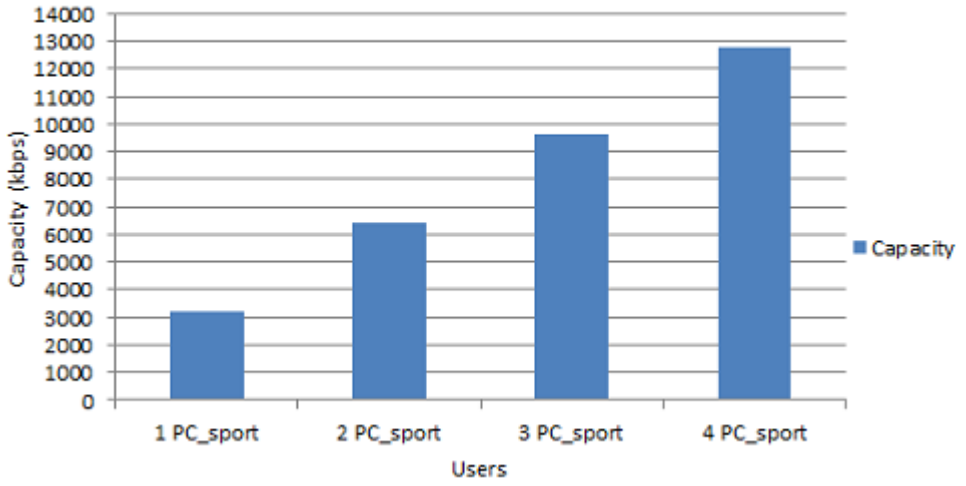

Figure 5.7: Capacity for fixed MOS in same devices

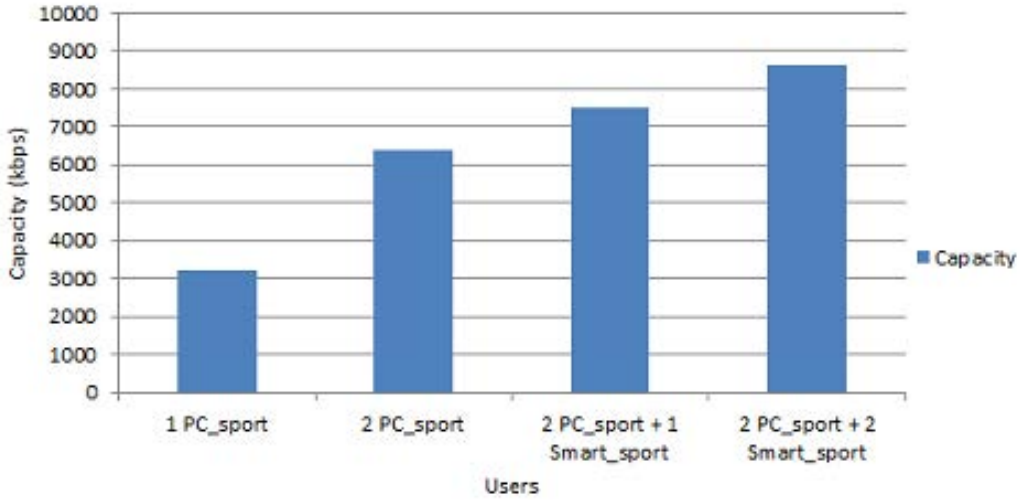

Figure 5.8: Capacity for fixed MOS in different devices

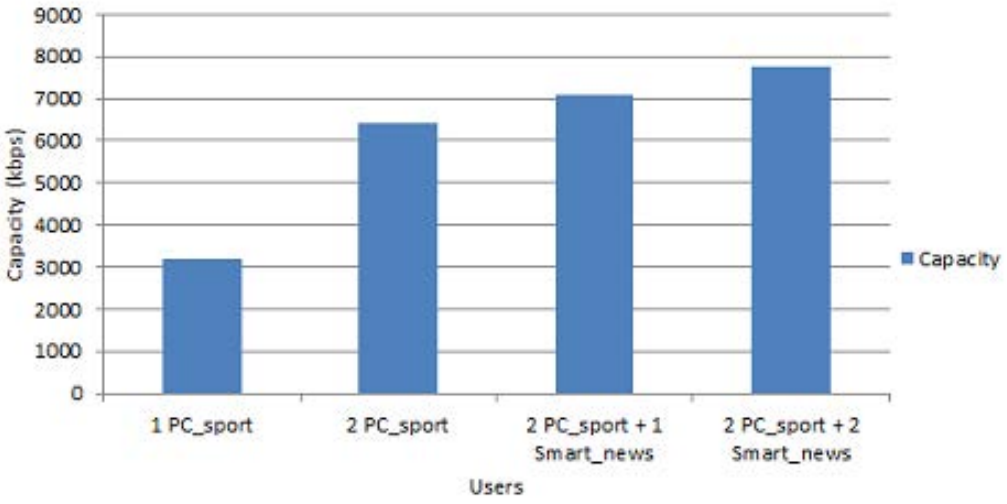

Figure 5.9: Capacity for fixed MOS in different devices and different contents

We can observe from Figure 5.7, that needed network capacity increases with the numbers of users, and in order to ensure a minimum MOS equal to 3.2, for three users, the needed network capacity is 9600 Kbps.

While, in Figure 5.8, the needed capacity for three users (where users consume the same content but in different devices) is 7600 Kbps for the same minimum guarantee MOS. The needed resource in this case, is less than the first one, this can be explained by the fact that, one of users uses smart phone, therefore less demanding in terms of resource to achieve the fixed MOS.

For the same use case (where one user streams news in the smart phone), we remark from Figure 5.9 that, the needed resource is equal to 7080 Kbps, this is less than the precedent case, due to the fact that the level of temporal motion that is higher for football than for news.

In this paragraph, we show that the content type and the device type can influence the needed network resource. As a consequence, resource allocation systems can strongly benefit from knowledge about these two parameters.

5.3 Proposed architecture and approaches for implementation

In this section, we first present the existing resource and admission control architecture for resources reservation in NGN (Next Generation Network) and learn from that architecture, in order to build our proposal. In addition, we propose approaches for implementation of our proposal.

5.3.1 Existing Resource and Admission Control Architecture

The ITU-T proposed a specified architecture for Resource and Admission Control Functions (RACF) supporting end-to-end QoS in NGN [81][82].

The RACF provides policy-based transport resource management for services. The RACF executes policy-based transport resource control upon user request through the Service Control Functions (SCF), that is in charge of requesting and executing services. The RACF entity determines transport resource availability and makes admission decisions. Then the RACF enforces policy decision to Transport Functions (TF). In the ITU-T NGN, dynamic policy control is a basic requirement. So the home network should be also considered for NGN based end-to-end QoS.[83]. For this, in our work we propose to extend this architecture for resource and control admission in home network.

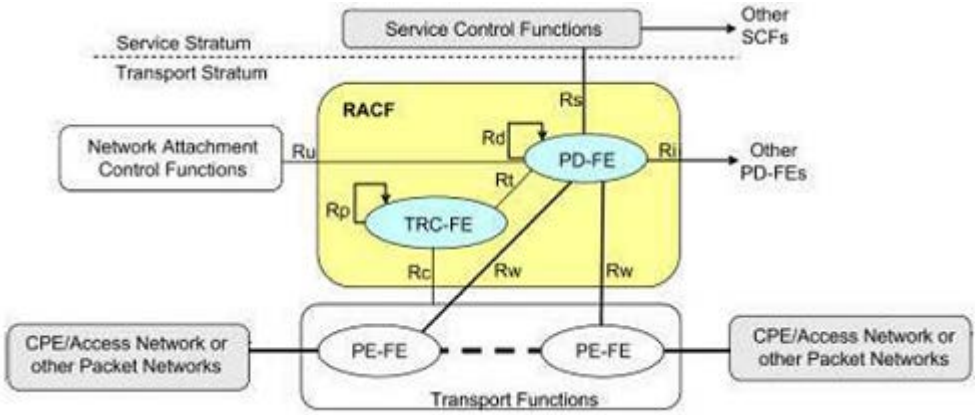

Figure 5.10: Resource and admission control functional architecture in NGN [ITU.T Rec. Y.2111]

As shown in Figure 5.10, RACF is composed of two main entities. The PD-FE (Policy Decision Functional Entity) and TRC-FE (Transport Resource Control Functional Entity). The functionality of PD-FE is to make policy decisions and the TRC-FE is to determine network resources information.

The policy rules can be based on network information, Service Level Agreement (SLAs), Service Information reveived from the SCF (Service Control Function). The PD-FE makes final QoS reservation decision based on information received from TRC-FE. Then the PD-FE sends request to PE-FE (Policy Enforcement Functionnal Entity to execute the final decision. In Figure 5.11, we represent a basic procedure for QoS reservation [84].

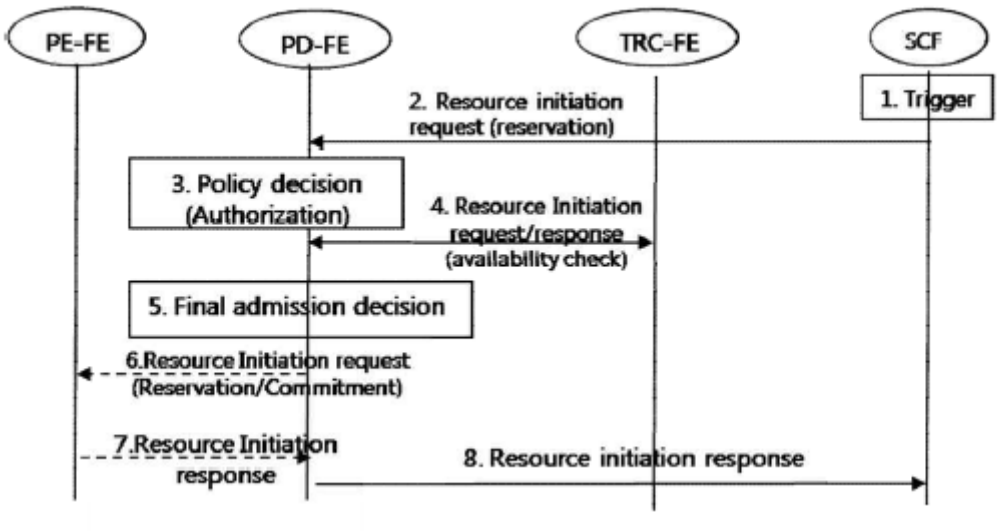

Figure 5.11: Procedure for QoS reservation

5.3.2 Proposed MDASH architecture and approaches for implementation

In order to propose the MDASH architecture for resources reservation and video adaptation in home network, we decided to learn from the ITU-T approach presented in section 5.2, for resource allocation. The general MDASH architecture is composed as follows:

- Context Manager: Gathers the context information from the Terminal Context, and Service Context modules.
- Context Storage: This module stores gathered context information from device, network and service.
- QoE optimizer: This component is in charge of computing the QoE of connected users, according to the developed QoE model.
- Optimization rates: This module finds optimal rates for each user by solving the optimization problem.
- RACF: Is in charge of network resources reservation and execution.

The proposed architecture in Figure 5.12 gathers context information respectively from device, service and network (step 1-3). Gathered contents are stored on the Context Storage module (step 4). The stored context information is sent to the QoE Optimizer module , that is in charge of computing the QoE of connected users according to their context information and resolve the optimization problem (step 5).

Optimal rates are signaled by the QoE optimizer to the Optimal rates (step 6). The Optimal rates are sent to PD-FE (step 7), that is in charge to apply Optimal rates returned by the QoE optimizer module. The PD-FE sends decision to PE-FE (Policy Enforcement Functionnal Entity to execute the final decision.

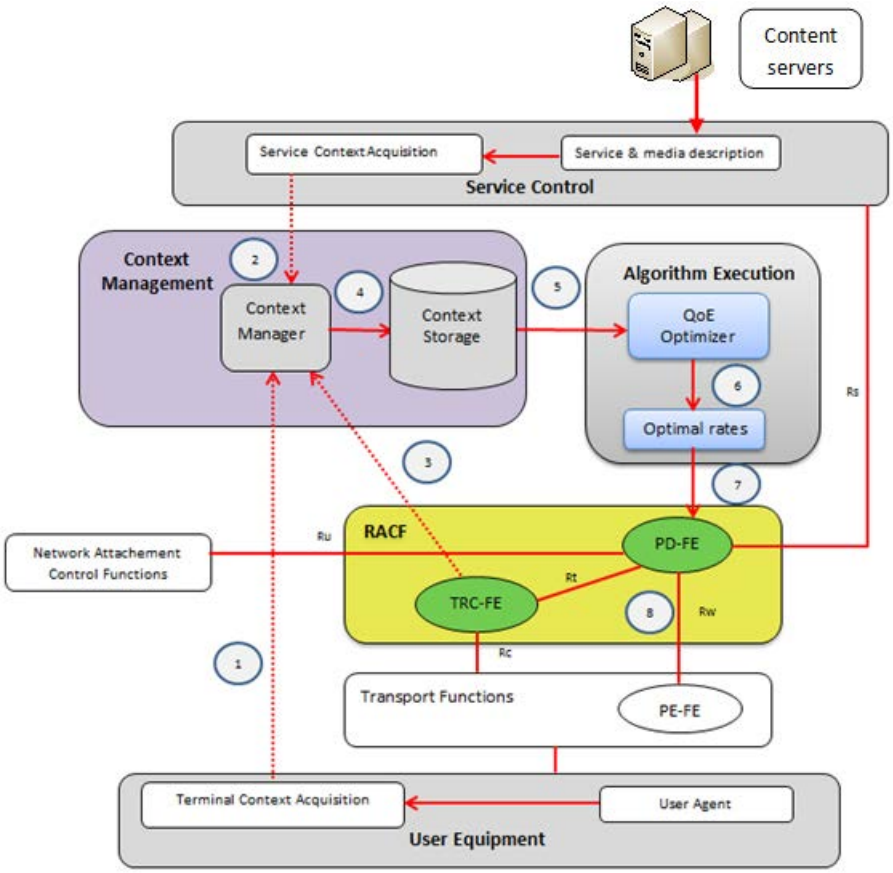

Figure 5.12: Proposed home gateway architecture

5.3.3 Approaches for implementation

The MDASH approach is composed by three parts: Contexual information collection, execution of the optimization algorithm (presented in section 5.2) and apply resource reservation rules from optimization algorithm, the principle is represented in Figure 5.13.

In this section, we describe steps for gathering context information in device, network and service side and propose different approach for execution of resource reservation.

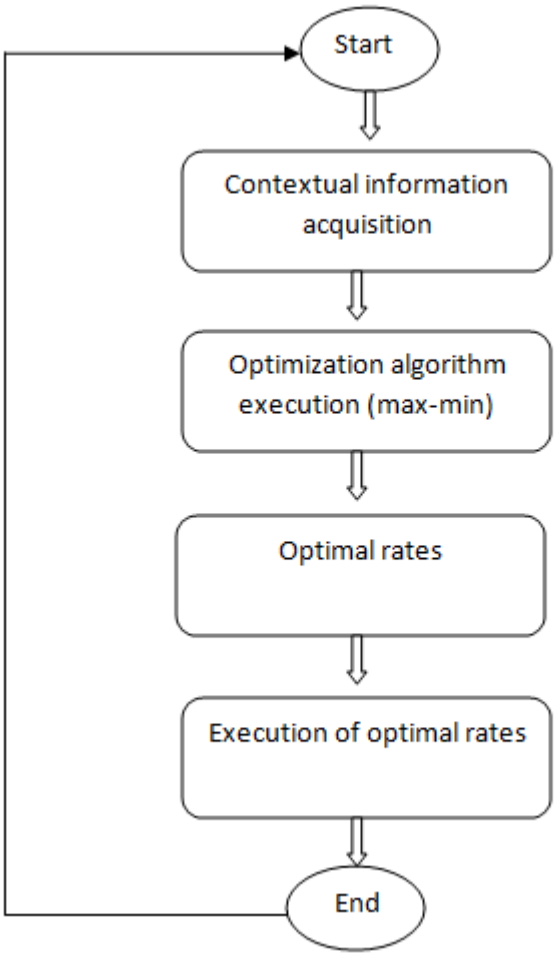

Figure 5.13: MDASH principle processes

5.3.3.1 Context information gathering and transmission

The used protocol for context information gathering and transmission is HTTP. Especially the HTTP POST is used to transmit context information. The HTTP POST message respects the XML format.

- Terminal context gathering and transmission

This procedure is proposed in order to update the context information in terminal side. The User Equipment is composed by two main modules: i) User agent: The most way of finding out the user's terminal type is user-agent [85]. The client inserts as a header field into HTTP request. The user header identifies the client software and version. ii) Terminal Context acquisition: This module is used in order to transmit the terminal context to the context manager. This procedure is repeated when the Terminal Context changes.

Figure 5.14 illustrates the Terminal context transmission message.

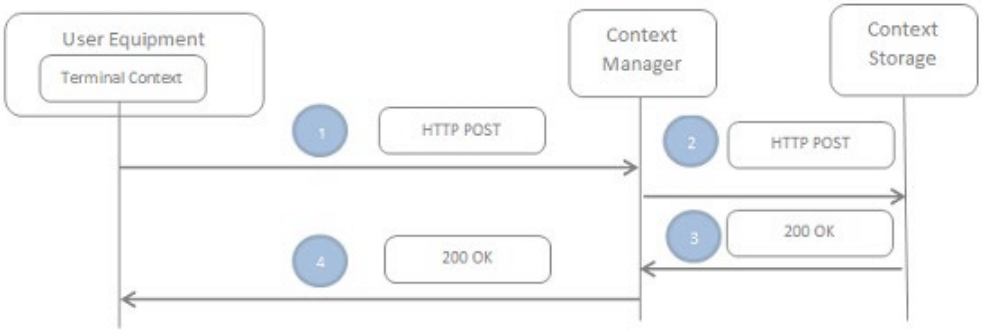

Figure 5.14: Terminal context transmission

- Service context gathering and transmission

The principle is the same than in the terminal context and is repeated when the Service Context changes. The HTTP POST message is also used. The type and characteristics of contents are extracted from service and media description entity received from the content provider.

The Figure 5.15 presents the principle of service context transmission.

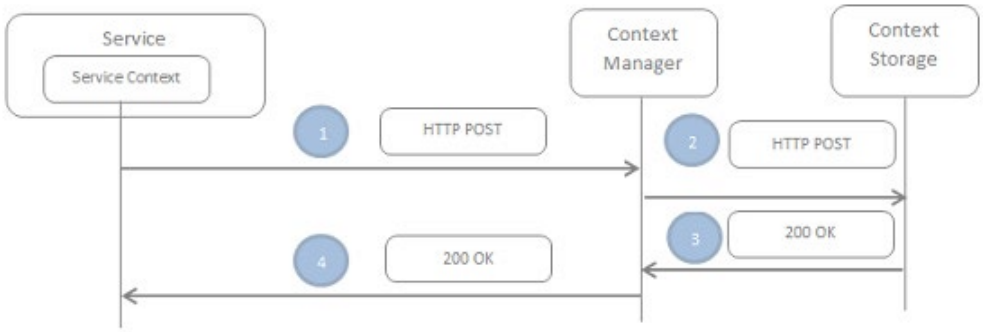

Figure 5.15: Service context transmission

- Network context acquisition

This procedure concerns the network context transmission to the context manager. The gathered parameter in this step corresponds to the available resources in the home network. The TRC-FE collects network information, such as network status such, network topology and network resource information. In [81], the recommended protocol for network information collection is Diameter. In Figure 5.16, we present the principle of network information collection.

The proposed architecture uses the same principle than the existing ITU-T Resource and Admission Control Functional architecture in NGN, in order to consider QoE metric on QoS reservation in the home network. The proposed approach maximizes the QoE of that user which has the lowest QoE among others.

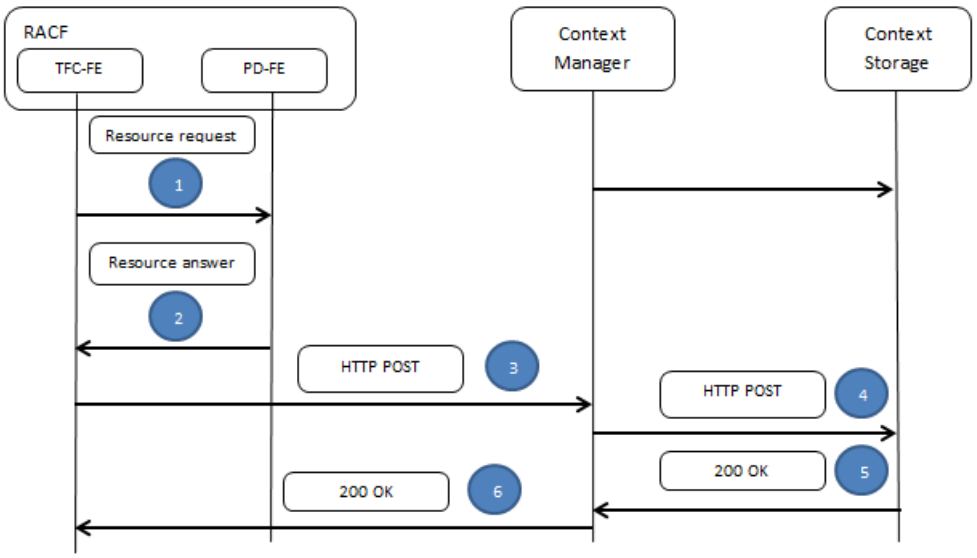

Figure 5.16: Network context transmission

5.3.3.2 Approaches for resource reservation execution

We propose three methods for execution of resource reservation in our MDASH algorithm:

- First approach: A proxy-based approach for redirecting clients HTTP requests to the closest lower representation from the manifest file, which match Video Optimizer results. The proxy intercept client's requests and redirect to lower representation.
- Second approach: In the second approach, the server transcodes the video, according to optimal rates computed by QoE Optimizer by solving the optimization problem. It supposes that, the rate adaptation is done after each optimization step.
- Third approach: The last strategy is to shape the flow of each client according to the Video Optimizer results. The Linux tool tc (traffic control) can be used to apply policy on the streaming server and the client support. The client reacts to the throughput changes and request from the manifest file a bitrate which matches the available client throughput.

5.4 Conclusion

In this chapter, we presented approaches, in order to optimize media delivery across multiple users sharing the same network, The proposed approachs consider end-user satisfaction in terms of MOS in the video adaptation process.

We first study some multi-user QoE metrics and provide general QoE multi-users optimization, that optimizes the network resources. The proposed approach maximizes the QoE of that user which has the lowest QoE among others.

Secondly, we propose an architecture in the home network, that learn from the existing ITU-T Resource and Admission Control Functional architecture in NGN for QoS reservation.

In order to validate the benefits of our approach, we evaluate the Mean Opinion Score (MOS) for different users (in terms of devices and content types) sharing the same network. DASH provides good results for less demanding contents and devices but fails for more demanding ones, while our approach maintains quality for less demanding users and improve the perceived quality for more demanding ones.

In addition, we considered the case where we need to find the necessary capacity for the same QoE for all users. We showed that, considering the contextual information in resource allocation can reduce the needed network capacity.

Chapter 6

Conclusion and Future work

6.1 Contributions summary

Video services are becoming increasingly popular on internet and users' satisfaction is the operators' and service providers' primary aim to reduce the churn, promote new services and improve ARPU (Average Revenue per User). QoE appears as a measure of users' satisfaction from a service through providing an assessment of user's perception. Therefore, there are needs to develop accurate methods for QoE assessment and demands of systems, that are able to show precisely the quality metrics that are really perceived on the user-side, in order to have an aggregated view on the overall content delivery quality. In addition to that, adaptation of the content and its delivery are needed in order to improve the perceived QoE. The adaptation process should consider the measured QoE coupled with context information on the user, devices, network and the content itself to take the adequate adaptation decision.

The contributions of this dissertation are outlined below:

We first analyzed the effect of quality metrics (buffering ratio, video startup time, and average bitrate) on user engagement for popular and unpopular contents during Roland Garros 2013 event. We showed that, the buffering ratio is a critical parameter which impacts user engagement independent on the content type (popular and non popular). Additionally, the video bitrate is a critical parameter which impacts the user engagement up to a certain threshold in the case of popular contents. We find also that video startup time is less significant until a certain level for popular contents.

In order to have an accurate assessment of QoE, we investigate in this thesis by experimentations, the influence of content characteristics, device type, network context and video bitrate on MOS that is the considered QoE metric. We also provided mathematical models that accurately captures MOS function of network and video degradation by considering contextual information in terms of content type and device type. We compared our proposal with a professionnal tool that provide QoE measurements. The results prove that our model is very well correlated to the exprimental data. The impact of startup time and buffering ratio

on end-users satisfaction were also analyzed. In general, for both desktop and smartphone, buffering ratio is a critical metric. The effect of video startup time is more negatively visible in the case of smartphone than the desktop.

As there are needs for content and delivery techniques, we propose in this dissertation QoE optimization based on multi-users QoE metrics. Our approachs improve perceived QoE for different video sessions sharing the same network, and then improves the QoE of that user among the others which has the lowest QoE. We validate the proposal through simulations and propose the architecture for our proposal, that is an extension of the existing ITU-T Resource and Admission Control Functional architecture in NGN for QoS reservation.

6.2 Open Questions and Perspectives

Techniques for measuring Quality of Experience and adaptation for video in Internet have a certain maturity.

This thesis has investigated several problems, that are important for providing techniques for measuring QoE and video adaptation with consideration of context information. Several research contributions for these problems are also provided. But some questions remain open:

In this thesis, we analyzed the impact of quality metrics on user engagement, through gathering and analysis of gigabytes of information over thousands of clients. In addition to collected quality parameters we need, as a next step, to gather other parameters in order to better understand what really happened on client side and what exactly impacted the user engagement. These information could include the CPU usage, the player window size and context information (device type, content type, user preferences, network status etc.). Other leads would be also for instance to fetch external data sources (e.g. social networks), in order to understand the link between socials networks information and user engagement.

In our work, we proposed mathematical models that captures MOS function of network and video degradation by considering contextual information in terms of the content type and device type. But, there are still parts of the models that need further improvements. For example, we have to consider additional context parameters (user preferences, user location, ...) in MOS modeling, because these parameters may influence the user satisfaction. A second open issue, as you know, subjective tests are time consuming and very hard to achieve, especially when there are large number of videos, large configurations to consider in terms of network status, type of video, type of device etc. In our tests for example, we consider only two types of devices (smartphone and desktop). There are needs to consider additionnal devices (Tablet, Connected TV etc.). Thus, we can have more accurate results and generalize our model.

Another perspective that must be solved in QoE modeling, before such method is made available in operationnal system, we have to know, how we can adapt the model, with new configurations. For example, when we have a new device, how will the model parameters evolve?

We proposed a video adaptation approach in home network based on MOS, where we can have different sessions sharing the same resource. Our proposal was validated by simulations. We did not have the time to implement that proposal in a real CPE (Customer Promises Equipment). As a perspective of our work, we have to study the integration of such modules in embedded software as the CPE and what will be the cost.

With the emergence of MPEG-DASH Dynamic Adaptive Streaming over HTTP, that is a new standard used for video streaming. In principle the MPEG-DASH cut a video into several segments recorded each in a separate file. Each segment contains a short playing time interval video. The content (video) is encoded in different bitrates and the customer selects automatically segment based on current network conditions. The developped model i.e., the VCUPM (Video Contextual Perceptual Model) already takes into account the impact of video bitrate for considered contextual information (device and content type) and thus can be used for each video segment for assessing QoE for that short playing time interval. There are also needs to develop a predictive model of QoE for video in internet, that takes into account the startup time, video bitrate, buffering ratio and rebuffering events.

Bibliography

[1] M. T. Diallo, H. Moustafa, H. Afifi, and K. Laghari, "Quality of experience for audio-visual services," *In EuroITV*, 4 to 6 July 2012, Berlin, Germany.

[2] "3GPP, TR 26.944 End-to-end multimedia services performance metrics," December 2008.

[3] "ITU Recommendation. P.10/g.100, Amendment 2 New definitions for inclusion," July 2008.

[4] "ETSI, Human Factors (hf), Quality of Experience (QoE) requirements for real-time communication service," January 2010.

[5] M. T. Diallo, H. Moustafa, H. Afifi, and N. Maréchal, "Adaptation of Audio-Visual Contents and their Delivery," *In Communications of the ACM magazine*, November 2013.

[6] L. K. ur Rehman and K. Conelly, "Towards total quality of experience: A qoe model in communication ecosystem," *In IEEE Communication Magazine*, pp. 58 – 65, April 2012.

[7] F. F. Mamadou Tourad Diallo and J.-B. Hennequin, "Impacts of quality of experience in user engagement in a live event," *In IEEE International Conference on Multimedia and Expo (ICME)*, July 14-18, 2014, Chengdu (Chine).

[8] M. T. Diallo, H. Moustafa, N. Marechal, and H. Afifi, "A hybrid contextual user perception model for streamed video quality assessment," *In IEEE International Symposium on Multimedia (ISM 2013)*, Anaheim, California, USA, 2013.

[9] M. T. Diallo, F. Fieau, E. Abd-elrahman, and H. Afifi, "Utility-based Approach for Video Service Delivery Optimization," *In International Conference on Systems and Networks Communications (ICSNC)*, October 12-16, 2014, Nice.

[10] L. Mittal and R. Mittal, "Internet protocol television (IPTV)," *In International Journal of Electronics and Computer Science Engineering*, 2012.

[11] V. Sze, M. Budagavi, and G. J. Sullivan, *High Efficiency Video Coding (HEVC)*. Springer, 2014.

[12] "ETSI Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); NGN integrated IPTV subsystem architecture," January 2010.

[13] S. SONG, *Advanced Personalization of IPTV Services.* PhD thesis, Télécom SudParis and Université Pierre et Marie Curie, 2012.

[14] H. Moustafa and S. Zeadally, *Media Networks: Architectures, Applications, and Standards.* CRC Press, 2012.

[15] "Technical report : Informa telecoms and media content delivery networks: Market dynamics and growth perspectives," 2012.

[16] Playready, "http://www.microsoft.com/playready/overview/," Last visit : February 2015.

[17] P. Y. L. Lann, N. Bihannic, and O. le Grand, "Techinal report : Core audiovisual convergence fdm1: Overview of existing webtv& mobiletv architecture," 2010.

[18] Anevia, "White paper : OTT streaming -2nd edition," tech. rep., September 2011.

[19] H. Moustafa, N. Marechal, and S. Zeadally, "Mobile multimedia applications: Delivery technologies," *IT Professional*, pp. 12–21, 2012.

[20] O. Oyman and S. Singh, "Quality of experience for HTTP adaptive streaming services," *IEEE Communications Magazine*, pp. 20 – 27, April 2012.

[21] "https://developer.apple.com/streaming/," Last visit : February 2015.

[22] F. Daoust, P. Hoschka, C. Z. Patrikakis, R. S. Cruz, M. S. Nunes, and D. S. Osborne, "Towards video on the web with html5," *W3C/ERCIM*, 2004.

[23] Smooth Streaming, "http://www.iis.net/downloads/microsoft/smooth-streaming," Last visit : February 2015.

[24] Silverlight, "http://www.microsoft.com/silverlight/," Last visit : November 2014.

[25] HDS, "http://www.adobe.com/devnet/hds.html," *Last visit : February 2015.*

[26] "3GPP TS 26.247, Progressive Download and Dynamic Adaptive Streaming over HTTP (3GP-DASH) , (Release 10)," November 2011.

[27] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hoßfeld, and P. Tran-Gi, "A Survey on quality of Experience of HTTP Adaptive Streaming," *In IEEE Communications Surveys & Tutorials*, Septembre 2014.

[28] H. Schulzrinne, S. Casner, R.Frederick, and V.Jacobson, "A Transport Protocol for Real-Time Applications.," *In: IETF RFC 3550*, 2003.

[29] RTMP, "http://www.adobe.com/devnet/rtmp.html," Last visit : December 2014.

[30] MMS, "https://msdn.microsoft.com/en-us/library/cc239490.aspx," Last visit : February 2014.

[31] "Real time streaming protocol (rtsp)," *In: IETF RFC 2326*, 1998.

[32] I. Martinez-Yelmo, I. Seoane, and C. Guerrero, "Fair Quality of Experience (QoE) Measurements Related with Networking Technologies," *Springer-verlag*, 2010.

[33] A. N. Netravali and B. G. Haskell, *Digital Pictures: Representation, Compression, and Standards (2nd Ed)*. Plenum Press, 1995.

[34] PEVQ, "http://www.pevq.org/video-quality-measurement," Last visit : November 2014.

[35] M. Pinson and S. Wolf, "A New Standardized Method for Objectively Measuring," *In IEEE Transactions on Broadcasting*, 2004.

[36] Z. Wang and Q. Li, "Video quality assessment using a statistical model of human," *In Journal of the Optical Society of America*, 2007.

[37] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual bluc metric," *In Image Processing, IEEE International Conference*, 2002.

[38] A. Rossholm and B. Lovstrom, "A New Low Complex Reference Free Video Quality Predictor," *In Multimedia Signal Processing, IEEE 10th Workshop*, 2008.

[39] L. Ding and R. Goubran, "Speech quality prediction in voip using the extended e-model," *In IEEE GLOBECOM*, 2003.

[40] H. A. Ahmed Meddahi and D. Zeghlache, "Packet-E-model for Wireless VoiP Quality Evaluation," *In IEEE lnternalionai Symposium on Personai,lndoor and Mobile Radio Communication Praceedings*, 2003.

[41] L. K. ur rehman, N. I. Khan, and N. Laghari, "Qualitative and Quantitative Assessment of Multimedia Services in Wireless environment," *MoVid, ACM Multimedia system*, February 2012.

[42] K. D. Singh, A. ksentini, and B. Marienval, "Quality of experience measurement tool for svc video coding," *In, IEEE International Conference on Communications (ICC)*, 2011.

[43] G. Rubino, "The PSQA project, INRIA Rennes-Bretagne Atlantique, http://www.irisa.fr/armor/lesmembres/rubino/mypages/psqa.html.," Last visit : December 2014.

[44] M. Fiedler, T. Hoßfeld, and P. Tran-Gia, "A generic quantitative relationship between Quality of Experience and Quality of Service," *In Network, IEEE*, pp. 36 – 41, March-April 2010.

[45] M. Fiedler, T. Hoßfeld, and P. Tran-Gia, "Quality of experience-related differential equations and provisioning-delivery hysteresis," *IEEE Network Special requirements for Mobile TV, 13th ACM, International Conference on Multimedia*, 2010.

[46] J. Kim, T.-W. Um, W. Ryu, and B. S. Lee, "Heterogeneous networks and terminal-aware qos/qoe-guaranteed mobile iptv service," *In IEEE Communications Magazine*, May 2008.

[47] S. Mohan and N. Agarwal, "A convergent framework for QoS-driven social media content delivery over mobile networks," 2011.

[48] J. Koo and K. Chung, "Marc: Adaptive rate control scheme for improving the qoe of streaming services in mobile broadband networks," *In IEEE Communications and Information Technologies (ISCIT)*, 2010.

[49] J. C. Tiantian Guo and C. H. Foh, "Scalable video adaptation in wireless home networks with a mixture of iptv and vod users," *In IEEE Globecom*, 2011.

[50] R. Hammi and K. Chen, "Some experience on video flow regulation with an active network approach," *Dagstuhl Seminar Proceeding*, 2005.

[51] E. J. A. Khan, L. Sun and E. Ifeachor, "Quality of experience-driven adaptation scheme for video applications over wireless network," *IEEE Communications IET*, 2009.

[52] T. Laakko and T. Hiltunen, "Adapting Web Content to Mobile User Agents," *In IEEE internet computing*, March-April 2005.

[53] R. Sylvia, H.Marc, K. Maning, and S. Scott, "Topologically-aware overlay construction and server selection," *In INFOCOM*, pp. 1190–1199., 2002.

[54] J.Xueying, L. Shiyao, and Y.Yang, "Research of load balance algorithm based on resource status for streaming media transmission network," *In International Conference on Computer Engineering and Network (CENet)*, pp. 503–507, 2013.

[55] A. Ksentini, C. Viho, and J. Bonnin, "QoE-aware Vertical Handover in Wireless Heterogeneous Networks," *In IEEE Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 95–100., 2011.

[56] B.Ciubotaru and G.-M. Muntean, "SASHA : A Quality-Oriented Handover Algorithm for Multimedia Content Delivery to Mobile User," *In IEEE Transactions Broadcasting*, June 2009.

[57] L. Suciu and K.Guillouard, "A Hierarchical and Distributed Handover Management Approach for Heterogeneous Networking Environments ICNS," *In International Conference on Networking and Services*, June 2007.

[58] Z. Aouini, M. T. DIALLO, A. Gouta, A.-M. Kermarrec, and Y. Lelouedec, "Improving caching efficiency and quality of experience with cf-dash," *In ACM workshop on Network and Operating Systems Support for Digital audio and video (NOSSDAV 2014),*, Singapore March 2014.

[59] M. T. Diallo, H. Moustafa, N. Maréchal, H. Afifi, and E. Abd-elrahman, "Optimization of tv multicast delivery," *Journal of Electronic System*, December 2013.

[60] D. Pichon, G. Rieublandou, J. Bonnin, and P. Seite, "Coordinated Service and Mobility Management to Maximize User Qoe and Optimize Network Resource Use (ieee) ,," *In IEEE 20th International Symposium, Personal, Indoor and Mobile Radio Communications*, 2009.

[61] "http://www.conviva.com," Last visit : November 2014.

[62] "Fortinet : http://www.fortinet.com," Last visit : November 2014.

[63] "Skytide: http://www.skytide.com," Last visit : November 2014.

[64] F. Dobrian, V. Sekar, I. Stoica, H. Zhang, A. Awan, D. Joseph, and A. Ganjam, "Understanding the impact of video quality on user engagement," *In ACM SIGCOMM Computer Communication*, pp. 362–373, August 2011.

[65] "Roland garros : http://www.rolandgarros.com/index.html," Last visit : December 2014.

[66] M. Lalmas, "User engagement, a scientific challenge," 2012.

[67] Peterson and Carrabis, "Measuring the immeasurable : visitor engagement, webanalytics demystified," September 2008.

[68] H. otto Georgii, *Introduction to Probability and Statistics*. Walter de Gruyter, February 2008.

[69] T. Hoßfeld, S. Egger, R. Schatz, M. Fiedler, K. Masuch, and C. Lorentzen, "Initial delay vs interruptions : Between the devil and the deep blue sea," *In IEEE Fourth International Workshop on Quality of Multimedia Experience (QoMEX)*, July 2012.

[70] H. Abdi, "Kendall rank correlation," *In Salkind, N.J. Encyclopedia of Measurement and Statistics. Thousand Oaks (CA): Sage*, 2007.

[71] E. M. R. Peter J. Huber, *Robust Statistics 2nd Edition*. Wiley, March 2009.

[72] A. Balachandran, V. Sekarz, A. Akellay, S. Seshan, I. Stoica, and H. Zhang, "Developing a predictive model of quality of experience for internet video," *In SIGCOMM Proceedings of the ACM*, pp. 339–350, October 2013.

[73] "http://www.ntt.co.jp/qos/qoe/eng/technology/visual/," Last visit : November 2014.

[74] F. Agboma, *Quality of experience management in mobile content delivery systems*. PhD thesis, University of Essex, July 2009.

[75] A. A. Webster, C. T. Jones, M. H. Pinson, S. D. Voran, and S. Wolf, "An objective video quality assessment system based on human perception," pp. 15–26, 1993.

[76] H. Knoche, J. D. McCarthy, and M. A. Sasse, "Can small be beautiful? assessing image size requirements for mobile tv," *In13th ACM, International Conference on Multimedia*, 2005.

[77] C.Sodergard, "Technical report: Mobile television : Technology and user experiences on the mobile-tv project," *VTT Information Technology*, 2003.

[78] E. W. Orange, "Technical report, aquatool ,orange labs, rennes," 2012.

[79] K. Atkinson, *An Introduction to Numerical Analysis*. John Wiley & Sons, 1989.

[80] H. T. Yamagishi K., Tominaga T and T. A., "Objective quality estimation model for videophone services," *NTT Technical Review*, Vol.5, No. 6, June 2007.

[81] ITU Rec. Y. 2111, "Resource and admission control functions in next generation networks (release 1)," 2006.

[82] ITU Draft Y. 2111, "Resource and admission control functions in next generation networks (release 2). ngn.gsi/doc-301," 2007.

[83] Y. Kim and I. Chong, "A Resource Control Mechanism on NGN basedHome Network for IPTV Service," *The 12th International Conference on Advanced Communication Technology (ICACT)*, February 7-10, 2010.

[84] K. Y. and C. I., "A Resource Control Mechanism on NGN based Home Network for IPTV Service," *The 12th International Conference on Advanced Communication Technology (ICACT)*, February 7-10, 2010.

[85] S. Dixit and T. wu, *Content networking in the mobile internet.* John wiley and Sons, 2004.

[86] "Network functions virtualization nfv, use cases," no. ETSI GS NFV 001 v1.1.1, October 2013.

[87] J. Penhoat, K. Guillouard, T. Lemlouma, and M. Salaun, "Analysis of the Implementation of Utility Functions to Define an Optimal Partition of a Multicast Group," *In International Conference on Networks (ICN)*, 2011.

Appendix A

Utility based approach for Video Service Delivery Optimization

Many operators are searching for new optimizations techniques that can achieve the balance between the main three actors in the chain (Content Providers, Operators and Clients). But, the massive deployment of Over-The-Top (OTT) technology is really representing a big threat for managed video services. Moreover, new opportunities brought by clients need to be studied in order to build a good utility between users needs, network characteristics and service requirements. Therefore, searching optimization algorithms and tools for managed video delivery networks is required.

In the state of art, there are techniques for video delivery optimization based on utility functions, that take into account the aim of each actor on the delivery chain (Content Providers, Operators and Clients). But some relevant actors parameters are omitted. Therefore, there are needs to define a new video delivery optimization which takes into account the key parameters.

In our optimization approach, we extend the existing approach, by integrating some relevant parameters, then a global utility function is calculated based on different constraints. Those constraints are based on separate utility function for each actor in video service delivery (content provider, operator and client), as described in work [9].

However, each actor has a global score for his vision, the overall optimization aims to satisfy the three actors. Our proposed methodology for this optimization is validated through simulation based utility function for obtaining the optimal values of our optimization problem. Then, a complete GUI interface is built based on the main parameters for each actor.

Hereinafter, we will explain the main challenges in video data centers in general and conduct a comparison between the main actors in video service delivery.

Comparison of Actors

It is important to analyze the main actors in video service delivery chain. Then, we can describe the objectives of each actor in order to introduce his utility and the overall work motivations. Here, two comparisons are mandatory in order to build our utilities and have clear problem statements as follows:

- Agility Comparison: The Agility is defined as the number of parameters and the ability of adaptation for the proposed system dynamically. So, the flexibility of service planning either for content adaptation or server placement is considered as an important factor in any video streaming chain. Thus, either for live streaming or VoD (Video on Demand), the easy adaptation and simple configuration of networks will enhance the overall system performance and users satisfactions at same time. Moreover, the correlation between the three actors in the video chain will lead to an optimal identification for both network capacities and users densities. Table A.1 compares the Agility of the three actors effects in terms of some major attributes as follows:

Attribute	Content Provider	Operator	Client
Capacity	Maximizing the throughputs	Minimizing the network load	Maximizing the number of clients
Quality	QoS SLA/TCA between CP & OP for an efficient content delivery with min and max thresholds of quality.	Quality of service measures for adaptive bit rates	Participating in QoS/QoE reports for enhancing the overall service delivery
Device	Hardware or Software consumed for contents visualizations or services on demand	Dynamic allocations for resources and network virtualization to cope with on demand servers caching or placements	Device capabilities to fit with different access networks and with virtual applications

Table A.1: Agility comparison for the three actors

- Cost Comparison: Table A.2 gives an overall cost comparison from each actor view as follows:

Attribute	Content Provider	Operator	Client
CAPEX cost	Min cost for content adaptations	Min transmission cost for each content	Min cost for required bandwidth
OPEX cost	Hosting servers for different layers of same content	Running cost for QoS SLA/TCA between CP & OP	Running cost for additional Bandwidth

Table A.2: Cost comparison for the three actors

Based on the previous two proposed comparisons and main issues in service delivery, we can formulate our problem statements as follows:

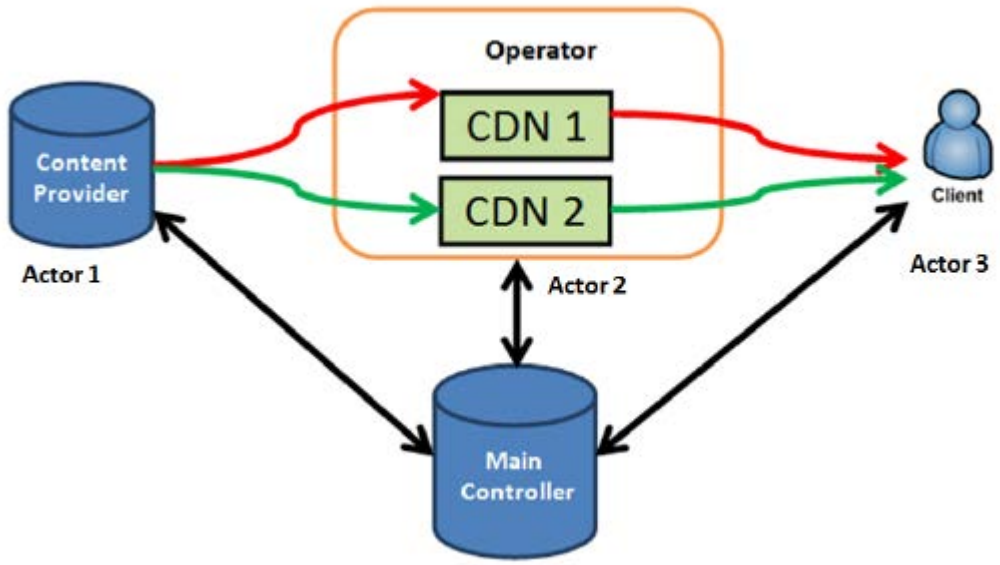

Figure A.1: Main three actors in the video chain

- Problem statement: We propose a global optimization utility function for each one of the three actors in the video chain. As shown in Figure A.1, the three actors in the chain are in collaboration for the best service delivery. Actor 1, the content provider asks Actor 2 (the operator) to deliver some video content requested by the third Actor 3 (client). We assume that the system is real time so requests can be handled through some controller unit

that manages sessions and handover decisions between CDNs based on our optimization function

Proposed methodology

The purpose of this section is to explain the steps of the optimization approach which takes into account the objective of Content Provider (CP), Operator (OP) and the Client (CL). Our approach is based on the definition of three entities, each with their goals as follows:

- The objective of Content Provider is to send the Content in the network with a minimum cost and still manage the Content expectations in terms of requirements (for example the minimum required throughput for the content).
- The objective of the Operator is to transmit content on its network (CDN1 or CDN2 in our example) while keeping the load as lower as possible.
- The objective of the client is to improve the Quality of Experience besides the QoS (Quality of Service).

The Utility-based Video Service Delivery Optimization (U-VSDO) will take into account the goals of each actor in addition to the main constrains. As shown in Figure A.1, the optimization decision will be managed by the Main Controller after solving the optimization problem. This controller can be for example an SDN controller as will be explained in Section IV for SDN Network Function Virtualization NFV [86].

So, we can solve the problem by the following steps: Problem Formulation We used the utility functions to calculate the scores of each actor; this is very useful to characterize the satisfaction derived from a parameter. The function must have the following characteristics:

- The function increases with parameter x and has a maximum of 1.
- When x is "low", the function tends to zero.
- The possibility to have normalized results between [0, 1].

Several functions meet these criteria. Moreover, we decided to use the utility function: $1 - e^{-x}$, as the work in [57][87], where x is a parameter of the function. In future work, we will further investigate the influence of others utility functions in our optimization problem.

Hereinafter, we introduce the details of each actor utility function based on the previous propositions either for utility type or normalization way. Then, a

global score utility will be calculated under the main constrains defined for each actor as follows.

As the work in [57], we have two types of parameters:

- The positives parameters: High values are better, example (throughput, available hardware, etc.), then for an utility function we took the parameter directly.
- The negatives parameters: Low parameters are better, example (cost, network load, etc., then for these parameters we choose $\frac{1}{cost}$ for example.

Content Provider

In this section, we provide the utility function associated to the content provider parameters.

$$S_{cp}(i,j) = (1 - e^{-\frac{1}{C_{cp}(j)}} + 1 - e^{-D_r(i)}).C_{cps}.D_s \tag{A.1}$$

$$S_{cp}(i,j) = (2 - e^{-\frac{1}{C_{cp}(j)}} - e^{-D_r(i)}).C_{cps}.D_s \tag{A.2}$$

where:

- $S_{cp}(i,j)$, is the score related to Content Provider for flow j in network i
- C_{cp}= UNIT cost per Mbyte, that is a cost of transmitting the content in the network (CDN1 or CDN2) in our example.
- C_{cp}= ($C_{cpmax}(j)$, $C_{cp}(j)$) =0 , when $C_{cpmax} < C_{cp} \Rightarrow S_{cp} = 0$
- C_{cpmax}, is the maximum cost that the content provider is ready to pay
- D_r, the available throughput
- D_s= ($D_{ref}(j)$,$D_r(i)$) =0, when $D_{ref} < D_r$
- D_{ref}, is the required video bitrate

Note that: (A, B) means, when $A < B$, $\Rightarrow (A, B)$=0

Network Operator:

In this section, as in the precedent one, we provide the utility function associated to network parameters.

$$S_{op}(i) = (3 - e^{-\frac{1}{C_{op}(i)}} - e^{-\frac{1}{NL(i)}} - e^{-H(i)}).NL_s.C_{ops}.H_s \tag{A.3}$$

where:

- $S_{op}(i)$, is the score related to Operator (OP) for network i

- C_{op}, is the cost from the operator side
- NL, is the network load
- NL_s= $(NL_{max}(i), NL(i))$=0, when $NL_{max} < NL$
- NL_{max}, is the maximum acceptable network load
- C_{ops} = ($C_{opmax}(i)$,$C_{op}(i)$)=0, when $C_{opmax} < C_{op}$
- C_{opmax}, is the maximum price that the operator is ready to invest
- H is the required hardware threshold
- H_s= ($H(i)$,$H_{min}(i)$) =0 , when H ¡ H_{min}
- H_{min}, is the minimum required hardware for considered service

Client

In order to estimate the client satisfaction, we propose the following equation:

$$S_{cl}(i,j) = \frac{NCUPM(i,j)}{Smax} \tag{A.4}$$

where:

- $S_{cl}(i,j)$ is the score for user in network i for flow j.
- $\varphi_N(i,j)$ corresponds to the satisfaction obtained by users in network i for flow j. It is a parametric model which computes the Quality of Experience function of contexts information, the model takes into account parameters such as the device type, the video content type and the quality of the network link as described in precedent chapter. To recall, the equation is presented as below:

$$\varphi_N(D_{ri}) = \alpha + \beta e^{-\delta \frac{D_v}{D_{ri}}}$$

- α,β and δ are the model parameters calculated by using subjective test data from different experiments.
- S_{max}, is the maximum value of φ_N which correspond to the normalized factor.

General Problem Optimization and constraints

In this section, we present the general optimization problem and summarize the main utility functions for the computed scores and their constrains that will be implemented in the next section and appeared in the GUI (Graphical User Interface).

The general optimization problem can be formulated as follows by total score:

$$S_T = \omega_1.S_{cp} + \omega_2.S_{op} + \omega_3.S_{cl} \tag{A.5}$$

where:

ω_1,ω_2 and ω_3: are the weights of entities in the global optimization and $\omega_1 + \omega_2 + \omega_3 = 1$

The weighting parameters define the importance of each actor in the optimization decision. In our work we decided that the Content Provider, the Operator and Users have the same weight, then $\omega_1 = \omega_2 = \omega_3 = \frac{1}{3}$

The objective of optimization problem is to:

$$maximize(\omega_1.S_{cp} + \omega_2.S_{op} + \omega_3.S_{cl}) \tag{A.6}$$

subject to:

$$C_{cp} < C_{cpmax} \tag{A.7}$$

$$C_{op} < C_{opmax} \tag{A.8}$$

$$D_r < D_{ref} \tag{A.9}$$

$$NL < NL_{max} \tag{A.10}$$

$$H_{min} < H \tag{A.11}$$

Implementation and Evaluation

To validate our work, we are going to optimize the utility function parameters through a simulation tools using Matlab. Then, the decision output of this optimization will take the form of graphical interface for doing many scenarios.

We implemented a complete Graphical User Interface (GUI) to be used by the operators in their networks design and optimization. This graphical tool is built based on Matlab code.

Figure A.2 illustrates the main construction steps as divided into two parts:

- Creating general parameters: which means defining the basic topology elements and factors in the three actors (CP, OP and CL) i.e. the main profiles for each video and CDN.

- Calculating results: Calculating the general score for all actors and show the selected CDN as best path for video profile. Actually, we simulate the global utility function and calculate the scores for different networks for our approach U-VSDO.

Moreover, and in order to facilitate the decision making output by each operator running our methodology, we developed GUI interface to cope with the three utility functions for the three actors main parameters as shown in Figure A.2.

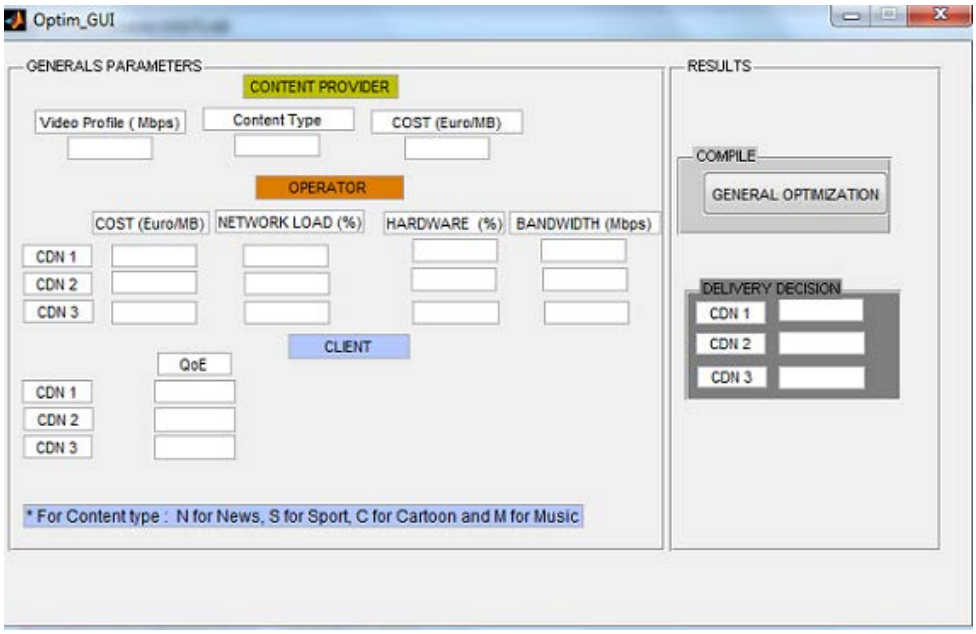

Figure A.2: GUI interface for U-VSDO approach

After finishing this simulation, we conducted a brief comparison between our approach U-VSDO and other similar techniques that used utility functions for decision making based multimedia delivery adaptation like like SASHA [56] and HDHO [57]. The detailed functionality about these two references methods are given in chapter 2.

The results indicated in Table A.3 highlighted the main parameters considered as supplementary by our approach U-VSDO over other ways.

Considered parameters	Optimization techniques		
PARAMETERS	SASHA	HDHO	U-VSDO
Operator cost	No	No	Yes
Content Provider cost	Yes	Yes	Yes
Content Type	No	No	Yes
Device Type	No	No	Yes
QoE	Yes	No	Yes
Network Load	Yes	Yes	Yes
Hardware status	No	No	Yes

Table A.3: Comparison between U-VSDO and other approaches

The Smooth Adaptive Soft-Handover Algorithm (SASHA) makes roaming through different network by computing an score. The delivery decision is based on the network that has the best score. In order to make decision, this method does not consider the cost that the operator has to invest to transmit the content, the used terminal is not taken into account and the hardware status in the network is not also considered for decision making.

As SASHA, the Hierarchical and Distributed Handover (HDHO) is also a distributed handover decision framework. In this approach some relevant parameters are omitted. In content provider side the cost of transmitting the content in a network is missed, in network side the cost and hardware status are absent, in client side the perceived Quality of Experience is not taken into account.

Our video delivery approach extends the HDHO technique, that already considers the network load and the cost that content provider is ready to pay. UVSDO integrates more parameters, that are relevant for video delivery decision making like: the operator cost (how much the operator is ready to pay), content type (contents don't have the same expectations), the type of device (devices don't have the same expectations), hardware status (available memory, CPU, storage capacity ...), the QoE (the end-user satisfaction).

Conclusion

In this chapter, we proposed the U-VSDO that is an delivery optimization mechanism. It solves the utility function optimization for the three common actors in video streaming chain, including their roles and objectives in video chain. The proposed methodology is evaluated through a simulation for global utility function and our approach is more accurate because, it considers more relevant parameters for video delivery decision making, compared to other methods like HDHO or SASHA.