Universidade de Aveiro
2014

Departamento de
Electrónica, Telecomunicações e Informática

André Ferraz Salvador

Avaliação da Qualidade de Experiência de Vídeo em várias tecnologias

"Deus quer, o Homem sonha
e a obra nasce."

— Fernando Pessoa

Universidade de Aveiro Departamento de Electrónica, Telecomunicações e Informática
2014

André Ferraz Salvador

Avaliação da Qualidade de Experiência de Vídeo em várias tecnologias

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requesitos necessários à obtenção do grau de Mestre em Engenharia Electrónica e Telecomunicações, realizada sob a orientação científica de Susana Sargento, Professora do Departamento Engenharia Electrónica e Telecomunicações da Universidade de Aveiro e co-orientação do Doutor Pedro Neves, PT Inovação.

o júri / the jury

presidente / president

Professor Doutor José Ferreira da Rocha
Professor Catedrático do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro

vogais / examiners committee

Professora Doutora Marília Pascoal Curado
Professora Auxiliar do Departamento de Informática da Universidade de Coimbra (arguente)

Professora Doutora Susana Isabel Barreto de Miranda Sargento
Professora Auxiliar do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro (orientadora)

agradecimentos / acknowledgements

É com esta grande oportunidade que agradeço o conhecimento/crescimento adquirido ao longo destes 5 anos nesta academia.

Em primeiro, quero agradecer à Professora Susana Sargento pelo apoio e formação que sempre me prestou.

À PT inovação, Engenheiro Pedro Fernandes, Engenheiro João Nogueira e Engenheiro Simão Cardeal pelo suporte e apoio que me foi crucial para este trabalho.

A todo o grupo de redes (NAP) que sempre esteve disponivel para ajudar.

A todos os meus amigos que sempre me apoiaram e me ajudaram.

Por último, à minha familia, Pais, Irmã, Tios, Primos, Avós e Namorada que sempre me deram o melhor e me motivaram sempre da melhor maneira.

Resumo

Nos dias de hoje a Internet é um dos meios com mais serviços associados. Conjugado a este facto, existe um acentuado aumento de utilizadores a aderir a este serviço. Nesta perspectiva existe a necessidade de garantir uma qualidade mínima por parte dos prestadores de serviços.

A Qualidade de Experiência que os consumidores têm dos serviços é bastante crucial no desenvolvimento e optimização dos serviços nas redes. É ainda de salientar que o aumento do tráfego multimédia, nomeadamente os *streamings* de vídeo, apresenta incrementos na probabilidade de as redes se congestionarem. Na perspectiva do prestador de serviços a monitorização é a solução para evitar a saturação total.

Neste sentido, esta dissertação pretende desenvolver uma plataforma que permite a monitorização do tráfego de multimédia do serviço do Meo Go, fornecido pela operadora Portugal Telecom Comunicações.

Neste trabalho foi necessário investigar e testar a arquitectura do *streaming* adaptativo sobre HTTP para ser possível obter métricas de qualidade de experiência. Este *streaming* adaptativo apresenta a técnica de *smooth streaming*, sendo esta arquitectura projectada pela empresa *Microsoft* e utilizada no serviço Meo Go.

Posteriormente foram monitorizadas as métricas que se obtiveram no player de vídeo. Esta análise foi realizada de forma objectiva e subjectiva. Nesta fase da implementação objectiva do método em que se pretende obter uma predição do valor de Qualidade de Experiência por parte do consumidor, foram seleccionadas as métricas oriundas do estado/desempenho da rede e do dispositivo terminal. As métricas obtidas entram num processo de tratamento que pretende simular a acção humana nas classificações da qualidade dos vídeos. De outra forma, subjectivamente, foi realizada uma pesquisa, com base num questionário, de modo a comparar os métodos. Nesta etapa foi gerada uma plataforma *online* que possiblitou obter um maior número de classificações dos vídeos para posteriormente se proceder ao tratamento de dados.

Nos resultados obtidos, primeiramente ao nível do player de *smooth streaming*, estes permitem analisar a técnica de implementação de *streaming* adaptativo. Numa fase seguinte foram criados cenários de teste para comprovar o funcionamento do método em diversas situações, tendo com maior relevância aqueles que contêm dinâmicas mais complexas. Na perspectiva dos métodos subjectivo e objectivo, estes apresentam valores que confirmam a arquitectura do módulo implementado. Adicionalmente, o desempenho do método em classificar a qualidade de serviço de vídeo *streaming*, ao longo do tempo, apresentou valores que se aproximam da dinâmica esperada numa acção mental humana.

Abstract

Nowadays the internet is associated with many services. Combined with this fact, there is a marked increase of the users joining this service. In this perspective, it is required that the service providers guarantee a minimum quality to the network services.

The Quality of Experience of services is quite crucial in the development of services in networks. Also noteworthy, the traffic increase in multimedia services, including video streaming, increases the probability of congesting the networks. In the perspective of the service provider, the monitoring is a solution to avoid saturation in network.

This way, this dissertation proposes to develop a platform that allows a multimedia traffic monitoring in the Meo Go service provided by the operator Portugal Telecom Communications.

The architecture of the adaptive streaming over HTTP has been studied and tested to obtain the quality of experience metrics. This adaptive streaming technique presents the smooth streaming, an architecture made by Microsoft company, and it is used in the Meo Go service. Then, it is monitored the metrics obtained with the video player. This analysis is done objectively and subjectively. In this phase, the objective implementation of the method allows to obtain the prediction value of the Quality of Experience by consumers. The selected metrics were derived from the state / performance of network and terminal device. The obtained metrics aim to simulate human action in video score quality. Otherwise, subjectively, it is conducted a survey based in a questionnaire to compare methods. In this phase it was created an on-line platform to allow the obtain a greater number of rankings and data processing.

In the obtained results, firstly in the smooth streaming player, it is shown the adaptive streaming implementation technique. On the next phase, test scenarios were created to demonstrate the functioning of the method in many cases, with greater relevance for those ones with higher dynamic complexity. From the perspective of subjective and objective methods, these have values that confirm the architecture of the implemented module. Over time, the performance of the scoring the quality of video streaming services approaches the one in a human mental action.

Contents

List of Figures

List of Tables

Acronyms

3GPP	Third Generation Partnership Project
AAC	Advanced Audio Coding
AACL	Advanced Audio Coding Low Complexity
ADSL	Asymmetric Digital Subscriber Line
AVC	Advanced Video Coding
BMP	Bitmap
CDMA	Code division multiple access
CDT	Content Delivery Techniques
CFTOOL	Curve Fitting Tool
CI	Confidence Interval
CIF	Common Intermediate Format
Codec	Coder-Decoder
CSS	Cascading Style Sheets
DASH	Dynamic adaptive streaming over HTTP
DRM	Digital Rights Management
DVR	Digital Video Recorder
EPG	Electronic Program Guide
FLV	FlashVideo
FPS	Frames per second
FTP	File Transfer Protocol
FR	Full Reference

GSM	Global System for Mobile
HAS	HTTP adaptive streaming
HD	High Definition
HTML	HyperText Markup Language
HTTP	Hypertext Transfer Protocol
IEEE	Institute of Electrical and Electronics Engineers
IIS	Internet Information Services
IP	Internet Protocol
IPTV	Internet Protocol Television
ITEC	Institute of Information Technology
JSFL	JavaScript Flash language
LCD	Liquid Crystal Display
LED	Light Emitting Diode
MCRG	Multimedia Communication Research Group
MOS	Mean Opinion Score
MP3	Music file (MPEG Layer 3)
MPEG	Moving Picture Experts Group
MPEG(ITU)	Motion Pictures Expert Group
MATLAB	Matrix Laboratory
OSMF	Open Source Media Framework
PT	Portugal Telecom
PSNR	Peak Signal-to-Noise ratio
PEVq	Perceptual Evaluation of Video Quality
PHP	Hypertext Preprocessor
PIFF	Protected Interoperable File Format
QCIF	Quarter Common Intermediate Format
QoE	Quality of Experience

QoS	Quality of Service
QT	QuickTime
RGB	Red, Green and Blue
RR	Reduced Reference
RTCP	Real-Time Transport Control Protocol
RTMP	Real-Time Messaging Protocol
RTP	Real-time Transfer Protocol
RTSP	Real Time Streaming Protocol
RTSPT	Real Time Streaming Protocol over TCP
RTSPU	Real Time Streaming Protocol over UDP
sftool	Surface Fitting Toolbox
SMPTE	Society of Motion Picture and Television Engineers
SSIM	Structural SIMilarity
STSQ	Short-Time Subjective Quality
TCP	Transport Control Protocol
TV	Television
TVSQ	Time-Varying Subjective Quality
UDP	User Datagram Protocol
UHDTV	Ultra High Definition Television
VGA	Video Graphics Array
VQM	Video Quality Metric
VOD	Video on Demand
VoIP	Voice over Internet Protocol
VOL	Video on Live
WCDMA	Wide-Band Code-Division Multiple Access
WMV	Microsoft Windows Media Video
WMS	Windows Media Services

XML	Extensible Markup Language
XP	Extended Profile

Chapter 1

Introduction

1.1 Motivation

Nowadays the internet is almost a basic service in the human species, and since it has been created, it has been growing and integrating more complexity. The internet network development has added multi-services such as Voice over Internet Protocol (VoIP), Hypertext Transfer Protocol (HTTP), File Transfer Protocol (FTP) and others.
The traffic through the Internet has also increased significantly. Figure 1.1 is a forecast by Cisco Mobile and the video presents the highest percentage of mobile traffic with 69.1% in 2018. The multimedia services products are increasingly being transported only over IP: "These technologies provide an all-IP infrastructure from the mobile device" [17]. The services of broadcast television are increasing in the world with the video-club service and video television service. The video transmission occupies a large of the bandwidth and traffic.

The streaming has been the most viable way to transport video to all devices: it does not need to store large amounts of video on the client, only a buffer (cache memory) is required in order to play a few seconds. This streaming video requires the network with large performance in terms of delay and bandwidth. The transmission over Transport Control Protocol (TCP) requires that the network exhibits low loss and high bandwidth for the minimum delay.

The quality of the streaming video cannot depend only on the quality of service, because the Coder-Decoder (Codec) mechanism and the performance of the terminal device are factors that can greatly influence. The many popular Codecs are used with loss to allow that the files have less size. This factor that degrades the video quality is not linear, because there are losses in the coding, but the ratio of bandwidth of video quality can be incremented with these intelligent methods of the Codec. There are codecs that do not present losses in the coding of video, but the video file will have a larger size. The service provider wants that the streaming broadcast has the highest quality with lower bandwidth, so it is important the relation between these two factors. The streaming quality is also affected by the client terminal, because the device performance is influenced by the number

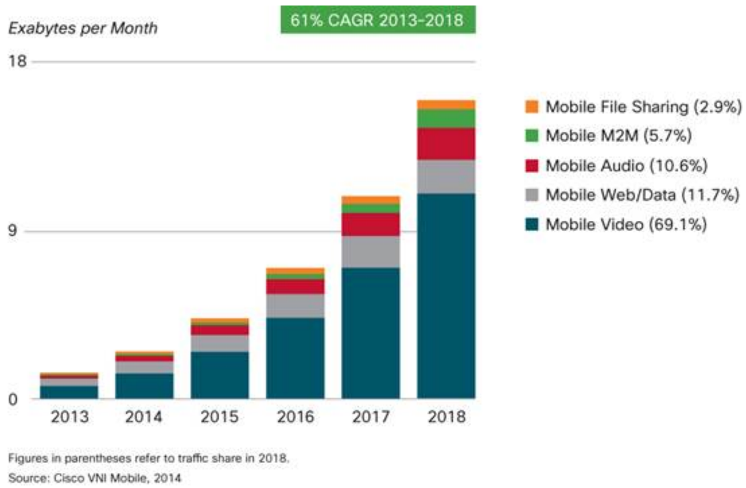

Figure 1.1: Cisco Mobile Applications Traffic Forecast [66].

---

of the Frames per second (FPS) and other factors.
The streaming video players have presented some changes to increase the quality of experience to the user. The techniques of the adaptive bit-rate allows that the quality of streaming video is changed according to the state of the network and terminal device.
The quality of user experience needs to be monitored for video streaming service to have the best quality of service which will increase the quality of user experience. The monitoring network is essential for a service provider to offer customers the best services. The provider of multimedia services has difficulties in determine the quality of experience for the user, so there are many companies that require scores when users use their services.
Meo Go is a service presented by Portugal Telecom (PT) Comunicações company that aims to provide television everywhere to customers with smart-phone, laptop and other mobile devices. This service provides a smooth streaming player from Microsoft company that has adaptative bit-rate streaming over HTTP. This service can provide the quality streaming on HD and even on FULL HD, which makes the quality of experience of this service one very important factor. PT Comunicações company aims to improve the service to have the best quality of experience and then increase the number of customers.

1.2 Purpose

The network operator has needs to monitor the quality of the experience, thus it aims to increase the quality in consumer experience. The Meo Go service presents difficulties in monitoring, because there is no module that allows to obtain the quality of experience. Therefore, PT has problems obtaining values which allows to monitor the service.

This dissertation proposes to create a system that allows monitoring the quality of experience in the Meo Go service to the PT Comunicações company. The proposed tool will allow to analyse the quality of the user experience of a service Meo Go.

The dissertation aims these goals:

- **Study the implementing Meo Go service:** to be possible to perceive the module of heuristics architecture and create the test scenarios to perceive the characteristics and metrics from the player.
- **Get quality metrics:** that are required for measurement of quality of experience from the user.
- **Implement the assessment method of the quality of streaming video:** first it implements a method to assess the chunks (2 seconds of video (default)) and then this will be adapted over time.
- **Get real values of videos quality by users:** these values will be used to compare the methods and calibrate the assessment approach.
- **Implement scenarios:** to observe the characteristics of the method over time. Create scenarios where network conditions are not ideal and determine the performance of the user's terminal.

1.3 Dissertation Outline

This dissertation contains six chapters and these are organized as follows:

- **Chapter 1:** presents de Dissertation contextualization, the motivation and the proposed work.
- **Chapter 2:** presents the state of the art of the network media services, the video concepts and quality of experience in streaming.
- **Chapter 3:** describes the structure of the smooth streaming heuristics.
- **Chapter 4:** describes the proposed architecture solution to obtain the quality of experience of the user.
- **Chapter 5:** describes the implementation of the proposed architecture; this is divided into two parts, where one is on the objective implementation method and the following refers to the subjective method.
- **Chapter 6:** depicts the results used to test the architecture of the smooth streaming, tests and calibrates the method implemented with different test scenarios.
- **Chapter 7:** summarizes the work done in this dissertation and presents the overall conclusions of the implementation and test. In the final stage it is proposed the features to be implemented subsequently in the near future.

Chapter 2

State of the Art

2.1 Overview

This chapter introduces the concepts and research work that has been developed before the preparation of this dissertation. It also describes several areas of the topic being studied. The organization of this chapter is as follows:

- **Section 2.2** presents the evolution of the types of streaming, features and advantages the architectures of these streaming technologies.
- **Section 2.3** introduces the concept of video, such as characteristics of the structure and methods to optimize resources.
- **Section 2.4** is concerned with examples of the streaming providers available in the market and the features that define the services.
- **Section 2.5** is related to the methods for the assessment of QoE by the user, and the types of reports generated in the export of the metric values.

2.2 Network Media Services

Nowadays, the streaming service is widely used. The delivery of this media (Content Delivery Techniques (CDT) [5]) uses Web services in three general delivery methods: traditional media streaming, progressive download and adaptative streaming.

2.2.1 Traditional Media Streaming

A good example of traditional media streaming is the Real Time Streaming Protocol (RTSP) [33]. The RTSP is characterized by the server control on the client during the session (Figure 2.1). The client can only start, pause and end the session, and that is why it is called a streaming state, because the client only sends its status. Once a session

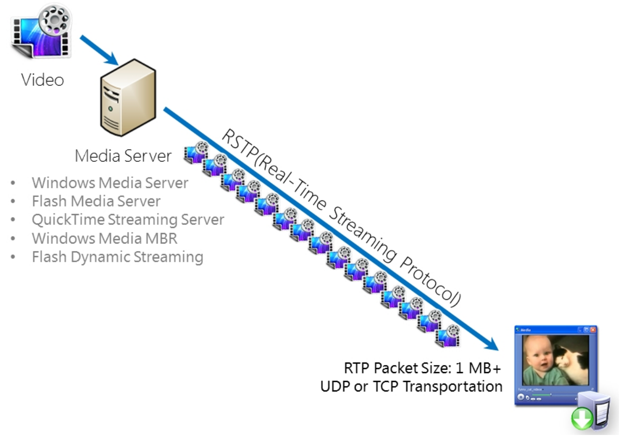

Figure 2.1: Traditional Streaming [42].

between the client and the server is established, the server starts sending the media as a steady stream of small packets (Real-time Transfer Protocol (RTP) [62]) (Figure 2.1). The video is encoded at certain bit-rates and the client will receive that quality of video content regardless the kind of CPU condition and bandwidth network. The connection is maintained persistent with each client. The size of a typical RTP packet is 1452 bytes, which means that in a video stream encoded at 2 megabits per second (Mbps), each packet carries about 22 milliseconds of video. RTSP data packets can be transmitted over UDP or TCP. The RTSP can be configured with one of the following layer transport protocols:

- Real Time Streaming Protocol over TCP (RTSPT)
  The TCP is preferred when proxies or firewalls block UDP packets. TCP packets are re-sent until received so they may increase the delay. The default port for RTSP is 554.

- Real Time Streaming Protocol over UDP (RTSPU)
  The UDP has advantages when a broadcast connection does not need a connection establishment, hence there is no delay in the case of the optimal network [10]. There is no need to maintain a connection state in the end systems, so the data of the state of the buffer and network is not needed. The packet header is only 8 bytes and in the TCP it is 20 bytes.

2.2.2 Progressive Download

Progressive download [72] is a technology that allows to transfer digital media files from a server to a client, but at the same time for the client to play the video before finishing the full download. It is mostly used as a Streaming Protocol HTTP. An HTTP connection is established and it will be transferred video files encoded with a present bit-rate. In the first phase the video file is encoded and transferred by a HTTP protocol to the client (Figure 2.2). The client can have a good network access and the state of the buffer will increase. This advantage allows the forward/backward of elapse time in the stream video and also enables to know the status of package delivery. The progressive download allows to use several codecs (such as H.264 (MP4), FlashVideo (FLV), QuickTime (QT), Microsoft Windows Media Video (WMV))(Figure 2.3) and bit rates, but only one per connection, without variability.

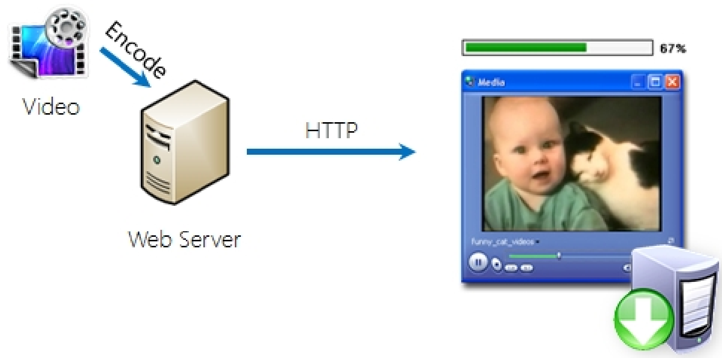

Figure 2.2: Progressive Download Architecture [42].

The progressive download can be configured with a bit-rate and codec to a defined device, thus optimizing the quality of streaming to the device, which does not happen in the traditional streaming (Figure 2.3). In the perspective of stream with excellent quality, this type of streaming is the solution, because there is no packet loss in the client, so it is possible to have excellent video quality, and therefore, large files present increased delay thereby causing network congestion.

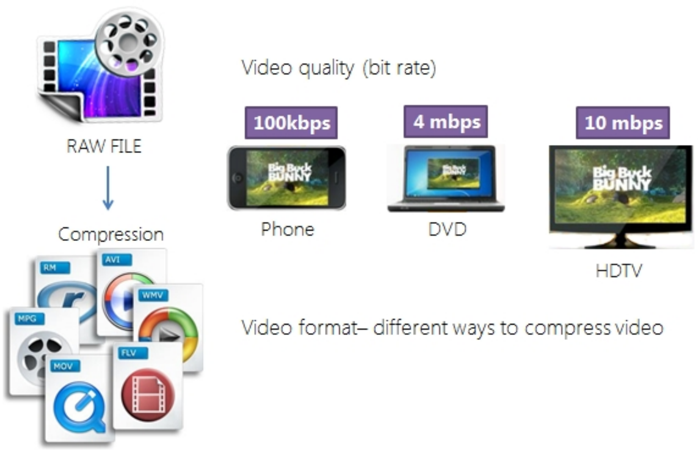

Figure 2.3: Features - Progressive Download [42].

2.2.3 Adaptive Bit-rate Streaming Technology

The adaptive streaming technology allows to adjust the quality of the video delivery that is based on network conditions and terminal device. This type of streaming is having a great success because it gives optimal results. This streaming allows all devices to receive video with the best possible quality of experience from the network and from the device at that moment. The HTTP/TCP tends to replace RTP/UDP as the main protocol for the delivery of video.

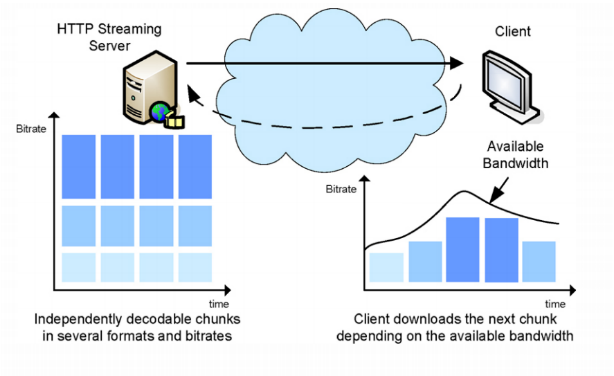

Figure 2.4: Adaptive Bit-rate Streaming Principle [32].

The Figure 2.4 presents the basic operation of adaptative streaming, in which the physical structure is composed of a server, network and client. The video file is contained in various video bit-rates, which it will transmit as a request from the customer on the server. The client will initiate the connection and then requests chunk videos with a bit-rate value.

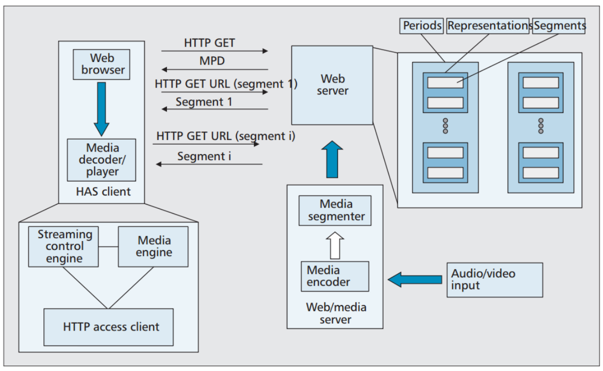

Figure 2.5: HAS framework between the client and web/media server [53].

Figure 2.5 shows a more complete structural scheme of adaptive streaming. The Client contains one access platform (Web browser) that establishes HTTP connections to the Web Server. These connections are requests of video chunks that are downloaded from the server. The engine to obtain the value of the bit rate in each request is the client, the streaming control engine. On the other hand, the server has the segments of the audio and video files in various bit-rates.

2.2.3.1 Apple HTTP Live Streaming

The HTTP Live Streaming was developed by Apple. This technique enables to send audio and video over HTTP from an ordinary web server for playback on portable devices and on desktop computers of the brand. This product supports both live broadcasts and video on demand (pre-recorded content), that can have multiple alternate streams at different bit rates, and the client software can switch streams intelligently as network bandwidth changes. Also, it provides media encryption and user authentication over HTTPS, and this enables that the client has security in the contents.

Briefly, the features of the player are:

- All devices have access to the Apple streaming video and audio.
- No special server is required.
- It allows to send a video on demand with encryption and authentication

The architecture of live streaming from Apple (Figure 2.6) on the server contains an encoder and a stream segmenter, that encode in MPEG-2 file and then separates the different chunks to deliver, and the media to be encapsulated for distribution. The distribution is responsible for accepting client requests and associated resources to the client. It also consists of standard web servers. The client software is responsible for the request of the media, stream for downloading those resource media, and after reassembling them, so that the media can be presented to the user in a continuous stream to playback.

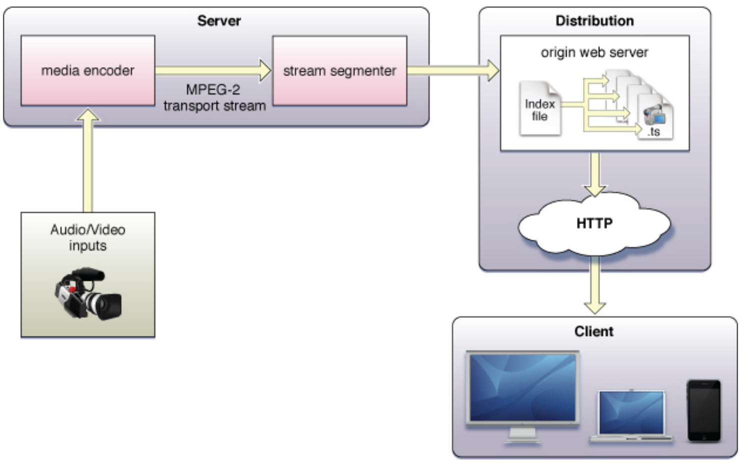

Figure 2.6: HTTP Live Streaming Architecture [18]

To use the mechanism in the server, Apple provides a free tool to make media segment files and playlists from MPEG-4 video or QuickTime movies with H.264 video compression, or audio files with Advanced Audio Coding (AAC) or Music file (MPEG Layer 3) (MP3) compression for video on demand from pre-recorded media. For live streams, it is used the MPEG-2 in the transport streams. The media streams can be encrypted with single keys randomly generated that allow automatic changes at intervals during the stream.

2.2.3.2 Adobe Dynamic Streaming

The HTTP Dynamic Streaming and Real-Time Messaging Protocol (RTMP) Dynamic Streaming [23] were developed by Adobe company[27].
Figure 2.7 shows the HTTP Dynamic Streaming architecture that contains RTMP and HTTP protocols. The RTMP Dynamic Streaming [23] is used for the live stream and it has better characteristics in delay [71]. Moreover, this stream has full support for multiple bit-rates. The content has real-time protection with Flash Access 2. It is supported by any Media encoder available on the market and does not require any special encoding.
Adobe HTTP Dynamic Streaming is also a streaming over HTTP that allows monitoring the data transmitted, so the Adobe presents this architecture as a solution for videos on demand. The Adobe allows using the mode live of this stream, in this case it has the Digital Video Recorder (DVR) feature. Also, it is supported with Flash Access 2 to encrypt data. This architecture requires one file (Manifest file in Figure 2.7) to allow the client to have the mapping of the chunks over the time and bit-rate. In the terminal part of the media on the client (Figure 2.7), this will have to use Open Source Media Framework (OSMF) to video playback. This code is open source because the Adobe company wants more video experiences to improve, extend video and media experiences to mobile devices, and improved monitoring of the content.
The Adobe software requires that the customer has the Adobe Flash Player 10.1 or later; in the part of the service it needs to run Flash Media Server 3.5 and later. In this case players are made over a JavaScript Flash language (JSFL) [27].

Some features Adobe has in the official website are the following [27]:

- Robust, scalable delivery
- Unparalleled reach
- Open source file specifications
- Adaptive bit-rate
- Support for standard HTTP caching systems
- Live or on-demand streaming support
- Client tracking and reporting
- Multiple video codecs

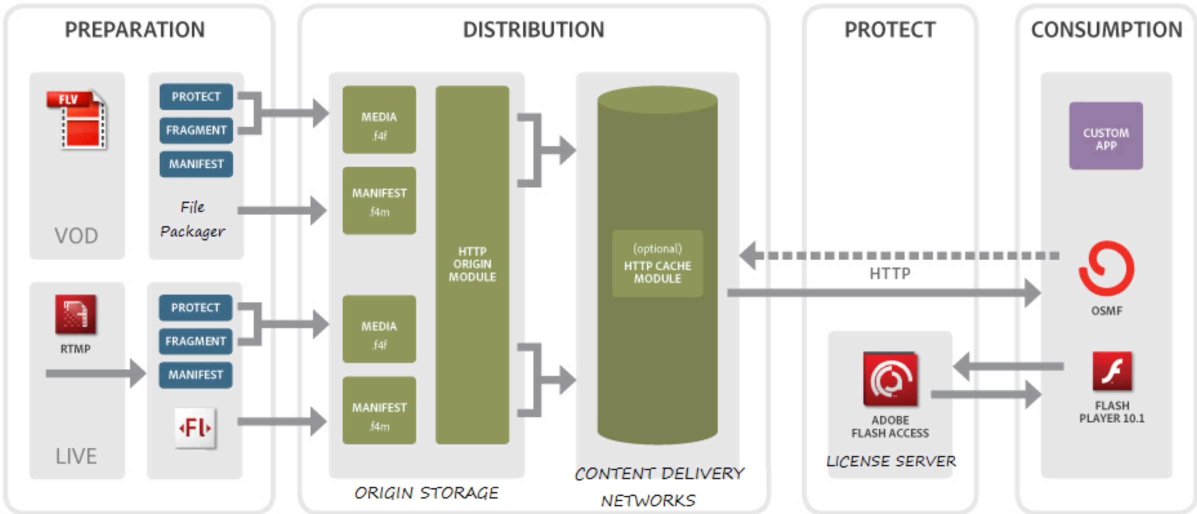

Figure 2.7: Adobe - Architecture [19].

2.2.3.3 Smooth Streaming

Smooth Streaming was developed by Microsoft, an extension for Internet Information Services (IIS) 7.0 [74]. IIS Smooth Streaming enables to detect dynamically the bandwidth and CPU conditions to switch the bit-rate, in real time. In the system it is used the HTTP protocol, the VOL and the VOD. The Silverlight is the player (Figure 2.8) that can be used by the costumer. The smooth streaming has a manifest file (Figure 2.8) that contains information about the streaming, as the values of bit-rates, number of chunks, tracks, and others. This information is required for the selection of the chunk (video, audio and text) that is requested by the client.

The heuristics module (Figure 2.8) is the main core of the Smooth Streaming Silverlight client development, and it determines when and how to switch bit rates. It follows that the main heuristics that decide the selection of the bit-rate are [74]:

- Bit-rate
- Frames per Seconds (FPS)
- Buffer Size
- Screen Size
- Load Processor

Figure 2.9 gives an example to demonstrate that, when the streaming starts, the silverlight player bit-rate has the minimum value, but the value adapts itself while the time elapses. There is a minimum period that allows to modify the bit-rate value as it shows the sample of Figure 2.9.

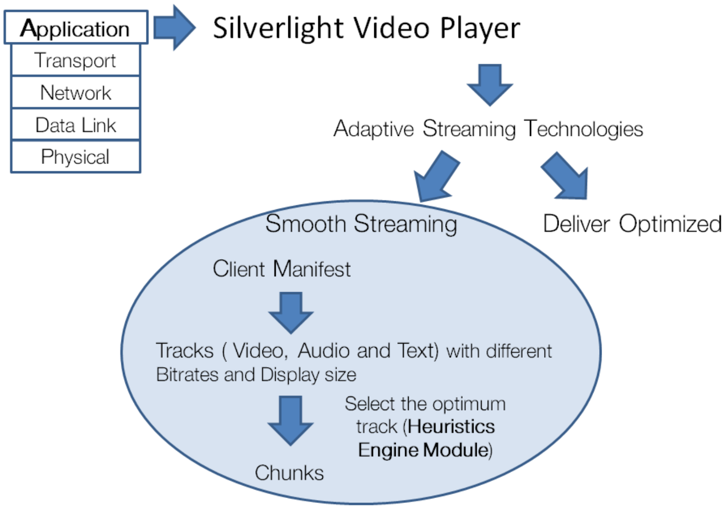

Figure 2.8: Smooth Streaming Architecture.

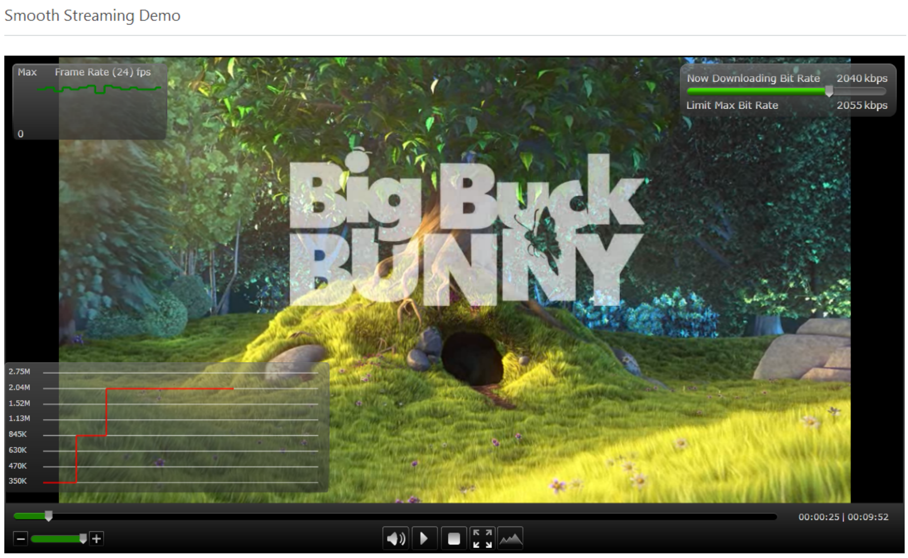

Figure 2.9: Demo - Player Silverlight [46].

The smooth streaming is a technique used in several services and it optimizes the quality of experience from the user. The Figure 2.10 compares the smooth streaming with traditional streaming service that was referred to in subsection 2.2.1. Smooth Streaming has the ability to adapt to the heuristic metric. However, in traditional streaming, bit-rates have to be set without knowing the conditions of the network and of the terminal device; these pre-defined bit-rate must be low to allow to have a larger number of customers.

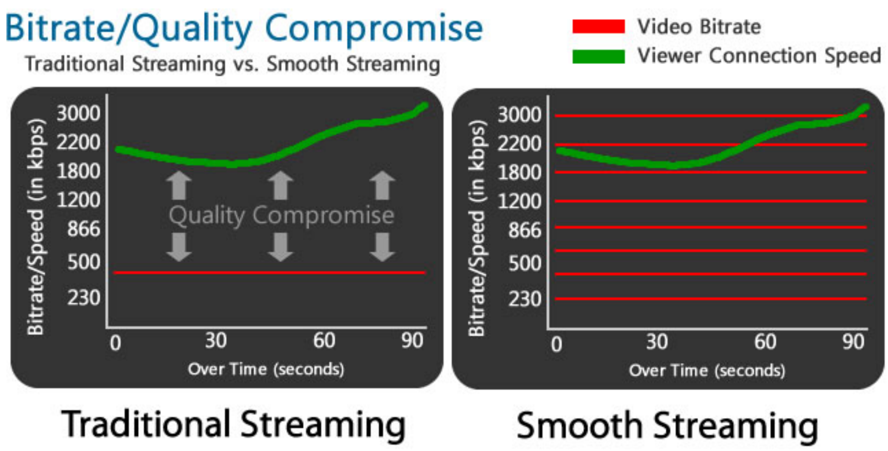

Figure 2.10: Bitrate/Quality Compromise[9].

2.2.3.4 DASH

The Dynamic Adaptative Streaming over HTTP (DASH) is an international standard that should enable interoperability among proprietary solutions by the ISO/IEC MPEG [34]. The concept of DASH is a technique for adaptive bit-rate streaming. The Institute of Information Technology (ITEC) and the Multimedia Communication Research Group (MCRG) of the Alpen-Adria-Universitt Klagenfurt has participated and contributed from the beginning to this standard [48]. It is the standard for adaptive bit-rate streaming that appeared to standardize existing solutions, e.g. the streaming adaptive by Apple, by Microsoft and by Adobe [34].

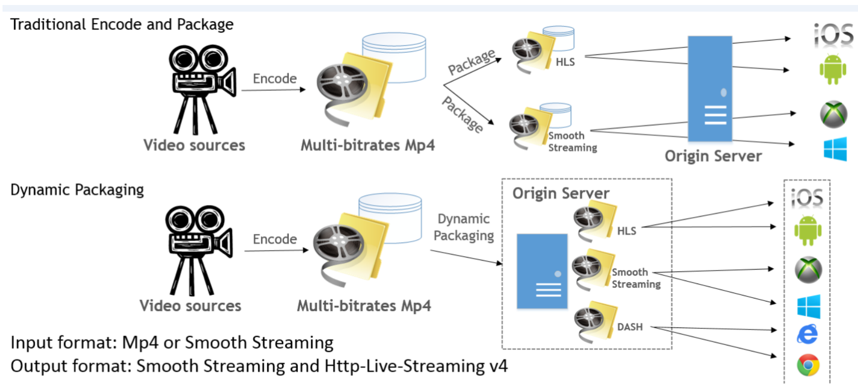

Figure 2.11: DASH - Context [42].

Actually, this technique has no new features, it just has greater portability for the normalization of streams. The figure 2.11 depicts the portability functionalities.

Figure 2.12 describes the 3GP-DASH solution to adapt to the standard specifications [2]:

- **Normative definition of a Media Presentation**, a structured collection of data that is accessible to the DASH Client (Figure 2.12).
- **Normative definition of the formats of a segment**, integral data unit of a media presentation.
- **Normative definition of the delivery protocol**, namely HTTP/1.1 (Figure 2.12).

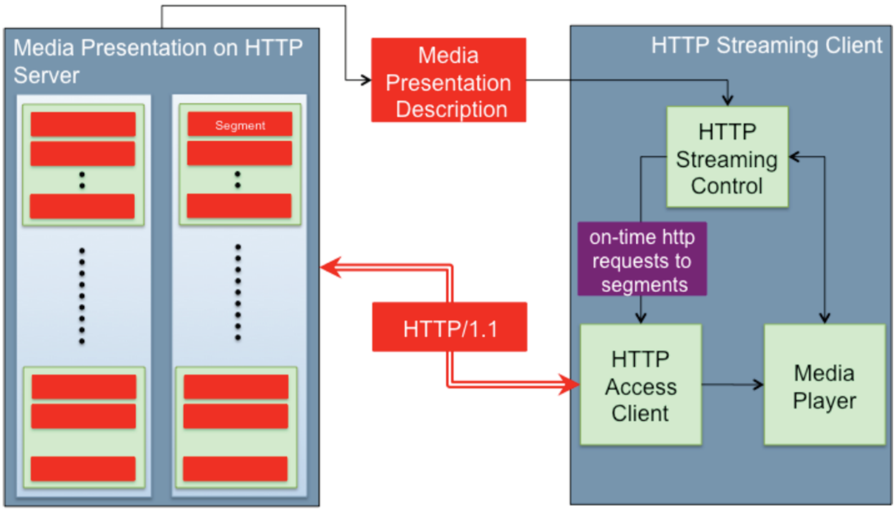

Figure 2.12: Solution overview 3GP-DASH [20].

MPEG-DASH consists of regulations that allow the use of adaptive streaming without the need to know the framework (Figure 2.11).

DASH in 3GPP is defined in two levels:

- **Clause 8.4.4.1 in TS 26.247** [4] provides support only for the media-streaming model and for the control for delivery of media content.
- **Clause 8.2.1 in TS 26.247** [4] provides that the Media Presentation Description is an XML-document defined in Clause 8.2.2.

2.3 Video concepts

This section introduces relevant aspects of video sampling and coding. The video is composed of sequences of images at a present rate; the images are frames of pixels. The colors of the pixels are represented in bits, then there is a reference that represents the size of the pixels in the image over time, the bit-rate. In the following, video optimization techniques can be used in the video:

- Sampling
- Coding

2.3.1 Sampling

A video can be seen as a set of images within the same space over time. In the same way, a digital video is represented by a sequence of images obtained by sampling, where this is made in two distinct spatial and temporal domains. The spatial sampling is the process by which an image is produced. Therefore the temporal sampling is the process that produces a sequence of images that create the illusion of movement. The number of frames per second influence video quality between 10-20 frames with the flicker phenomenon, but disappears after 25 frames. This is due to the refresh rate of images in the human eye, i.e., when two images are successively presented to the eye in a short period of time that retains the image.

Color image

Images can be represented in the color image. There are several color images, but the most used are Red, Green and Blue (RGB) and YCbCr (or YUV (Where Y is the gray-luminance, U is the blue-luminance and V is the red-luminance differences)). Human vision follows the principle of superposition, so it is possible to mix the primary spectrum (RGB). Typically YUV signals are created from RGB source. There are several methods for YUV-to-RGB conversion transformation [56].

2.3.2 Coding

The video is presented with a set of images that change at the rate of frames per second, increasing its size over time. The use of coding techniques allows to reduce the video information: by redundancy and irrelevancy. The irrelevancy explores the limitations of the human visual system, eliminating parts of the video signal that human eyes can not recognize. The redundancy explores the fact that the redundant information in video signal can be removed, and it is possible to reconstruct the signal without loss of information or introduction of distortions. When removing this type of information to the video signal, the eye will perceive information from the compressed signal as if it was the original signal, but the signal can not be reconstructed just like the original signal that caused it, so there is therefore loss of information.

The following are some video encoders in the market:

- MPEG4 AVC (H.264) [69]. Presented by the ITU-T Recommendation H.264 and by ISO / IEC International Standard 14496-10, the (MPEG-4 part 10) Advanced Video Coding (AVC) is currently the most effective video encoder. There is also MPEG-2, which is used in some cases, such as DVB, RTMP streaming by adaptative Adobe, but MPEG-4 is its successor. The feature developments in H.264 codec over MPEG-2 are described in the article [69]. The main features are:
  - Motion compensation, the size and shape of blocks can vary.
  - Reduction of spatial redundancy uses an entire transform that reduces the influence of errors.

  - Quantification has a larger number of levels, the MPEG-4 has 52 levels and the MPEG-2 has 31 levels.
  - Entropy encoding uses a more complex coding, which is more efficient compared with the static code included in the MPEG-2.
  - De-blocking filter uses an adaptative filter for reducing block effect of the image, which degrades in the MPEG-2.

- Windows Media 9 / VC-1 [8]. Compression standard video / audio normalized by the Society of Motion Picture and Television Engineers (SMPTE) [49] and implemented by Microsoft, WMV 9 [40]. This standard is an alternative to H.264 and is also an improvement compared to MPEG-2. The basic functionality of the VC-1 involves block-based compensation, and a scheme similar to the spatial transformation that is used in other video compressions, such as MPEG-2 or H.264 movements. However, the VC-1 includes a number of innovations and improvements as follows [68]:
  - Adaptive Block Size Transform
  - 16-Bit Transforms
  - Motion Compensation
  - Loop Filtering
  - Interlace Coding
  - Advanced Coding Frame B
  - Fading Compensation
  - Differential Quantization

  The VC-1 presents profiles and levels that support the encoding of many types of video. The profile has the codec and the resources that are available to allow optimizing types of services. Figure 2.13 lists VC-1 profiles and levels.

2.4 Streaming Services

Television services are increasingly being present at residential population through services with cable, satellite, Internet Protocol Television (IPTV), and others. Streaming media has been experiencing an increase of users due to the portable equipments. Providers of streaming services are investing heavily on the market and there are already dedicated services.
Two examples of multimedia streaming providers are presented below:

Netflix is an American provider of on-demand Internet streaming media. It is presented as one of the big providers of streaming media services and has a large range of hardware compatible with the system. This service can be used in a compatible television,

Profile	Level	Max Bit Rate	Representative Resolutions by Frame Rate
Simple	Low	96 Kbps	176 × 144 @ 15 Hz (QCIF)
	Medium	384 Kbps	240 × 176 @ 30 Hz 352 × 288 @ 15 Hz (CIF)
Main	Low	2 Mbps	320 × 240 @ 24 Hz (QVGA)
	Medium	10 Mbps	720 × 480 @ 30 Hz (480p) 720 × 576 @ 25 Hz (576p)
	High	20 Mbps	1920 × 1080 @ 30 Hz (1080p)
Advanced	L0	2 Mbps	352 × 288 @ 30 Hz (CIF)
	L1	10 Mbps	720 × 480 @ 30 Hz (NTSC-SD) 720 × 576 @ 25 Hz (PAL-SD)
	L2	20 Mbps	720 × 480 @ 60 Hz (480p) 1280 × 720 @ 30 Hz (720p)
	L3	45 Mbps	1920 × 1080 @ 24 Hz (1080p) 1920 × 1080 @ 30 Hz (1080i) 1280 × 720 @ 60 Hz (720p)
	L4	135 Mbps	1920 × 1080 @ 60 Hz (1080p) 2048 × 1536 @ 24 Hz

Figure 2.13: Profiles and Levels (VC-1) [68].

mobile phones and tablets, but it is not present in all countries, only those who have an agreement with providers of internet services [45].

Meo Go is provided by the company Portugal Telecom. Meo Go [11] is the service that allows the Meo TV Service features, offers 60 channels of live television and Video-Clube, inside and outside the home, on the tablet, smartphone and PC.

The service is available for all MEO TV Service Clients and presents strategies to only use the traffic on the networks of PT Comunicações:

- Free to access inside home.
- Upon subscription of Meo Go Multi (data traffic included in the PT Comunicações network) to access everywhere.
- Available only in Portugal.

The referenced Figure 2.14 in official site provides some features, multi-platforms, as shown below:

- Watch Live TV, including Premium Channels;
- Watch movies from MEO VideoClube
- Access to the Electronic Program Guide (EPG)
- Home DVR management
- Allows control of TV via a digital keypad or through gestures.

- Available for any PC/Mac, and the most significant mobile operating systems, Android, iOS (iPhone and iPad), Windows Phone and Windows 8 (Figure 2.14).
- Any network of any operator: available in WiFi and 3G/4G - Free Traffic to Customers in PT .

Figure 2.14: Meo Go application on smart phone (iOS system) [11].

The smooth streaming adaptive (Section 2.2.3.3) is the technology of the MEO service. PT communications has servers configured with smooth streaming, and in relation to customers, Silverlight [41] is required.

2.5 Quality of Experience

"The degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the users personality and current state"[50], QoE definition by Qualinet

With the increase of the media streaming applications, it is crucial to monitor the quality of audio / video services. Methodologies to assess the quality of services are essential when it is required to provide a high quality service. The models are appropriate to the monitoring of streaming media when these are not intrusive (example, the parametric audio packet-layer model [39]). It is not feasible for service providers to measure the network qualities intrusively, because complexity increases as it is necessary to compare the original video with the received one and to know the experience of user.

In the assessment of video quality, we must consider the characteristics of human visual system to obtain a more accurate quality estimation. The presence of the concealment effect of another stimulus can be invisible, as it is an important characteristic for the human visual system as the spatial and temporal concealment. There are several articles that report studies on the evaluation of the quality of experience, but only analyses a few heuristics, rarely more than two. For example, article [31] only covers the coding distortion and degradation of packet loss.

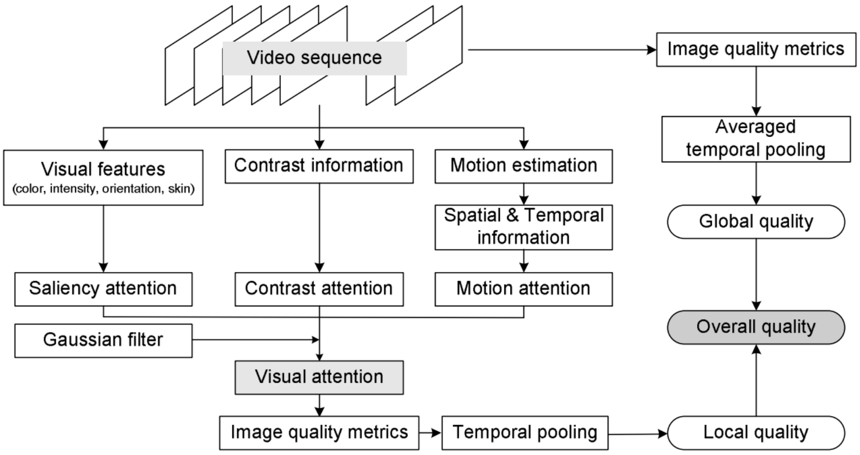

Figure 2.15: Flowchart of the proposed video quality model in the article [73].

The complexity of these methods increases with the improvement of the results. The work in [73] presents a figure (Figure 2.15) with the proposed structure with two types of quality measurement, quality Local and Global. In this work it is assumed that the global quality is a result from distortion over all frames, in sequence, and the local quality can be affected by the distortion into the frame. This structure references only the distortion of the video. Figure 2.16 presents an example of the debug stream on the sport stream, which aims at analysing the different distortions of the video to approach a subjective method. The relevant distortions have different negative contributions in quality of experience from the user.

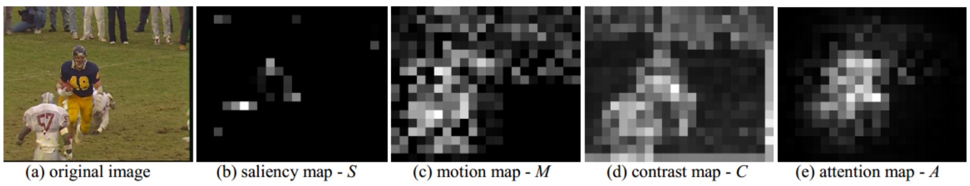

Figure 2.16: Frame image in Football sequence and attention maps [73].

2.5.1 QoE Reporting

The development of methodologies for assessing QoE allows to obtain performance metrics of network and terminal device. The service provider requires the adaptive streaming QoE monitoring to improve their services. This monitor is intended to be incorporated in the company system, so that the processing is done in real time or afterwards. The standardized reporting of QoE [1] is the solution presented on exporting values, and the adaptation for a DASH (section 2.2.3.4) stream recommends using a standard file [1]. This export requires standardized data that are obtained in the metrics, described in Figure 2.17. 3GPP recommends performing an export of metrics data in XML.

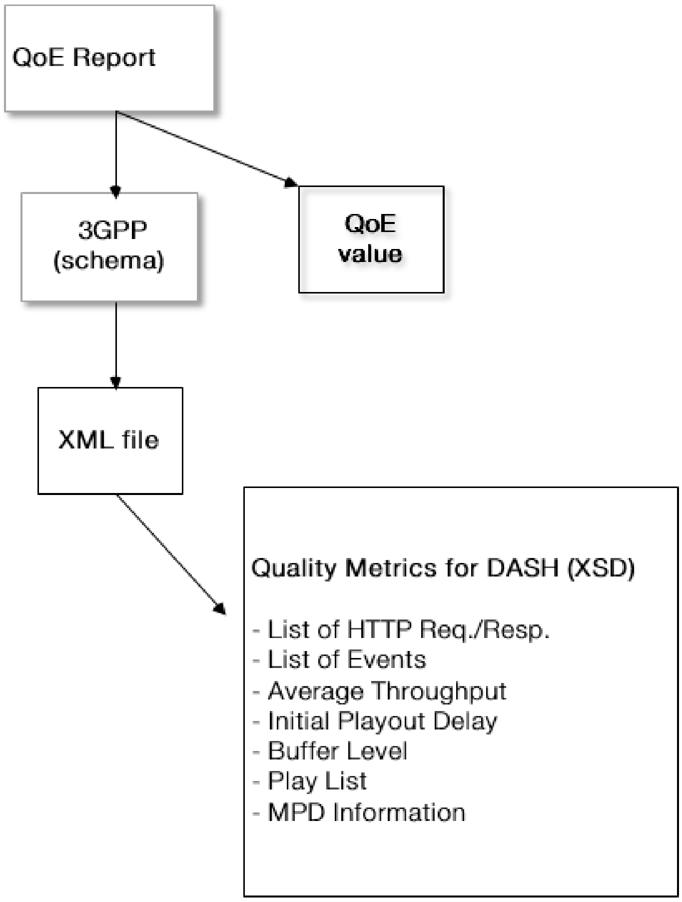

Figure 2.17: QoE Reporting [1].

2.5.2 Assessment methods for video quality

In adaptive video streaming, quality of user experience is assessed using subjective and objective methods. Subjective methods [44] have no specific characteristics to analyse, they are only based in humans with assessment of a streaming video, containing a high number of assessments. The objective method requires metric values to classify the streaming video, but it is necessary to calibrate the metrics with the subjective method, usually there are databases already trained, as an example is Perceptual Evaluation of Video Quality (PEVq) [52]. The video streaming systems are complex because they can depend on

many of the factors unknown, such as codecs, screen type (Light Emitting Diode (LED), Liquid Crystal Display (LCD)), and others. Human visual perception is very complex. There is a need to understand the human brain structure to simulate its experience, such as human memory. This type of evaluation is crucial for the service provider, and taking only the value of the metrics is not enough.
Example of a case: A low bit-rate for a client of a laptop with full HD (19 inch) screen displays a very low quality of experience, but this bit-rate for a client of a smart-phone with 3.5-inch screen has the medium or high quality of experience.
Figure 2.18 presents the structure of adaptive assessment types of streaming. It is noted that there is always connection between the two methods. As indicated in the in Figure 2.18, the subjective method cannot replicate without the use of the objective method.

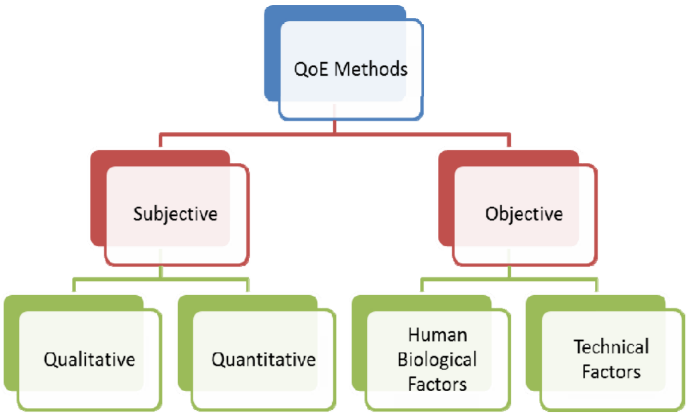

Figure 2.18: QoE Assessment Methods [65].

2.5.2.1 Subjective methods

Subjective assessment methods are commonly based on surveys, interviews, and statistical sampling of users to analyse their perceptions from the service [31]. There are two techniques for subjective studies (Figure 2.18):

- Qualitative Techniques
  Data are presented as verbal words and behaviour. There are no results represented in numbers, but there are opinions, comments and questions. This technique can be advantageous in the so-called ratio of positive to negative comments (CCA - catalog, categorize, analyse)[61]; for example when the bit-rate is low and the display window is small (pixels not recognized), if the user does not like to see on the small screen, the quality result from the user can be of low quality, caused by the one different user, people generally give a better value to quality of experience.

- Quantitative Techniques
  Data are presented as numbers and statistics. These studies can be performed in

a laboratory environment to measure human feelings and perceptions. Tipically a questionnaire is applied with ratings on scales (Figure 2.19) for the data to be processed. To create scenarios with standard values, there is a recommendation ITU-T P.1202 (10/2012) (parametric, non-intrusive assessment bitstream of video media streaming quality) [24] by Motion Pictures Expert Group (MPEG(ITU)).

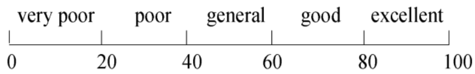

Figure 2.19: One-hundred mark scale [31].

2.5.2.2 Objective methods

Objective evaluation of video is based on mathematical models that estimate results of subjective evaluation of the quality rankings that are based on metrics that can be measured objectively and can be automatically evaluated by a computer program. There are two branches of objective assessment as follows:

- QoS/Technology centric techniques [31]
  In this particular case of the objective method it is used tools, mathematics and techniques to simulate the value of experience from a user. It is usually analysed the image quality or video, and these methods can be Full Reference (FR) or Reduced Reference (RR). In FR it is required the original file to be compared, but in the RR it is only required characteristics that are presented on the terminal of the client. There are several ways to calculate the degradation of video with FR, as indicated below:
  - Peak Signal-to-Noise ratio (PSNR) is a FR method that obtains the ratio of noise between two images, transmitted and received [47].
  - Structural SIMilarity (SSIM) is a FR method that looks at differences by using the YUV color space (Section 2.3) [67].
  - Video Quality Metric (VQM) [51] is a standard developed by the Institute for Telecommunications Sciences, National Telecommunication and Information Administration (TIS/NTIA), it allows to be a FR and it is standardized by the American National Standards Institute and it has been adopted in two ITU Recommendations, namely ITU-T J.144 and ITU-R BT.1683 [59].
- Physiological and cognitive techniques (Figure 2.18)
  This sub-method allows the user to be examined neurologically and cognitively, thus it analyses the behavior of humans in relation to QoE. Tests such as electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging(fMRI), and near-infrared spectroscopy (NIRS), examine the state of

the user (sensors) and collect values for measuring QoE, that is the most expensive method [6].

However, these do not give results as reliable as the methods of subjective evaluation. In fact, the great advantage of the objective methods is the ease of monitoring the signal of interest. Therefore, there is a large investment in this area in order to make the results more reliable as possible by these methods.
Operator demands metrics streaming evaluation to improve service at delivery to consumer. The quality of user experience is influenced by metrics that may not depend on network and subsequently the operator does not have direct access. This dissertation aims to create one solution to improve the operator network management. QoE was studied but only upon individual metrics (two at most), so we decided to use several metrics influenced by human and network factors in order to achieve a more accurate QoE. This dissertation will be organized from objective and subjective assessment to minimize errors. This section defines methods that are considered in the work; however, there needs to be a complement, i.e., to calibrate the method and to simulate. This can be considered as an algorithm which allows to simulate a value of quality of experience for the user.

Summary

This chapter describes the existing technology and the requirements in the proposed solution for quality of experience. The technology over the player (Section 2.2) is quite crucial to obtain all possible metrics, and selected only those that will contribute to analyse the quality. The characteristics of a video (Section 2.3) have to be considered in the method, the main features that the user will assess, for example the bit-rate, FPS, and other video features. As the next, step mention is made to existing streaming services (section 2.4) on the market, so in this dissertation it is intended to develop a method for the Meo Go; in this implementation there is the need to create a report QoE. This can follow the standards to which reference is made in Section 2.5. Finally, it will be necessary to select the architecture of the assessment methods for the purpose of this dissertation, as the service provider (PT) Meo Go aims to insert a final algorithm solution in the system.

Chapter 3

Smooth Streaming Architecture

3.1 Introduction

The proposed solution to quantify the quality of experience requires the metrics informations and we can get them through the study of the player structure. This structure is essential to the customer to receive the adaptative streaming service. The adaptative streaming is characterized by an engine which optimizes the quality of experience. This Chapter presents the player structure and it is splitted into the following sections:

- **Section 3.3** will present the overview of the player structure, operation mode and format files.
- **Section 3.4** will present the communication layer between a client and a media server and the opposite, and the structure of the connection.
- **Section 3.5** will describe the module heuristics present in the smooth streaming architecture, structure with engine selections metrics.

3.2 Overview

The proposed solution will be implemented in the scenario of the operator, Figure 3.1 shows this integration that will be inserted in an intrusive probe to obtain the quality of experience values in a particular network, such as the export of network metrics that allows to obtain a quality of service.

Smooth Streaming dynamically detects the conditions of bandwidth and performance of the terminal device and seamlessly switches, in real time, the video quality of a media file that a client receives. However, this quality will be emulated in a module, that allows to combine the values of all metrics in one value, including the user experience.

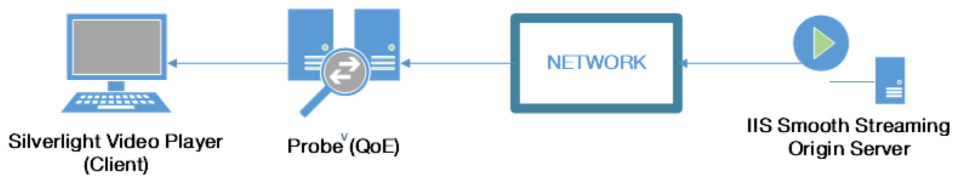

Figure 3.1: System structure.

3.3 Presentation structure

Stream structure is comprised by several components: web media servers, HTTP connections, several applications for terminals (laptop, mobile phones, ...) (Figure 3.1). The figure 3.2 represents an overview of the architecture: there is a Silverlight (Application) player that is on the HTTP layer. This presents an adaptive technique to stream the video frames, which is denoted by smooth streaming.

The technology involved is based in multiple rates (Figure 3.3) that can be sent between media server and client terminal. In the server, there is a video with high resolution (High Definition (HD), Full HD, Ultra High Definition Television (UHDTV)), that will be later processed by an encoder with multiple bit-rates which will create multiple tracks with different bit-rates (Figure 3.3). On the client, the player decides heuristically the optimal quality for the connection stream. In this connection there are consecutive chunks of video requests that may be of different qualities. The qualities are defined by track where each one contains a list of chunks. This list size depends on the time of streaming and if the chunks present standard packages; then, usually it contains 2 seconds of video.

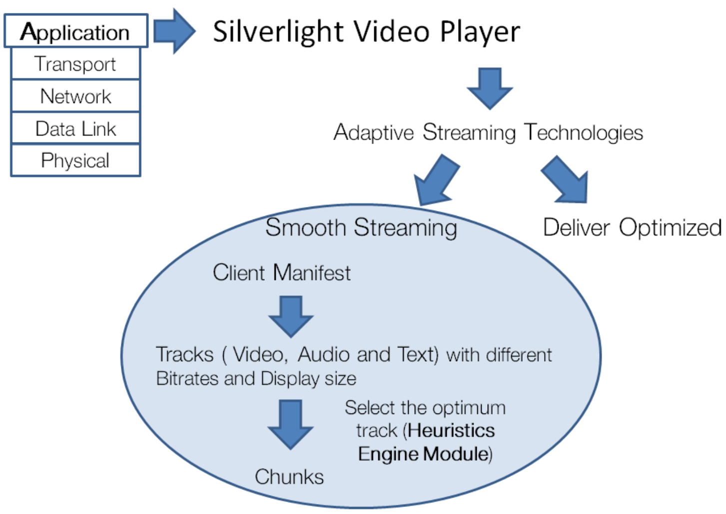

Figure 3.2: Overview - Silverlight Video Player.

When a customer wishes to play a streaming video, the player requests a chunk of a track using an interface file (Manifest, explained in section 3.4). This innovative architecture allows to receive one track of stream that depends on the network performance and the capacity of the terminal device (client). When the client initiates a request, it is necessary to mention the desired quality that is decided by heuristics engine on the client player. Figure 3.2 presents the streaming scenario where the player decides the chunk per track which depends on specific characteristics, such as screen size and bit-rate, subsequently it is requested the chunk.

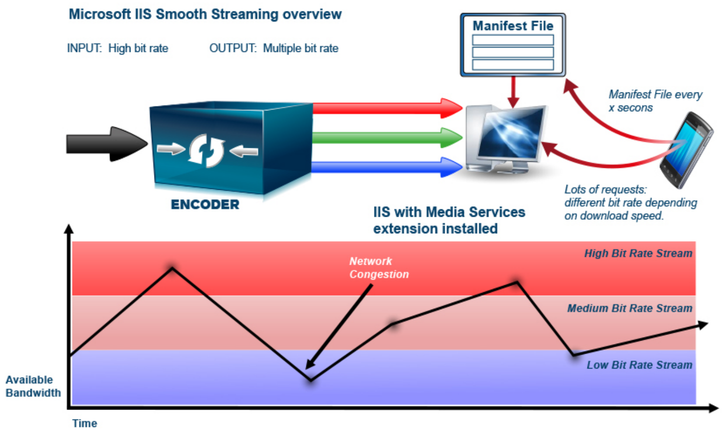

Figure 3.3: Smooth Streaming Architecture [74].

3.3.1 Bit-rate selection

The server stores information of different quality in tracks, for example, the first track contains the lowest bit-rate and the second contains the medium bit-rate. Each track is constituted by chunks (2 seconds of video) of time-dependent streaming session (mapping of chunks). Smooth Streaming allows to map the quality as a function of time; as seen in the Figure 3.3, there is an axis with bit-rates and another with the elapsed time.

The player selection, based on the heuristics module, will select the qualities that provide a better service experience over time. The chunk selection is done by mapping a matrix (tracks over time - Figure 3.4), where the change of qualities of video stream playback is never stopped. This is the innovation that makes the streaming smooth. The diagram in Figure 3.4 shows the mapping of blocks based on the time elapsed.

This mapping is done at the level of the player; the heuristics engine indicates the desired track that is playing at the moment and the time-elapsed in playback; thus, it requires the following chunk and may have different bit-rates. For example, based on the figure, the first player starts the play with a track of 300 Kbps (Bit-rate); thereafter, the network increases its performance after 4 seconds, and the player decides to change to a track of 1 Mbps (Bit-rate), without interruptions, and only now it changes the playback track.

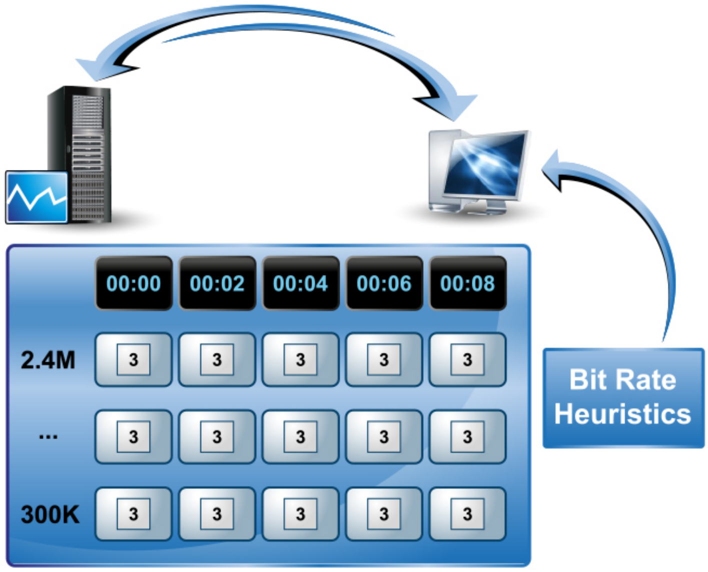

Figure 3.4: Mapping selection bit-rate [74].

3.3.2 Disk file format

The server stores the original file and the generated tracks. The tracks have chunks with 2 seconds (default), designated by ISMV files. ISMV files are saved in the Protected Interoperable File Format (PIFF), which is the Microsoft's standard for multimedia content delivery (based in the MPEG-4 specification [3]), and may be encoded using the Microsoft Expression Encoder. This type of standard platforms within the Microsoft offers large protection and optimizes performance. In addition, the format allows for encryption and DRM protection [36], when the service is authenticated by the client. A streaming session is characterized by having the transmission of audio (Radios, ..) or audio + video (TV, ..): each chunk contains only audio or video stream; when there is one connection with audio + video, the player will download it of each type of chunk. The following are the codecs that can be used:

- Video + Audio (*.ismv extension): The encoding can be used with the VC-1 codec or H.264; these profiles are based in the compression methods. In the case of H.264 (MPEG-4 Part 10), it presents a profile (Extended Profile (XP)) for streaming media that achieves high compression rates; in the case of VC-1, it has profiles which are defined by bit-rates and screen sizes. Meo Go service is based on H.264 (MPEG-4 Part 10) which uses the profile XP optimized stream.

- Only Audio(*.isma extension): In this case the codec only needs to encode audio. WMA Pro 10 was created by Microsoft for their applications and third parties, and the case of AAC-LC that presents best compression than MP3 traditional, which is also used in the integration of the audio in MPEG-4.

3.3.3 Chunk format

The chunks are derived from lists of tracks that belong to the contents of the lists, and these are identified with reference to the corresponding track and the elapsed time. Figure 3.5 refers to the fragment format; the structure is divided into two parts, metadata (*moff*) and media data (*mdat*). The *moff* contains the stream identification, the track identification and elapsed time identification, in the header. The *mdat* contains the encoded file of the video to playback. The size of the chunks varies with the desired bit-rate and, when it is decoded, the duration always presents a normalized time (2 seconds). The chunks are finally transported one by one and can be played regardless, which means that it is not needed to wait between chunks to load completely for playback. This fragment belongs to the list of tracks on the server, and when it reaches the customer, it is inserted into a play queue (Buffer Engine).

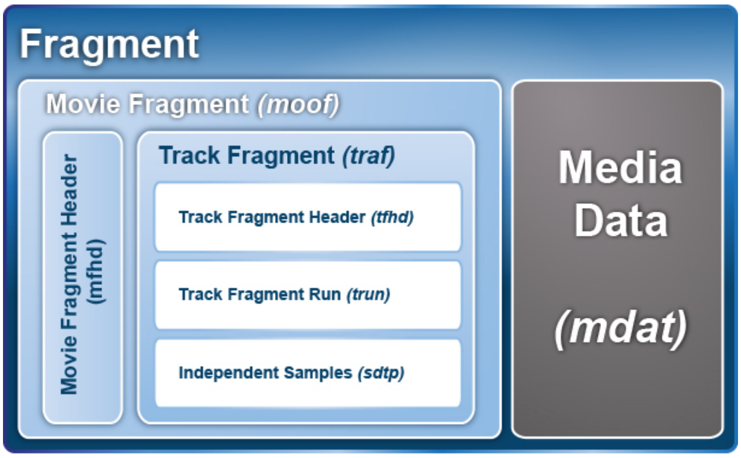

Figure 3.5: Smooth Streaming Wire Format [74].

3.4 Server and client manifests

The manifest contains the information file to export from the server to the client; this is the first file downloaded by the player (client), subsequently it saves the information in order to initiate and continue playing the video stream. This initialization file playback stream is termed as element for presentation and it has the following information:

- Metadata for each track (video, audio, and text);
- Describes the possible choices for quality level in the video/audio tracks;
- Describes the different fragments (c (chunk) element)) of this track;
- Metadata for each track fragment;

Briefly, the client (server) loads the manifest and then it selects the chunk based at information acquired in the manifest, the quality (Bit-rate and windows size) and elapsed time. Figure 3.6 presents an example of the structure of the XML Manifest created with Expression Encoder (Microsoft). Initially, there is an identifier to the elapsed time of the stream session which matches at live streaming channel, in VOD case. VOL corresponds to the elapsed time with the oldest content channel. The existing ID in the manifest allows to know the current chunk (VOD) or older (VOL). Depending on the service provided, example, audio and video + audio, there is an index for each type of stream with the information on the qualities of existing tracks: in this case a track with a bit-rate of 2.436Mbps with a resolution of 1280x720 pixels and another track with one bit-rate of 1.636Mbps with a resolution of 960x544 pixels; there is also the index related with stream chunks information, duration and offset number. The specification of the streaming Wire / Soft File format defines the manifest XML language; XML has advantages in the file extension.

```xml
<!-- Created with Expression Encoder version 2.1.1216.0 -->

<SmoothStreamingMedia  Duration="5964580000">
        <StreamIndex Type="video" Subtype="WVC1" Chunks="299">
                <QualityLevel Bitrate="2436000" FourCC="WVC1"
                        Width="1280" Height="720"/>
                <QualityLevel Bitrate="1636000" FourCC="WVC1"
                        Width="960" Height="544"/>
                ...
                <c n="0" d="20000000"/>
                <c n="1" d="20000000"/>
                ...
        </StreamIndex>
        <StreamIndex Type="audio" Subtype="WmaPro" Chunks="299">
                <QualityLevel Bitrate="64000"/>
                ...
                <c n="0" d="21362358"/>
                ...
        </StreamIndex>
</SmoothStreamingMedia>
```

Figure 3.6: Sample - Client Manifest (XML)

After the manifest is obtained, it starts the engine download of chunks. HTTP request follows the structure:

`http://{StreamURL}/{PublishingPointName}.isml/QualityLevels`
`({BitrateValue})/Fragments({Video_or_Audio_or_Text}={Timestamp})`

For example:

`http://video3.smoothhd.edgesuite.net/Big%20Buck%20Bunny%20Adaptive.ism/`
`QualityLevels(2436000)/Fragments(video=5924583334)`

In this sample, there is a request of the chunk (*Big Buck Bunny* video) with 2 seconds (*d="20000000"*), started at the fourth second (5964580000-5924580000 = 40000000 (2 seconds)), with 2.436Mbps of bit-rate, that is hosted in `video3.smoothhd.edgesuite.net` server.

3.5 Heuristics

3.5.1 Heuristics Engine

This module aims at the selection of the qualities present in the Manifest file. Figure 3.7 shows the structure of the heuristics module operation. The sub-module allows to get eligible tracks of different heuristics, and then each heuristic filters the list of tracks according to their requirements. In the final phase, it is suggested to track the chunk. This influences only the combined network module characterized by the network and the buffer. The selection is direct when the suggested list has only one element; if it does not have tracks in the list, it selects the lower quality; if the list has more than two elements, it is selected by the network and the buffer (combined network/buffer).

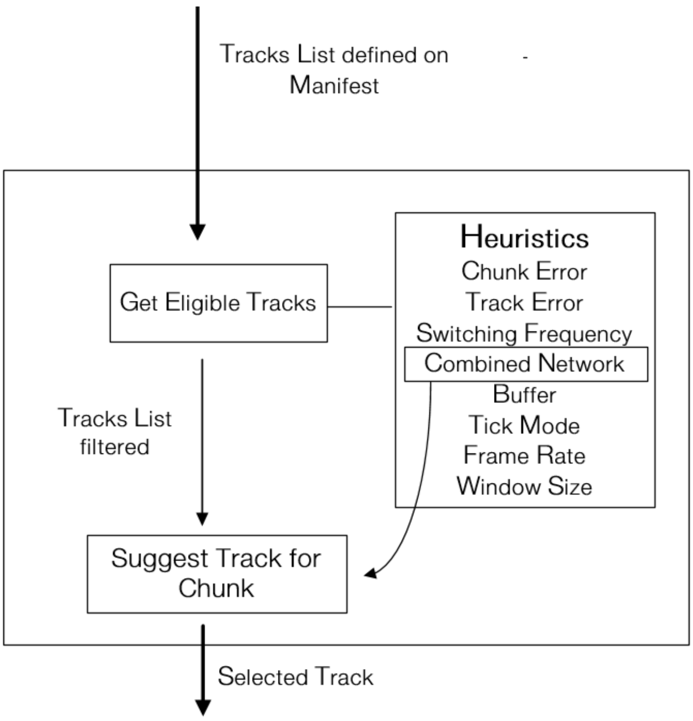

Figure 3.7: Heuristics Engine.

As an example case, consider a playback of a stream track in the mobile phone with a small screen. The heuristics engine suggests a track with small resolution (lower quality), because the list is filtered with only one track by Windows Size Heuristic.

3.5.1.1 Chunk Error

The heuristic aims to validate if the selected chunk is correct in relation to the elapsed time at playback, because it may occur one old chunk.

3.5.1.2 Track Error

The network can always see its performance changed during playback of the video stream; moreover, tracks quality also change. Thus, the tracks are suspended when a number of attempts are unsuccessful to reproduce. This heuristic filters the tracks list that are suspended; thus, video quality does not constantly changes in the streaming session, as it allows increasing the quality of user experience.

3.5.1.3 Switching Frequency

In a network there are several peaks of performance and saturation, which affect the quality of the video stream, this heuristic will confirm that the change is accomplished in the specific interval. This interval is different when it increases or decreases the quality of the stream. Switching Frequency is very advantageous to increase QoE, because a network can provide good conditions that could increase the bit-rate quality for only a few seconds. For example, the video stream can only change the bit-rate between 4 in 4 seconds. In an unstable network, it allows the user to not observe instability in the video, and it is similar at the effect of a low-pass filter quality on time.

3.5.1.4 Combined Network Buffer

This heuristic aims to relate the metrics buffer and network. These two factors contribute significantly to get eligible on track and the final suggestion of the track. This heuristic aims to verify the conditions of the network and the buffer combined. For example, if the network has a high performance, it can overrun the buffer and exceed the limit of capacity, so a complex management is required for the two factors in the heuristics. The following is the description of the network and the buffer.

Figure 3.8 shows the combined architecture of the two heuristics, which is crucial to suggest the track for chunk. First, it acquires the lists of suggested tracks from the network and buffer heuristics; then it presents the cases that may occur:

- **High network performance** - When the network module provides the full track list (all possible qualities), it is suggested the buffer's tracks list.

- **High buffer performance** - When the buffer module provides the full track list (all possible qualities), it is suggested the network's tracks list.

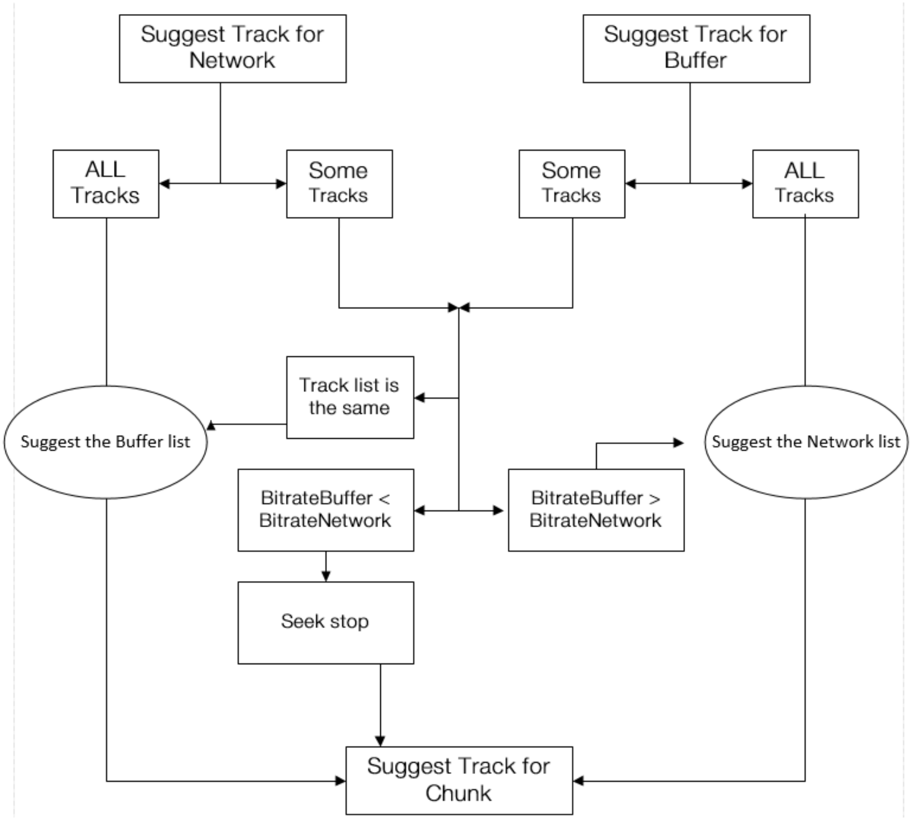

Figure 3.8: Architecture - Combined Network Buffer.

- If both heuristic modules present a list of some tracks, the lists are compared with the next cases:
  - If the network and buffer modules suggest the lists with the same maximum bit-rate at the track, then it is suggested the same list.
  - Otherwise, if the maximum bit-rate value of network heuristic is higher than the maximum bit-rate value of the buffer heuristic, there is an exception. The chunk can not be loaded in the buffer, since there is no storage capacity, and the solution is to stop seek chunks, until re-start the module heuristics. The other case, when the maximum bit-rate value of the buffer heuristics is higher than the bit-rate value of the network heuristic, it is suggested the track list of the buffer heuristic.

Briefly, the list selected is the one with the lower bit-rate track, but there is an exception: when the network and buffer list are not filled and the maximum bit-rate in the network list is grater than the buffer list.

The maximum bit-rate in the tracks list selected is added at the scheduler for download.

Network

The network can obtain various metrics; in this case it is obtained the perceived bandwidth and latency (estimation of network performance). The bandwidth is the total capacity of the connection, which is the amount of data flowing through a network connection over time, measured in bits per second (bps). In this case, the player calculation estimate based on time in downloading packages to the buffer (Equation 3.1) is the following:

$$PerceiveBandwith = \frac{Size_{bits}}{Duration_{seconds}} \tag{3.1}$$

The perceived bandwidth is estimated in streaming session (through the exchange of packages between the client and the server), by the package size ($Size_{bits}$) and elapsed time ($Duration_{seconds}$) (Equation 3.1). The latency value includes several types of delays, typically incurred in processing network data: the wireless connections are usually more affected by path loss and the player obtains a high latency in requests. In this case, the latency obtained in the player is affected by the server plus network; the estimated value is obtained by sending the requests from the player to server; usually the data packages size are small to allow lower estimated error.

Buffer

The buffer heuristic specifies the management of the main stream and manages the state of the buffer. The management mechanism is relevant to the level of loading buffer. There is a range of confidence of the state of the buffer to prevent the stop of the video stream in cases where the network is affected at short intervals. In the heuristics, the track chooses based on the behaviour of the buffer, growing or lowering the level of content.

3.5.1.5 Frame Rate

As a way to analyse the performance of the terminal device (client), there is the heuristic of FPS, which mainly analyses the graphics hardware. If the terminal device is limited on the FPS low values, it may be due to graphics performance. The solution is to switch for a track with lower bit-rate, and thus, it increase FPS. The equation 3.2 (video uncoded) presents this relation of the bit rate with the FPS: the bit-rate value depends on the FPS, the hardware performance is limited bit-rate (bps); when there is a higher bit-rate value, it limits the FPS. Bit-rate is a value defined by pixels per seconds. It is obtained by the window size multiplied by frames per second and the compression obtained in the codec. The heuristic only depends on the user device: in this case the FPS obtained value is an advantage for the assessment of the QoE, because the operator has access at device performance.

$$Bit-rate = \frac{Width \times Height \times FPS \times Compression_{Reduction}}{1024} \tag{3.2}$$

3.5.1.6 Window Size

The universal use of multiple devices with different screen sizes and resolutions is the need to limit adverse events: to play a video stream with high resolutions, the device screen can limit the video stream with lower resolution. This heuristic aims to create a filtering track in order to eliminate the track of stream video with larger size than display; this information is present as a feature of the track from Manifest file. For example, the heuristic can prevent a mobile phone screen with 320x480p to receive a video stream with 1280x720p.

3.6 Conclusions

This chapter studied the existing architecture of the smooth streaming player. First, the structure of the technology was introduced: it was referred the procedures in the selection of the bit-rate, the file formats and the format chunk. Then, the manifest was presented, as well as the features and overview of the connection between the client and the server. The structure of the heuristic module was also explained as the process of obtaining heuristics. This chapter presented the introduction / basis for the study of the method of obtaining the service quality. The study of the architecture of the smooth streaming player was crucial to understand the heuristics and the process for obtaining the metrics for the player.

Chapter 4

Quality of Experience

4.1 Introduction

This chapter will present the architecture of the predictive method to determine the quality of experience based on heuristics (section 3.5) of the video with automatic rate adaptation. In the proposed method it will be necessary to choose the type of assessment (section 2.5.2) and the metrics from the smooth streaming player. Then, the architecture will be presented and it will be explained how the predicting process of quality of experience is performed. The structure of the chapter is the following:

- **Section 4.2** In this section, the methods of quality assessment will be remembered together with the circumstances in which they will be applied.
- **Section 4.3** In this section, it will be presented the heuristics that will be used in the video stream and their advantages in the architecture.
- **Section 4.4** In this section it will be proposed the architecture that allows to obtain a QoE in the method of adaptive streaming video.

4.2 Assessment method

In this dissertation, the creation of the prediction method for the values of quality of experience requires the use of calibration methods, for example the subjective assessment and objective evaluation (section 2.5.2). Figure 4.1 shows these two evaluation methods. The subjective method is used with a database with real human indicators. The objective part has more concrete factors to be used in the assessment, for example: bit-rate, interrupts, initial playout delay, delay time and ratio between screen size and resolution. Subjective assessment is a reliable source to compare results of the objective method. However, there is the need to comply with ITU recommendations to improve the results obtained from surveys. ITU recommends always more than 50 surveys [64] to ensure sufficient confidence.

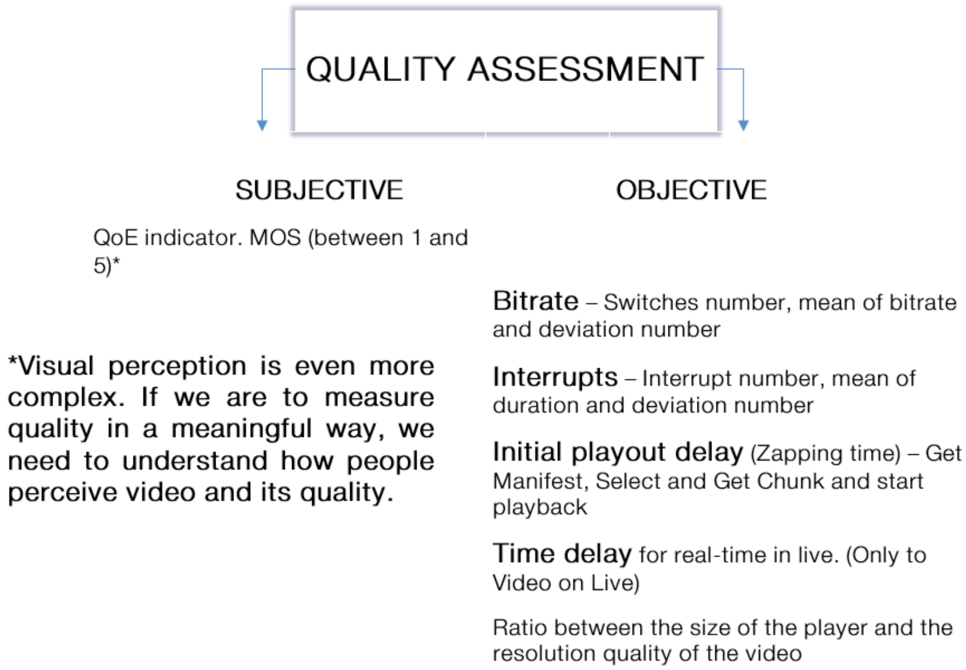

Figure 4.1: Assessment method.

4.3 Heuristics

The selection of heuristics is essential to design the method. The method can only use the metric that are obtained in the player structure. As mentioned in section 2.5, there are several works that relate studies in this area, but they usually consider only one or two metrics. The increase on the number of metrics increases the complexity of the method, but reduces the forecast error. In this area of research, the methods become quite complex when it is desired to obtain the quality of user experience.

4.3.1 Bit-rate

Bit-rate metric contributes largely to the quality of service of the network, as well as the quality of user experience. Video bit-rate is encoded has a constant value for each track to transmit and apply the smooth streaming technique, but fixed bit-rate is a problem with videos that have a higher rate of change of pixels, because these are affected with loss in compression. In Meo Go service, the encoded bit-rate values are included in the range of 250 Kbps to 3 Mbps.

4.3.2 Screen resolution and bit-rate [ratio]

The relation between the screen resolution and the video track is relevant to estimate the quality of experience for the user. Thus, it is possible to relate the quality of experience from a device with a small screen and a device with a large screen. The screen size is usually standard, and it depends on the user distance from display (this effect was studied in article [58]), so it is required that the stream resolution is within range of the screen device.

4.3.3 Frame rate

In the service streaming of the channels television, for example Meo Go, there are different channel types (news, sports, ..) with different quality of the video (quality changes with channel type), so the compression is performed by the main frames and frames that only contain pixels changed (frame-by-frame compression [28]). A video streaming with large pixels changed in frame-by-frame (for example the sport channel) has a lower rate of compression and it results in the encoded bit-rate to decrease the video quality. In the smooth streaming, the bit-rate of the video encoded is fixed and the quality of experience can alter with the type of channels. The solution to this problem is to simulate channel types with profiles and adjust the FPS with the Equation 4.1. FPS is updated with a variable (K_{FR}) that has a range between 0 and 1; for example, FPS adjusted to a lower value in sports video ($K_{FR} < 1$).

$$F_r = K_{FR} \times F_{RameRate} \tag{4.1}$$

In the real scenario, it is necessary to know the update influence, because K_{FR} range can vary with a small or large frequency. An example follows in the Figure 4.2 for the various types of channels. The Figure 4.2 contains the influence of the channels genre by a work developed in [57]. In the graph, the documentary genre is the channel with less Mean Opinion Score (MOS) value, but the Confidence Interval (CI) covers all genres.

The channels genre has low variation (MOS), but this metric can be used to increase the estimated precision of the value of quality of experience. This influence may be compared to the limited frames: it is in the sudden movements in the video that this effect is observed; to disregard it, it can be inserted into the FPS adjustment (Equation 4.1); when there is a low quality, it can be emulated by lower FPS.

4.4 Method & Procedure (QoE)

In this sub-chapter, it will be presented an architecture to predict the values of quality of experience. This value is influenced by many factors such as heuristics (listed in section 4.4.1.1) and human characteristics including memory. The classification (Figure 4.3) of a video streaming session at any instant is influenced by previous samples, so this architecture presents a first phase which classifies video chunks, with 1 second for example, and then updates this value based on previous classifications. The architecture is influenced by characteristics of video and human memory. Moreover, this method needs to be calibrated when it is deployed. The purpose of the final solution is an objective method that can be inserted in the company PT Comunicações platform.

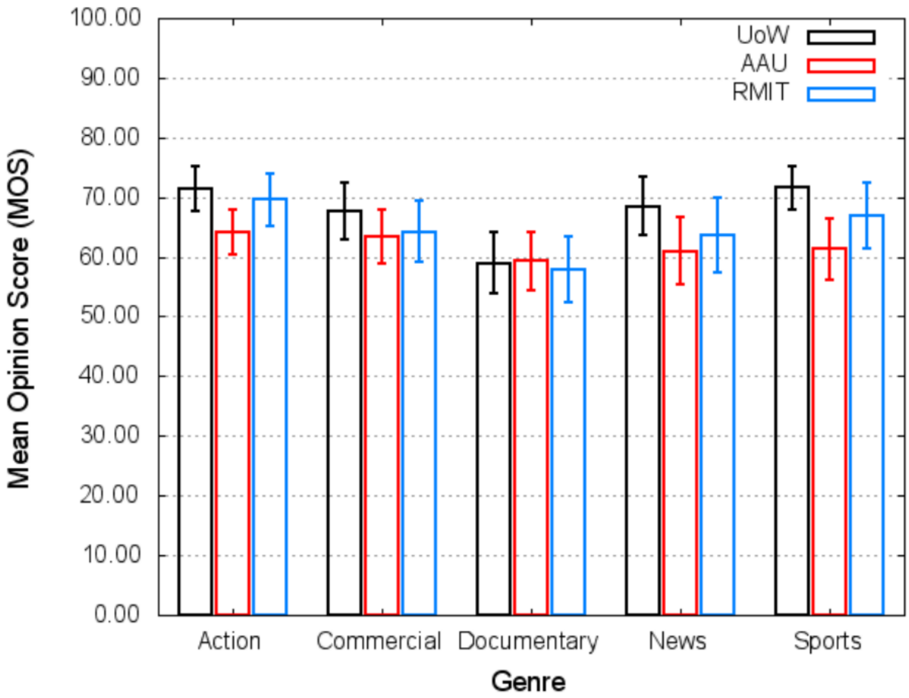

Figure 4.2: Enhancement of QoE for each Genre for all three user studies [57].

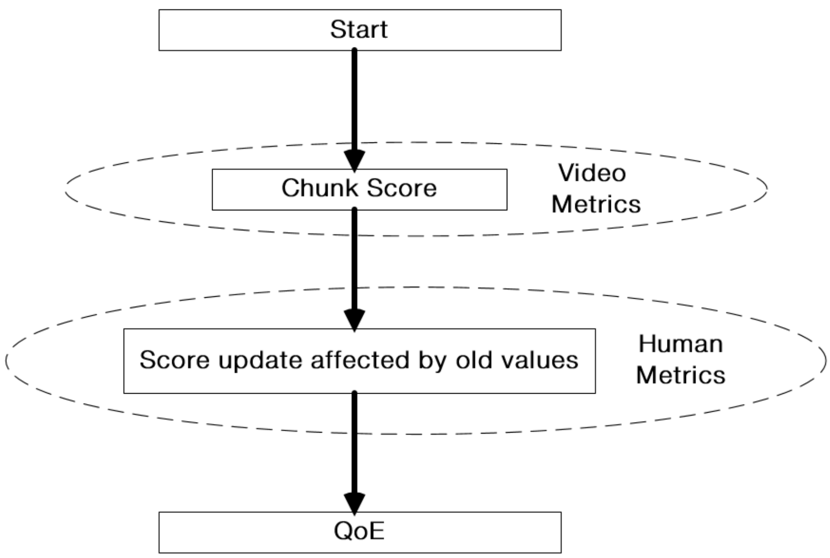

Figure 4.3: Method & Procedure - Architecture.

4.4.1 Chunk Score

Initially, the method aims to create an objective assessment of video chunks that classifies features that are independent over the time; for example, the bit-rate and the FPS.

This evaluation aims to be objective and create an equation that simulates the effects of the video characteristics. The equation is obtained by analysing the behaviour of the metrics and subjective methods and must be used for values calibration. There are several tools that contain databases and allow to obtain the quality of experience of the video.

PEVq software allows to use the trained database with reliable values by the ITU recommendations [52]. There are several examples of works that investigated effects of video, such as the one in [43] that refers to a study assessing the quality metrics in the context of freezing video, which evaluates the quality of user experience based on the number of freezing.

4.4.1.1 Perceptual Evaluation of Video Quality (PEVq)

PEVq [52] is an example of software used in databases by OPTICOM Company. It analyses the degradation of the original video (reference) and the video received (impaired); therefore, it is always necessary to have a reference (intrusive system).

Figure 4.4 contains an intrusive test scenario using PEVq software. The evaluation is carried out on samples collected near the server and close to the client. This type of intrusive systems is complex to implement in practice by an operator; the non-intrusive systems have many advantages in terms of the level of implementation.

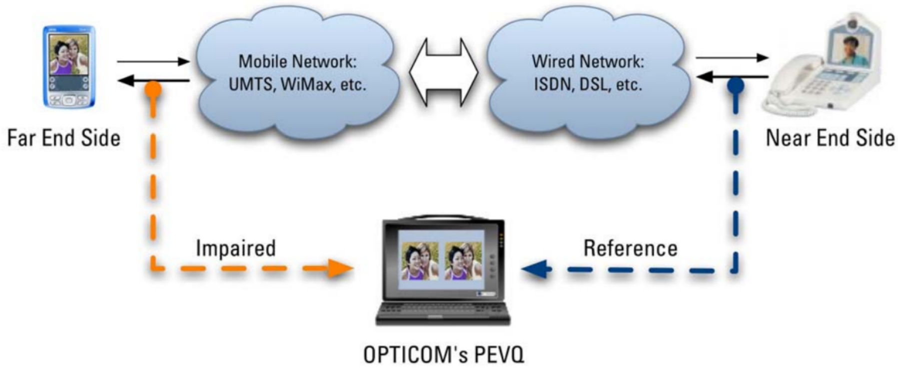

Figure 4.4: Intrusive scenario [52].

In the case of smooth streaming technology, the files do not degrade video, and received packets are always complete (TCP) but with a delay (there are only retransmissions).

Figure 4.5 shows an overview of the software PEVq, characterized by an intrusive system (second signal) over the time. This process starts with the temporal and spatial alignment in videos to be possible to analyse the perceptual differences using YUV color image (Y, Cr, Cb) (section 2.3.1). In the next module, PEVq software classifies the distortion (distortion classification module) using a database trained by humans. These databases trained by human allow to get MOS values for specific factors (metrics), for example bit-rate and FPS. Finally, there is a module that integrates the classifications over time to return a final value (PEVq Score).

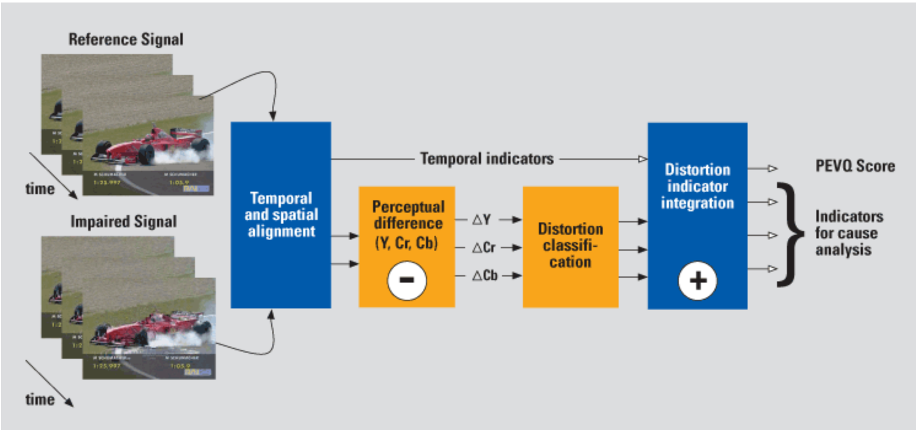

Figure 4.5: PEVq Architecture [52].

In smooth streaming, the tracks are transmitted with different bit-rates; the degradation of the video is done on the server and then it is transmitted. Thus, PEVq software is possible to use in the video score with the comparison between degradation of the network and the quantified files in several qualities. This software may allow calibration of the chunk score, as noted above. The server contains a track with a maximum bit-rate and other lower bit-rates that were distorted with decreasing bit rate, so it can equate to a distortion in the network.

The software prediction is necessary to verify the characteristics of the video and compare them with the metrics of the smooth streaming player. The requirements of the OPTICOM platform are as follows:

- Presents a mapping scale of 1 (bad) to 5 (excellent) for the perceptual evaluation of video quality.
- The input parameters only accept uncodec video (YUV Raw Files and AVI) with 2.5 up to 30 fps rate (require conversion from the MP4 to the RAW).
- Ability to use profiles with Common Intermediate Format (CIF) (352x288 pixels), Quarter Common Intermediate Format (QCIF) (176x144 pixels) and Video Graphics Array (VGA) (640x480 pixels), but in smooth streaming it has various resolutions and these can go to Full HD (1920x1080 pixels).

In the evaluation process, the chunk is calibrated with PEVq software. In the assessment of video quality, the video is analysed by the quantity of pixels per second (bit-rate (bps)), by the processing performance of the player on the client with the FPS, by the relation of the size of windows device with default size defined in the track, and by the network congestion status by the re-buffering metric. Briefly, the chunk score is obtained by the following metrics:

- Bit-rate [bps]
- Frames per Second [FPS]
- Re-buffering time [seconds]
- Windows/Resolution [ratio]

4.4.2 QoE update over time

The chunk score mechanism is based on metrics of the video and do not depend on time. Quality of experience depends on the samples at current and previous time, and the human mind is responsible for this behaviour. It follows an example to understand the effect of human memory: when the quality of a stream increases to one higher quality, it is better than to decrease to the minimum, but it has the same average value. The method, in a first step, obtains the score chunk value depending on the video characteristics and then, in second phase, it updates the score chunk values based on a filter of human memory with previous values (this method has a sampling frequency (fa = 1Hz)).

The mechanism of human memory is an area with many research and development works; for example, the work in [25] proposes a method to build a database of human mental data including the human losses by forgetfulness. In this step, the main objective is to update / get the final value of the QoE, based on a model with time varying and lowest approximation error. Time-Varying Subjective Quality (TVSQ) report presents the effect for variation in quality over time based on the filter [13].

4.4.2.1 Time-Varying Subjective Quality (TVSQ)

The work in [14] presents a paradigm for time-varying subjective quality estimation [38][60] (Figure 4.6), but the dynamic part of the model is only TVSQ. TVSQ simulates human performance with several video conditions, for example, re-freeze frame and re-bufferings that, after these values, are stored in databases to process the methodology. General model of Hammerstein-Wiener[70] (Figure 4.7) was used to find a human filter solution, and this model has a filter that requires non-linear inputs and outputs. In this model, it is obtained specific quality ($q^{st}[t]$) in the first phase, which can be corrected ($\hat{q}[t]$) to approximate the value of MOS (Figure 4.7). This work notes that the first 15 seconds of human memory have more impact in the quality of experience.

Figure 4.8 shows the IIR filter with 30 samples (assumed: f = 1Hz) that can simulate the human memory memory [26]. The response of the filter is only about the first 30 seconds that influenced the update of MOS value.

Initially, in the evaluation of the video stream, the method starts the playback without

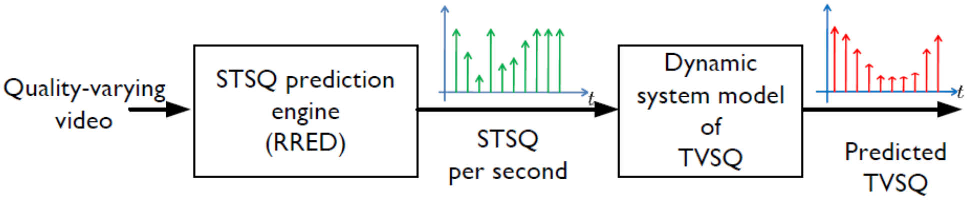

Figure 4.6: Time-varying subjective quality estimation [14].

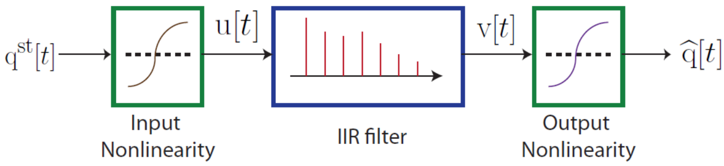

Figure 4.7: Hammerstein-Wiener model for TVSQ prediction [13].

---

previous values. There is no historic of the session information when the video stream is started, so the lack of previous values is initially a problem, since the last 30 samples (f = 1Hz) are necessary for the method to apply the filter of human memory.
The solution proposed adjusts the influence of the samples to 100 percent; for example, the first evaluation (starts at 0 seconds) of the chunk depends only on the current value. The next evaluation (starts at 2 seconds) depends only on the previous and current value. These values are adjusted to 100 percent, with the previous value influencing 63.75 % (40 %, default by filter, more 23.75 % by adjustment), and the actual value influencing 36.25 % (12.5, default by filter, more 23.75, adjusted).
Figure 4.8 shows a graph (filter behaviour) with the influence by previous seconds in the human memory.
Other works verify that the influence of ancient samples is only present for an amount of time. For example, the work in [26] of the human brain presents several types of memories, short and long term. In this case, when the user watches a video, the probability of information to stand in the long term memory is very low and it is only used the short memory.

Figure 4.9 shows the architecture proposed of the method that obtains the QoE. The architecture can be divided in two processing stages. In the first phase, the method uses metrics of the network and the user device in the mathematical expression to obtain the video quality. In the second phase, the method uses the previous value, obtained in the first phase, to update based in the previous QoE values and in the influence of human memory. This influence is based on 30 previous samples that are updated in each second.

In the chunk score update (method implemented), sampling frequency is very crucial to select the chunks, because these chunks, contain 2 seconds of the video; thus, the chunk update can be done in seconds (figure 4.10). There is an exception when the time of re-

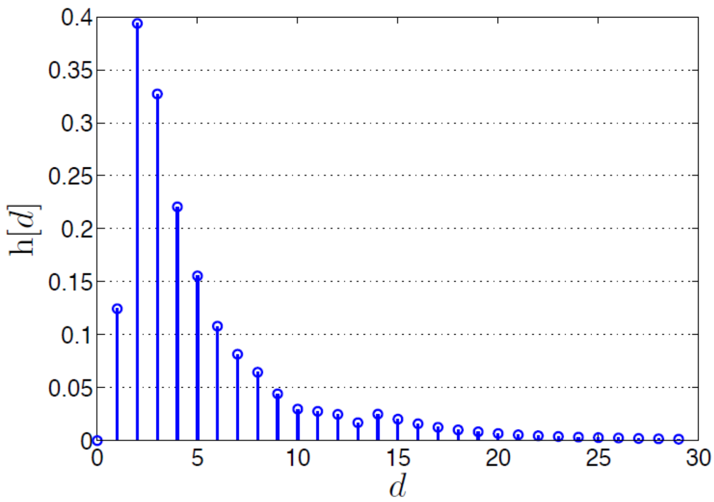

Figure 4.8: The impulse response of the IIR filter in the first 30 seconds [13].

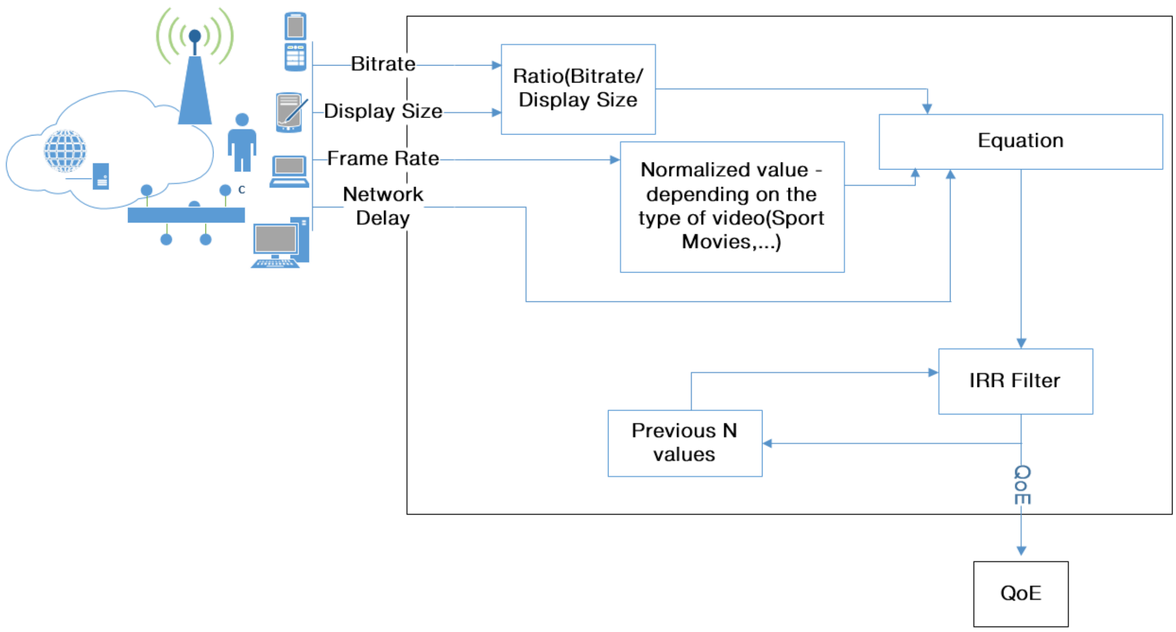

Figure 4.9: Architecture to obtain the QoE.

bufferings does not have integer value; so the chunk can be evaluated 10 times in the 1 second (100 ms sampling (figure 4.10)); subsequently it is performed averages of 10 samples. Briefly, QoE update occurs each 1 second and the chunk score (characteristics of the video) has a sampling rate of 10 samples within 1 second (100 ms).

In this architecture, it is necessary to create an equation that relates all metrics and then returns the video quality. After this, it is necessary to check the influence of the human memory filter. The influence of the human memory filter can be checked by the real scenarios that use the subjective methods (surveys).

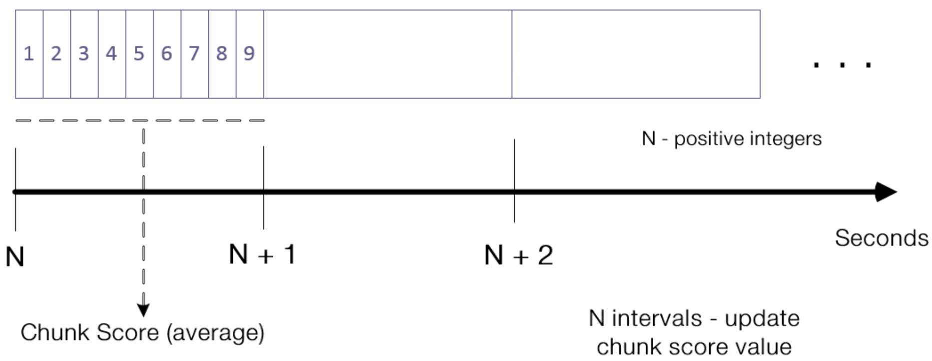

Figure 4.10: Method proposed - Sampling Frequency.

4.5 Conclusions

In this chapter, the architecture of the QoE evaluation method was proposed to obtain the quality of experience in adaptative streaming. Initially, the section 4.2 described the possible types of evaluation in the QoE measurement and its adaptation to the service Meo Go. PT company aims to integrate the method into their systems, so the method has more facilities to import to ArQoS [55] system (probe). The section 4.3 presented the metrics that the player can export to the proposed method. This selection is based on the smooth streaming architecture (section 3.5) and the more influenced metrics to QoE. Subsequently, the architecture of the method of obtaining the QoE was presented in two main phases: one that influences the quantification of the video quality and another that considers the influence of the human memory. In conclusion, the architecture can be implemented in a module that provides an adaptive quality video stream to the operator.

Chapter 5

Implementation of Heuristics Monitoring

5.1 Introduction

This chapter presents the implementation of the solution to predict the QoE value of adaptative video streaming. As mentioned in the chapter 4, the objective and subjective evaluation are closely related; the objective method is always preferred, but subjective methods are required to calibration. So, this chapter aims to implement the proposed solution and it is divided into two parts:

- **Section 5.2** In this section, the objective method will be implemented based in the proposed architecture (Chapter 4.4). The prediction is performed in two phases: the first is an evaluation of the chunk based in characteristics of the video; in the second phase, the chunk score is updated with the previous samples (based in a human memory filter).
- **Section 5.3** In this section, the subjective method will be implemented using questionnaires with test scenarios. To facilitate the collection of the results, it is developed a web page with videos that contain different characteristics and a bar for QoE submission evaluation.

5.2 Objective QoE Assessment

The assessment required using the objective method allows to monitor the network performance metrics and the terminal device (Client). This quality management of user experience has advantages for the provider of the video streaming service, thus it can improve the packages delivery of the stream service. In this section, the proposed method will be implemented in two phases; Initially, this method allows to score the video, based of the network and terminal, Bit-rate, FPS, re-buffering time and size of windows device. In a second phase, it is performed an update to the chunk value based on a filter of the human memory. This implementation process is treated mathematically with both MATLAB and EXCEL tools.

5.2.1 Score Chunk

In the first phase of the prediction value implementation, the method is based on specific features of the video. This method aims to be implemented in a mathematical expression for possible integration into the PT communications system. The score chunk aims to assess the specific characteristics of the video. In the assessment of the metrics, OPTICOM tool is used to get a QoE method. This tool also allows to take benefit of database in order to calibrate the QoE method. The following video characteristics are evaluated:

- Bit-rate [bps]
- Frames per Second [FPS]
- Re-Buffering [Seconds]
- Windows Size & Resolution Track [Ratio]

5.2.1.1 Analysis - Heuristics parameters

Briefly, this work is divided into two parts. First a general equation is obtained and then it is trained. The best way to obtain a general equation that approximates the metric of a QoE value, is first to observe the behaviour of each metric separately; because this way, it is much easier to analyse the metric behaviour. Only then, we combine the results together. Finally, the general equation is trained by using MOS values, that were acquired through the database provided by the PEVq software. This method is a good solution, because it uses the real opinion of many people.

Bit-rate
Bit-rate is a metric that measures the streaming adaptation, and the behaviour can be approximated by an inverse tangent function. This behaviour is defined by the increase of quality. Figure 5.1 shows the increase of the MOS with the bit-rate and it is obtained by the work in [54], which developed and tested the metric as a function of several factors, by real human opinions.

Frame rate
FPS introduces a behaviour as the one that can be seen in Figure 5.2. This curve was obtained taking into account several works, such as [29]. This metric tests the behaviour of the filter that displays the human eye (around 24 FPS), and large FPS values allow an optimal MOS value. The curve can be estimated by a logarithmic function.

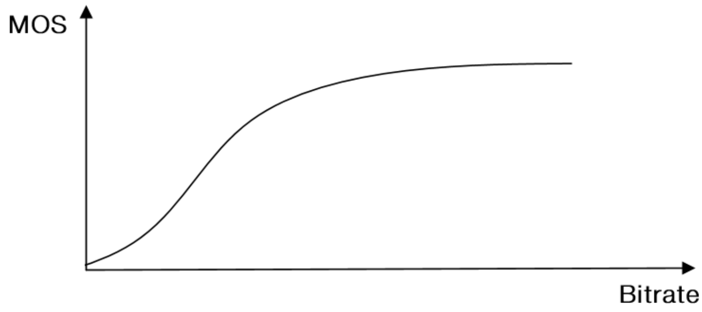

Figure 5.1: MOS value with Bit-rate.

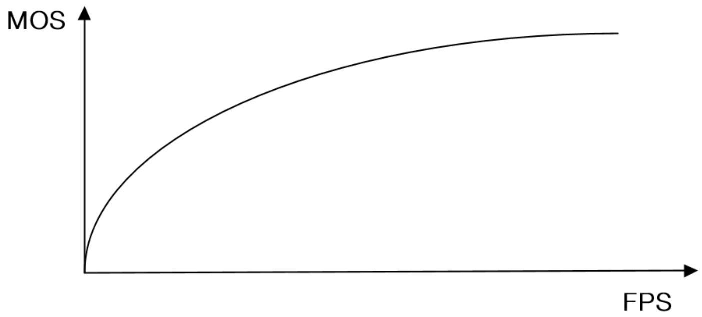

Figure 5.2: MOS value with FPS.

Based on the above analysis, it is possible to verify that the behaviour of the two heuristics can relate with Figure 5.3. With two metrics, the behaviour of frame rates interferes with the evolution of the bit-rate, because the frame rate is crucial if the video allows to have a sequential session.

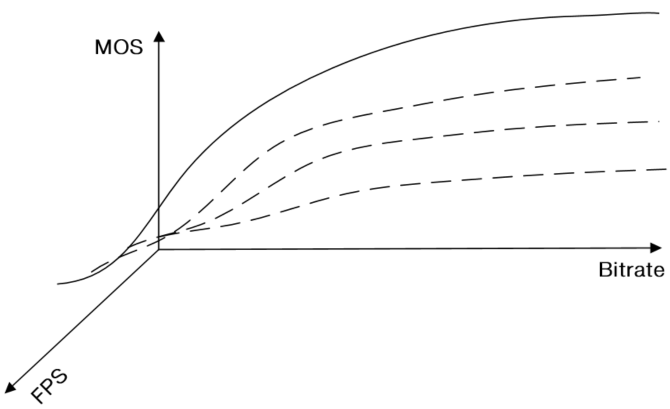

Figure 5.3: MOS with Bit-rate & FPS.

Re-buffering

The re-buffering is the metric that has the worst impact on QoE, because a user does not want to see the video with stops (misses the reality of the session). With the increase of bit-rates of the stream (HD and Full HD), the network may congest and then the videos are affected by the stops in the session. People prefer lower quality and to see a continuous video stream.

Figure 5.4 presents the MOS behaviour with the re-buffering. The behaviour starts with a reliability zone (shift at Figure 5.4) that keeps the MOS value and it is caused by human perception. People knows that a video has to do a bit of buffering. After the start of the session, the re-buffering can degrade very much the quality of experience. Thus, the behaviour can be represented by a logarithmic function, since it is different to wait 2 or 4 seconds or 20 or 25 seconds.

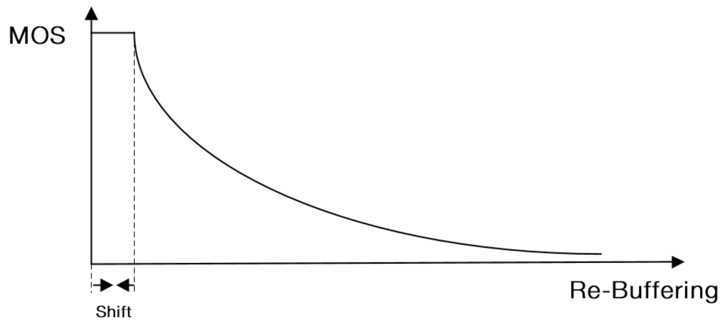

Figure 5.4: MOS value with Re-buffering time.

Ratio between windows size & default size

Nowadays, the resolution of video is very crucial in the QoE. There are large displays and small displays, but the broadcast streaming can be played on each of them. Figure 5.5 shows the bit-rate with the horizontal screen resolution, so the quality of experience depends on the client device. The ratio needs to be a value between the size of the window and the resolution recommended. The behaviour of the ratio with the MOS can be represented by a logarithmic function. Figure 5.6 shows this behaviour that was obtained by the work developed in [37], but also it can be confirmed by the following expression: "Therefore a MOS of 3 for a small screen means that the video is about as far from the quality of the best video at that screen resolution as is a video of MOS 3 that is shown on a large HD screen" [12].

Compression

The codec compression is an important part of QoE, but that will not be implemented, since it is considered that the work has the same video codec AVC/H.264. Therefore, irrespective of the codec used, it always uses the same stream. Figure 5.7 presents the

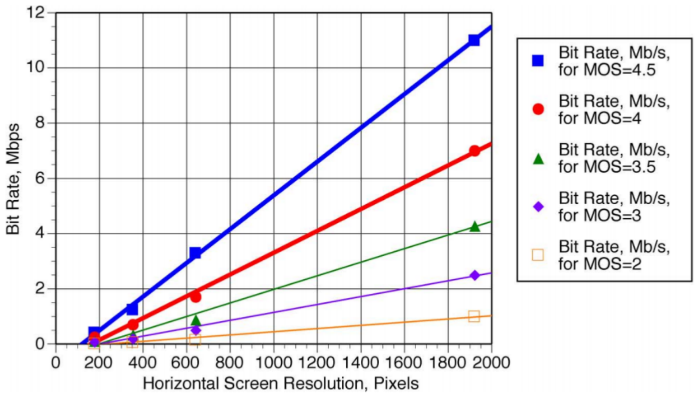

Figure 5.5: H.264 bit rate required for desired level of MOS at different screen resolutions [12].

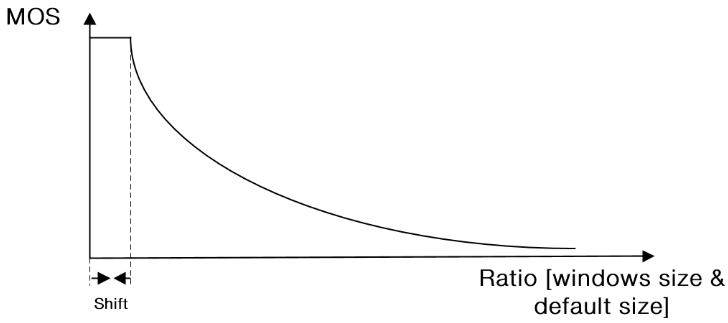

Figure 5.6: The value of MOS with Ratio [windows size & default size]

---

quality of several video codecs; the codec with lower losses is H.264; this is the one used in Meo Go service.

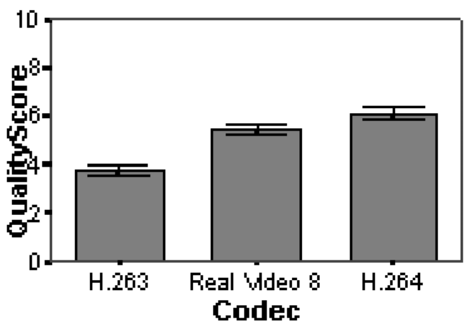

Figure 5.7: Video test: codecs, picture ratio and bit-rates in visual quality [30].

General equation

The characteristics of each metric were analysed in the previous sections. The method proposed aims to relate together the features of the video, in the mathematical expression 5.1. In this case, the characteristics that directly depend on the video are: the bit-rate and FPS (referenced in the session 3.5.1.5), and the mathematical expression 3.2 relates the two metrics through a multiplication. Network performance is analysed by the buffering metric, in which the time of re-buffering is crucial to calculate the quality of experience. This metric will present a subtraction in the general equation. This subtraction is also justified by the independence that exists between the bit-rate and the FPS. The fourth metric is the ratio between the size of the user's device window and the resolution pre-defined by the track. This metric creates the blur frames in the video when the screen size is higher than the standard.
General equation 5.1 combines all metrics, then, it will be necessary to calibrate the values (*v1, v2, ... , v8*), which will be performed in the following sections.

$$\begin{aligned} Score_{chunks} = & v_1 + v_2 \arctan\left(Bitrate \times v_3\right) \times \log\left(v_4 \times FPS\right) - \log\left(v_5 \times Rebuffering + v_6\right) \\ & - \log\left(v_7 \times Ratio_{size} + v_8\right), \end{aligned} \tag{5.1}$$

5.2.1.2 PEVq

In the general equation, the values of unknown parameters are crucial to the final equation of the chunk score. This section implements the calibration method described in section 4.4.1; the proposed calibration is based in the mechanism that uses the PEVq software (MOS database).
This software was built on a trained database which analyses the differences between two videos and returns the MOS; as stated in section 2.3.1, the distortion is evaluated in YUV. Thus, the calibration is performed by creating uncompressed video files with different characteristics, and then it evaluates them in the PEVq program.

The calibration values of 4 videos of different genres were analysed; the bit-rate, FPS, re-buffering time and the ratio between the size of the terminal window and the resolution of the track (resolution default). As there are four parameters and the method is non-linear, the MATLAB tool was used in the implementation.

In a first step, the video was obtained in an uncompressed file (RAW) of a high-quality video [HD (720p), FPS = 30]. In the process to analyse, this video file must have 2 or 8 seconds, because the adaptative streaming uses packages that contain 2 seconds of video with fixed bit-rate.
In a second step, Microsoft Movie Maker was used to edit videos and the reference video was created with the higher characteristics (Figure 5.8) to not lose the quality of the reference video.

In a third phase, the AVS Video Converter (version 8.3) program was used to convert the video to an uncompressed format. Now, it is possible to have an uncompressed video

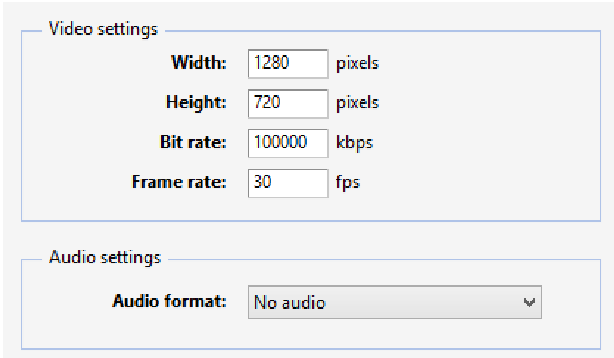

Figure 5.8: Video Characteristics (Movie Maker).

version with 2/8 seconds to use in the PEVq software. In the video streaming, the codification is essential to data, so the uncompressed video will be converted to the format used in Meo Go, Advanced Video Coding (AVC), and then turn to be decompressed; this phase allows to consider the codification effect.

PEVq software requires two converted files: one as the reference file and the other as the test file. The software will calculate the MOS between these files by YUV distortion. The test file contains different characteristics to allow to evaluate the metrics of the quality of experience and it was used the Movie Maker software to edit these videos.
When it is aimed to add a time re-buffering in the video, it is simulated by a black image; it should be noted that it was necessary to keep the same resolution and quality of the image.

In this process, it was necessary to overcome the time taken to run several phases, which took longer than planned due to the use of uncompressed video, large files and slower processing. The size of a compressed video is obtained by multiplying the time elapsed with the result of equation 5.2(the variables are in the Table 5.1), as mentioned in chapter 2.3. A compressed video bit-rate is lower than uncompressed bit-rate; the calibration process used always values of decompressed bit-rate, because the video is analysed by the YUV color image.

Variable	Name	Description
Color	Color	Color Depth (bits/color) with Color Model (RGB,YUV,..) ratio
FPS	Frame rate	Frame Frequency (Frames per Second)
PixelSize	Pixel Size	Dimensions - Wide x High pixels

Table 5.1: Variables of the uncompressed data rate

$$Uncompressed_{DataRate} = Color \times FPS \times PixelSize; \tag{5.2}$$

Calibration values

To calibrate Equation 5.1, it is necessary to have a good amount and range of values to test, but the method implemented has an operation range of the video characteristics that were defined by current streaming characteristics. These ranges are:

- Bit-rate (coded at H.264/AVC1) - between 200 kbps until 5 Mbps;
- FPS - between 2 FPS until 30 FPS;
- Re-buffering - between 0 seconds until 20 seconds;
- Device windows size and track resolution ratio - between 1 until 19.96;

MOS values were obtained and generated by the database (PEVq software) with the different video characteristics. It is used the four videos with different characteristics.

The Surface Fitting Toolbox (sftool) function of MATLAB was used to estimate the values of R-squared obtained from the Eq. 5.1 in relation to the original. This MATLAB function presents a graphical interface to display 3D graphics.

Bit-rate & FPS

The general equation consists of four variable values and various fixed values (calibration values). The best way to perform this calibration is to use a tool that accepts 4 input variables (Bit-rate, FPS, Re-buffering and ratio size), but there is no recourse to a tool that enables this advantage.
The relation between the Bit-rate and the FPS is calculated when the Re-buffering and the ratio size does not influence the QoE, so it is assumed that the Re-buffering is 0 seconds and the ratio size is 1.

In the Figure 5.9, the results obtained, by PEVq, are characterized by inverse tangent (as in the Section 5.2.1.1). These results are used to calibrate the Equation 5.3.

Subsequently, the program calculates the fixed values $v_1,\ v_2,\ v_3,\ v_4$, that allow to calibrate this singular equation. The variable existence, v_4, is an offset to build the singular equation 5.3 in the general equation 5.1.
The R square value was obtained by the *sftool* program to verify if the equation approaches the information of the database trained; the value is 0.9 (near to 1), so it confirms the confidence in the equation.
The order of metrics calibration starts first with the more relevant metrics. The ratio of the window size and the resolution of the track has less relevance, and the FPS and Bit-rate have more relevance in the quality of experience of the user.

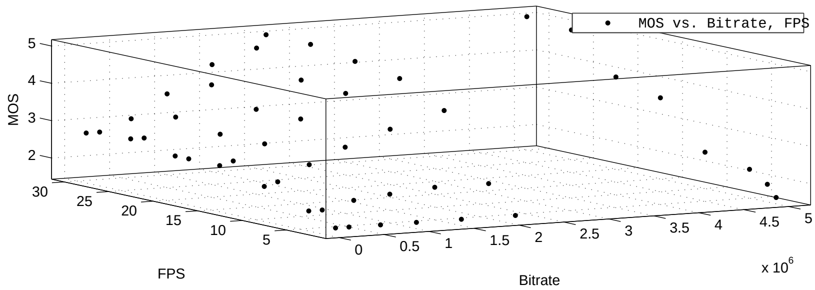

Figure 5.9: Graph - (Bitrate,FPS,MOS).

$$
\begin{aligned}
&Assume\ in\ the\ equation\ 5.1:\\
&\qquad Rebuffering = 0\\
&\qquad Ratio_{screen} = 1
\end{aligned}
$$

$$f(Bitrate, FPS) = v_1 + v_2 \times atan((Bitrate \times v_3)) \times log(FPS \times v_4) \tag{5.3}$$

$$
\begin{aligned}
&Coefficients\ (with\ 95\%\ confidence\ bounds):\\
&v_1 = 2.038(1.722, 2.354)\\
&v_2 = 1.027(0.8296, 1.224)\\
&v_3 = 1.42^{-06}(7.265^{-07}, 2.114^{-06})\\
&v_4 = 0.3031(0.2014, 0.4049)
\end{aligned}
$$

$$
\begin{aligned}
Goodness\ \ of fit &:\\
SSE &= 6.912\\
R_{square} &= 0.8817\\
AdjustedRsquare &= 0.874\\
RMSE &= 0.3876
\end{aligned}
$$

The dots in the 3D Graphic in Figure 5.10 represent the real values of the trained database, and the surface represents the result of the Equation 5.3 with the calculated coefficients.
The real values and the ones of the equation have slight differences: for example, in the point (2.5Kbps, 30 FPS, MOS = 4.7), the equation has a lower value.

Bit-rate & FPS & Re-buffering

In this section we will include the metric re-buffering in the equation. The new equation

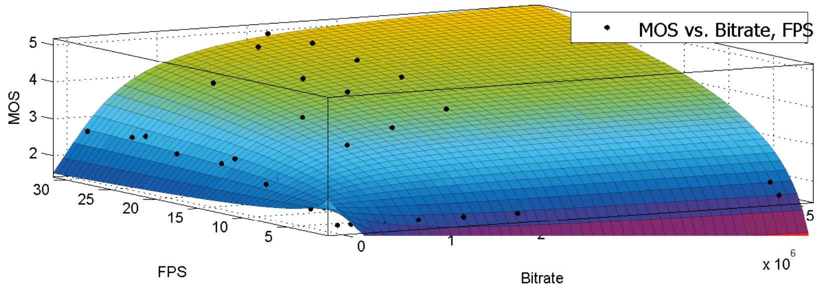

Figure 5.10: Surface - (Bitrate,FPS,MOS).

f contains (bit-rate, fps) and the logic of re-buffering, so the equation calibrated will have 3 metrics. To obtain the Equation 5.4 we assumed that the $Ratio_{screen} = 1$. The calibration was processed with the PEVq values by the *sftool* software.
The program allows to return the R square value that is used to confirm the approximation between the values used in the calibration and the curve obtained by the Equation 5.4; in this step the value 0.9083 of the R square is a value that offers assurance. Figure 5.11 presents the results that confirm the expected behaviour in the Section 5.2.1.1, a logarithmic function with a large decrease.

$$Assume\ in\ the\ equation\ 5.1:$$
$$Ratio_{screen} = 1$$

$$f2(Bitrate, FPS, Rebuffering) = f(Bitrate, FPS) \ - \ log(v_5 \times Rebuffering + v_6) \tag{5.4}$$

$$\begin{aligned} Coefficients\ (with\ 95\%\ confidence\ bounds):& \\ v_5 &= 3.064(1.291, 4.837) \\ v_6 &= 0.5407(-0.9883, 2.07) \end{aligned}$$

$$\begin{aligned} Goodness\ \ of fit:& \\ SSE &= 0.07071 \\ R_{square} &= 0.9312 \\ AdjustedRsquare &= 0.9083 \\ RMSE &= 0.1535 \end{aligned}$$

Bit-rate & FPS & Re-buffering & Size(ratio)

Finally, the equation has 4 metrics, when it is inserted the last metric (ratio of the

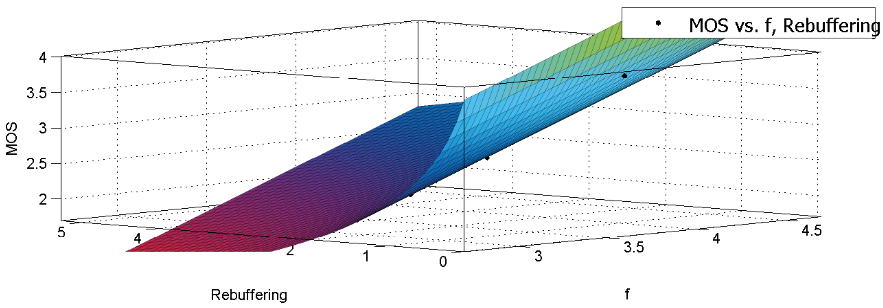

Figure 5.11: Surface - (f,Re-buffering,MOS).

screen size and resolution of the track). Thus, the Equation 5.4 (f2) will be referenced to the calibration of previous values. In the process for calibrating the general equation, the R square value had 0.91 of average in the 3 steps, which provides confidence in the deployed process. Figure 5.12 provides a curve with the effect of the screen size of the method and the values obtained in the PEVq software.
Accordingly, the equation is calibrated with the metric value of screen ratio, and then it is obtained the general equation.

$$\begin{aligned} f3(Bitrate, FPS, Rebuffering, SizeRatio) = & f2(Bitrate, FPS, Rebuffering) \\ & - \; log(v_7 \times SizeRatio + v_8) \end{aligned} \tag{5.5}$$

$$\begin{aligned} & Coefficients\ (with\ 95\%\ confidence\ bounds): \\ & \quad v_7 = 0.05652(-0.001239, 0.1143) \\ & \quad v_8 = 1.756(1.242, 2.269) \end{aligned}$$

$$\begin{aligned} Goodness\ \ of fit : & \\ SSE : & \ 0.01182 \\ R_{square} : & \ 0.9091 \\ AdjustedRsquare : & \ 0.8636 \\ RMSE : & \ 0.07687 \end{aligned}$$

Equation

Equation 5.6 is a general equation that relates of the metrics behaviour and it is obtained by the expansion of the Equation 5.5. The calibrated equation is obtained with the determined values of *v1, v2, v3, v4, v5, v6, v7, v8*. This equation gives the MOS value with respect to the characteristics of the video, and there is the need to measure the evaluation of user experience in a streaming session. Thus, this equation calculates a chunk value that will later be upgraded with a filter of human memory.

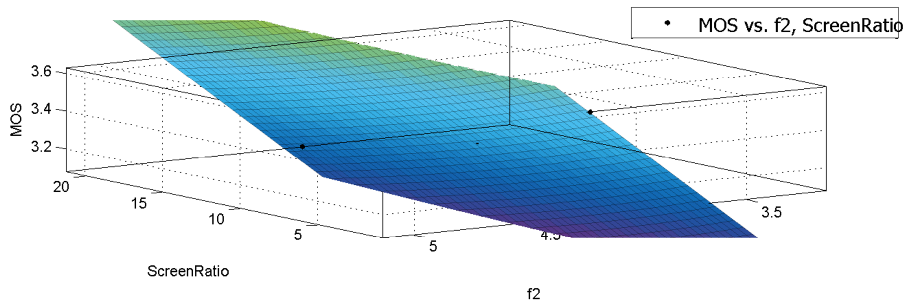

Figure 5.12: Surface - (f2,Size(ratio),MOS).

$$v_1 = 2.038 \qquad v_2 = 1.027 \qquad v_3 = 1.42^{-6}$$
$$v_4 = 0.3031 \qquad v_5 = 3.064 \qquad v_6 = 0.5407$$
$$v_7 = 0.05652 \qquad v_8 = 1.756$$

$$\begin{aligned} Score_{chunks} =& v_1 + v_2 \arctan\left(Bitrate \times v_3\right) \times \log\left(v_4 \times FPS\right) \\ & - \log\left(v_5 \times Rebuffering + v_6\right) - \log\left(v_7 \times Screen_{size} + v_8\right), \end{aligned} \tag{5.6}$$

5.2.2 Filter of Human Memory

The algorithm that filters the human memory (Figure 4.8) was originally developed in the MATLAB software. It is created a script with the equation 5.6 to score chunks, and the obtained values in this equation are updated by the filter of the human memory. The frequency is defined by a sampling period of 100 ms. In each second, there are 10 samples of the chunk score, and it is used the mean to update the QoE value. Regarding the filter, it is important to note that, at the beginning, there are no previous values. The implemented solution is done by the sum of weights that already have values. After 30 seconds, we have all necessary QoE values.

5.2.2.1 MATLAB function

In the implementation of the method, first of all, it is necessary to evaluate the chunk. Then, the value based on the filter of the human memory is updated. It is used a Equation 5.6 to obtain the score of chunks. Those scores are stored and subsequently passed through the filter of human memory (Section 4.4.2.1, Figure 4.8). Thus, a function was constructed to accept arrays as an input, in which each position represents 100 ms.
It is possible to create more complex scenarios to be able to predict the quality of consumer service. These scenarios will be made in the streaming session, where quality is constantly changed, such as bit-rate, FPS, and other metrics, to prove the implementation of the

method.
Figure 5.14 shows the structure method to obtain QoE that was implemented in a MATLAB function. This implementation is divided in two steps: in the first step, the metrics are obtained and then the chunk score is rated; In the second step, the filter of human memory uses the mean of the chunks to estimate the QoE value.

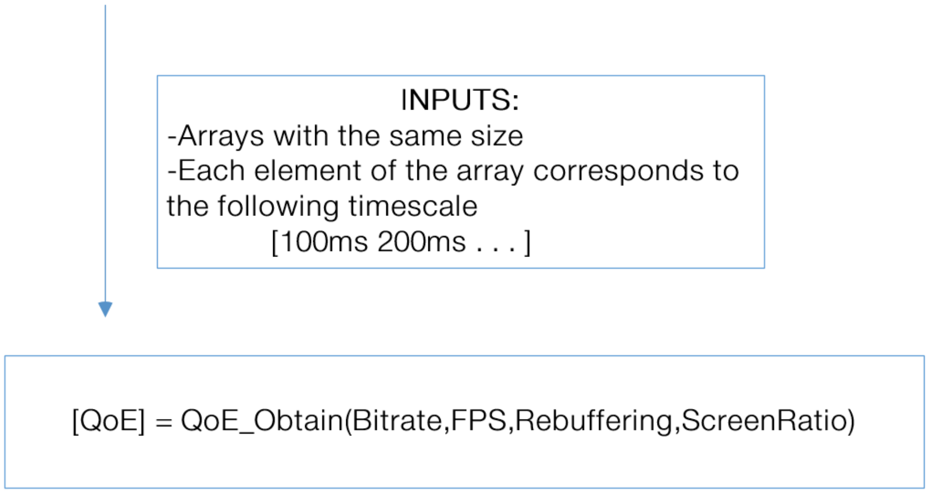

Figure 5.13: Function interface

The implemented MATLAB function is characterized in the Figure 5.13, where it receives as input parameters, the Bit-rate, FPS, Re-buffering and Screen ratio. These parameters will be ranked every 100 ms using the Equation 5.6, and later the average is performed with 10 samples every second. This average will be updated in the filter of human memory based on samples from previous quality of experience. At the end, there is a value of QoE updated every 1 second; this time interval is below the value of the length of chunks (2 seconds).

5.3 Subjective implementation

Subjective assessment assesses the validity of the implemented method and it aims to obtain real values of quality of experience for the user. There are several scenario tests created with different characteristics of videos. The test scenarios will be presented in Section 5.3.1. We consider the ITU recommendations to perform subjective validator; where a main recommendation [24] with is more than 50 replies in the questionnaire. It was created a web page in order to simplify the access and to easily scale the number of replies in the questionnaire. This web page presents 20 video tests that are defined by video characteristics, like the bit-rate, the FPS, the re-buffering and the screen-size.

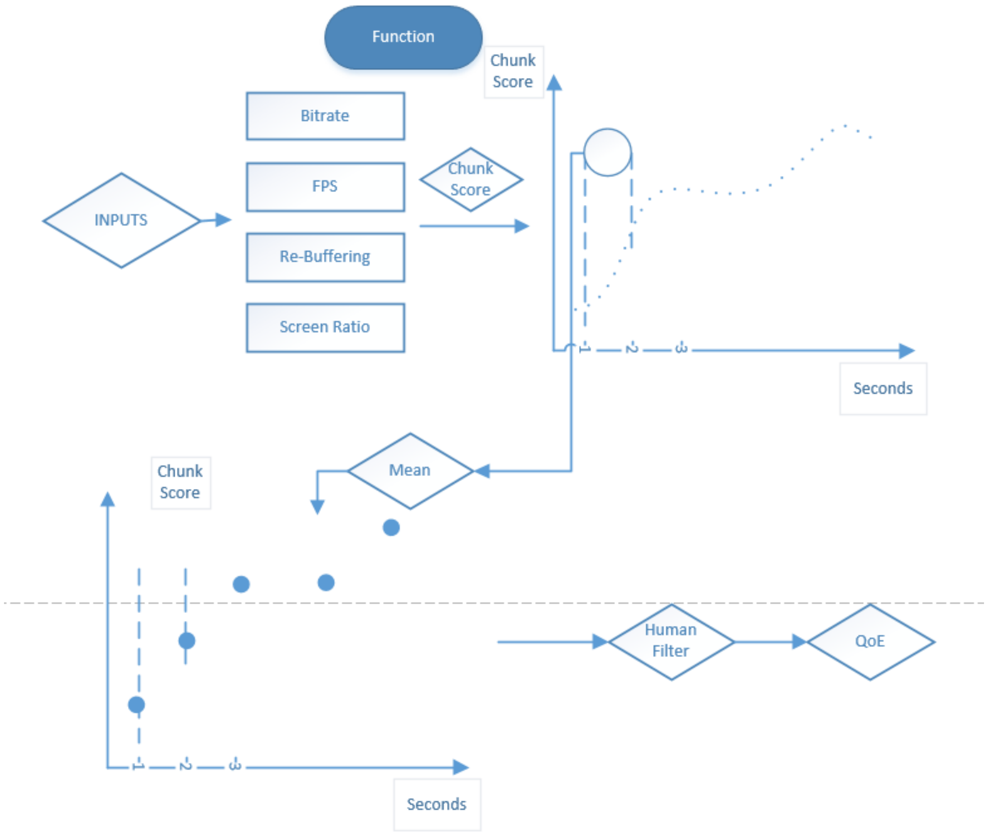

Figure 5.14: Structure method.

5.3.1 Test Scenarios

The logical structure of the test scenarios is crucial to obtain the values of quality of experience, in order to allow sequences and characteristics related to the implemented method in cases of more complex streaming videos. Thus, the existence of 20 videos with variable time (5 minutes), allows the user to not be bored when watching them and to not be influenced by the previously watched videos. The strategy of implementation of each scenario will be explained individually. This method of subjective assessment will have the purposed of verifying the objective method and give a final calibration. The Table 5.3.1 depicts the characteristics of the tests, but they are going to be presented in more detail, individually.

There were created tests in order to follow the time of the recommendations made by MPEG (ITU) [24]. Note that, the table 5.3.1 presents the expected value of QoE that was calculated using the method implemented in this dissertation. The chapter 6 will compare the expected values with the obtained ones. Two videos were selected to create the scenarios, and their selection was based on scenes with different natures and complexities: the category of sport (test scenario 1 to 15) and the category of animation (test scenario 15 to 20).

id	Bit-rate	FPS	Rebuffering	Resolution	Exp. Value
	[bps]	[Frames/s]	[Seconds]	WxH pixels	QoE
1	5Mbps	25	0	1280x720	4.8
2	1Mbps	15	0	1280x720	3.202
3	500Kbps	2	2	360x240	1.027
4	1Mbps	15	0	480x360	2.308
5	1Mbps	25	0	1280x720	3.547
6	200K/500K/1M/2.1M	10/20/25	1	1280x720	1.893
7	5Mbps	25	1	720x480	3.222
8	1Mbps	15	2.5	480x360	3.642
9	500Kbps	25	0	1280x720	2.733
10	500Kbps	25	5.5	1280x720	1.121
11	5Mbps	25	0	1280x720	4.8
12	1Mbps	25	3.5	480x360	1.398
13	500Kbps	25	1.23	480x360	2.09
14	500K/1M/5Mbps	25	0	480x360	3.63
15	200K/2.1Mbps	10/20	2	720x480	4.142
16	500Kbps	20	0	1280x720	2.733
17	50Kbps	15	0	720x480	2.436
18	50Kbps	5	0	720x480	1.728
19	50Kbps	5	0	720x480	1
20	50Kbps	15	0	1280x720	1.769

Table 5.2: Tests Scenarios in the survey

- Scenario 1

  This first example is designed to be an initial video with a very good quality, so the user can create a reference to the original video. This reference is determinant to the classification of the following videos, because, the user is mentally comparing the features with the original one. This video has a rate of 5 Mbps, 25 FPS and a resolution of 1280x720p.

- Scenario 2

  In this scenario, it will be decremented two metrics of video to analyse the user rating. In this scenario the characteristics of the video designed with a medium quality with a bit rate of 1 Mbps, 15 FPS with a resolution of 1280x720p.

- Scenario 3

  This third scenario is designed for streaming with low quality of experience for the user to realize the worst quality. It is used 2 seconds of re-buffering, a smaller screen (360x240p), a low bit-rate value, about 500 Kbps, and 2 FPS.

- Scenario 4

  This scenario aims to test the ratio of screen resolution: it is used the second scenario, but with the decrement of the window size to 480x360p in order to understand if it affects the users experience.

- Scenario 5

  The metric FPS is very important for the quality of user experience. In this scenario, the influence of FPS will be tested. This scenario is created to improve the second one. The FPS value will be increased to 25 FPS (the rest of the video characteristics are the same) in order to understand the changes in users experience.

- Scenario 6

  In this case we will present a scenario with quality changes. Initially, the user will have to wait 1 second for the re-buffering, then he begins to play the minimum quality (200 Kbps, 10 FPS and 1280x720p) for 2 seconds, then it is raised the quality to 500 Kbps, with 10 FPS and a resolution of 1280x720p also during 2 seconds. Then it is raised the bit-rate to 1 Mbps (20 FPS, 1280x720p) with the same time. After those 2 seconds, bit-rate reaches a peak of 2.1 Mbps (25 FPS, 1280x720p) for 2 seconds again. Later, it ends with a 4 seconds of video that presents a bit-rate of 1 Mbps (20 FPS, 1280x720p).

- Scenario 7

  In this scenario, the video used is the same as the first scenario, but the window size is changed to 720x480p, and it is added 1 second of re-buffering time. This scenario aims to measure the quality of experience with these variations.

- Scenario 8

  In this scenario, the video is the same as in the second scenario, but it is added a 2.5 seconds re-buffering, and the size of the screen is decreased to 480x360p. In this case, the aim is the same as in Scenario 7.

- Scenario 9

  In this scenario, the bit-rate is the same as in the third one (500 Kbps), but FPS is increased to 25, it is added 5 seconds of re-buffer and it is used a window size of 1280x720p. This test scenario aims to check the QoE value with the decrease of bit-rate of 500 Kbps to 1 Mbps.

- Scenario 10

  In this scenario we will determine the results on variations of a high re-buffering value (5.5 seconds).

- Scenario 11

  Scenario aims to assess the user memory. The main objective is to observe the discrepancies of two scenarios, this one and the first scenario that contains the same video.

- Scenario 12

  In scenario 12, the objective is to evaluate the quality of experience in the mobile operator perspective, because it is presented the resolution of a smart-phone (480x360p) and network congestion that affects the time of re-buffering, there are the characteristics that affect the mobile network. The other scenarios contain better resolutions, such as the HD (1280x720p) and the SD (720x480p).

- Scenario 13

  This scenario presents variations in all parameters, except the bit-rate, like the third scenario. This is characterized by having bit rate of 500 Kbps, 25 fps and size of 480x360p window, with a re-buffering of two seconds in 1.23 seconds and at 2.30 seconds.

- Scenario 14

  In this case, the bit-rate starts with 500 Kbps for 2 seconds, then increases to 1 Mbps for 2 seconds, ending it up the last 2 seconds with 5 Mbps. The FPS does not change, it is fixed at 20.

- Scenario 15

  This scenario presents variations in the bit-rate and FPS over the time, it is the case that start with a low quality and then increases to a medium quality. Initially there is a re-buffering of 2 seconds. In this case the bit-rate starts with a 200 Kbps (10 FPS) for 2 seconds, and then increases to 2.1 Mbps (20 FPS) for the next 2 seconds.

- Scenario 16

  In this case it is only changed the window size. The objective of this scenario is to understand the changes in users experience with the variations of the screen resolution. In this case initially the video is played in 1024x720p for 2 seconds, then switched it to 1920x1080p for 2 seconds, after 2 seconds on 1024x720p.

- Scenario 17

  This scenario presents a very low bit-rate (50 Kbps) in order to check whether there is influence on playing animation videos with so low quality (FPS = 15).

- Scenario 18

  This is similar to the 17th video, but it is decremented the FPS to 5 FPS.

- Scenario 19

  This is similar to the 18th video, but with a re-buffering of 5 seconds.

- Scenario 20

  This is the last scene, and it uses the 17th video's scenario increasing the size of the window to 1280x720p.

5.3.2 WEB Page

The page 5.15, hosted in our server group is characterized by having an interface with twenty videos and, below each one, a scale of rating quality (MOS scale 1-5) is displayed. This page contains videos uploaded on Youtube with different characteristics. One test was done on Youtube to check if the video quality was changed. After the video file being uploaded, we downloaded it to a local folder on a computer to analyse if the characteristics were changed. The struct on the web page is dynamic to the user, thus Hypertext Preprocessor (PHP) was used to create the MOS scale buttons to allow the user to classify the videos. Later it was necessary to submit the data to a database. MySQL was used to store the data of each survey response.

So the different languages to create this resource are as follows:

- HTTP to create the base structure;
- Cascading Style Sheets (CSS) to create the visual style, colors, and fonts;
- PHP to create buttons and selection that will subsequently read the data.
- MySQL to store the answers in a database.

The videos were loaded in the same web page, based on iFrame element. This allows to reproduce compressed videos archives (AVC/H.264 and AAC). As in this case we only aim to analyse video quality, the audio part will never be analysed, because the video is not influenced by the audio.

The structure of the web page can be seen in the figure 5.15. The header contains a presentation of this survey, and below it contains the videos with different scenarios. There is a bar evaluation for the user to choose the value of the video quality, and at the end, the page contains a button to submit the results.

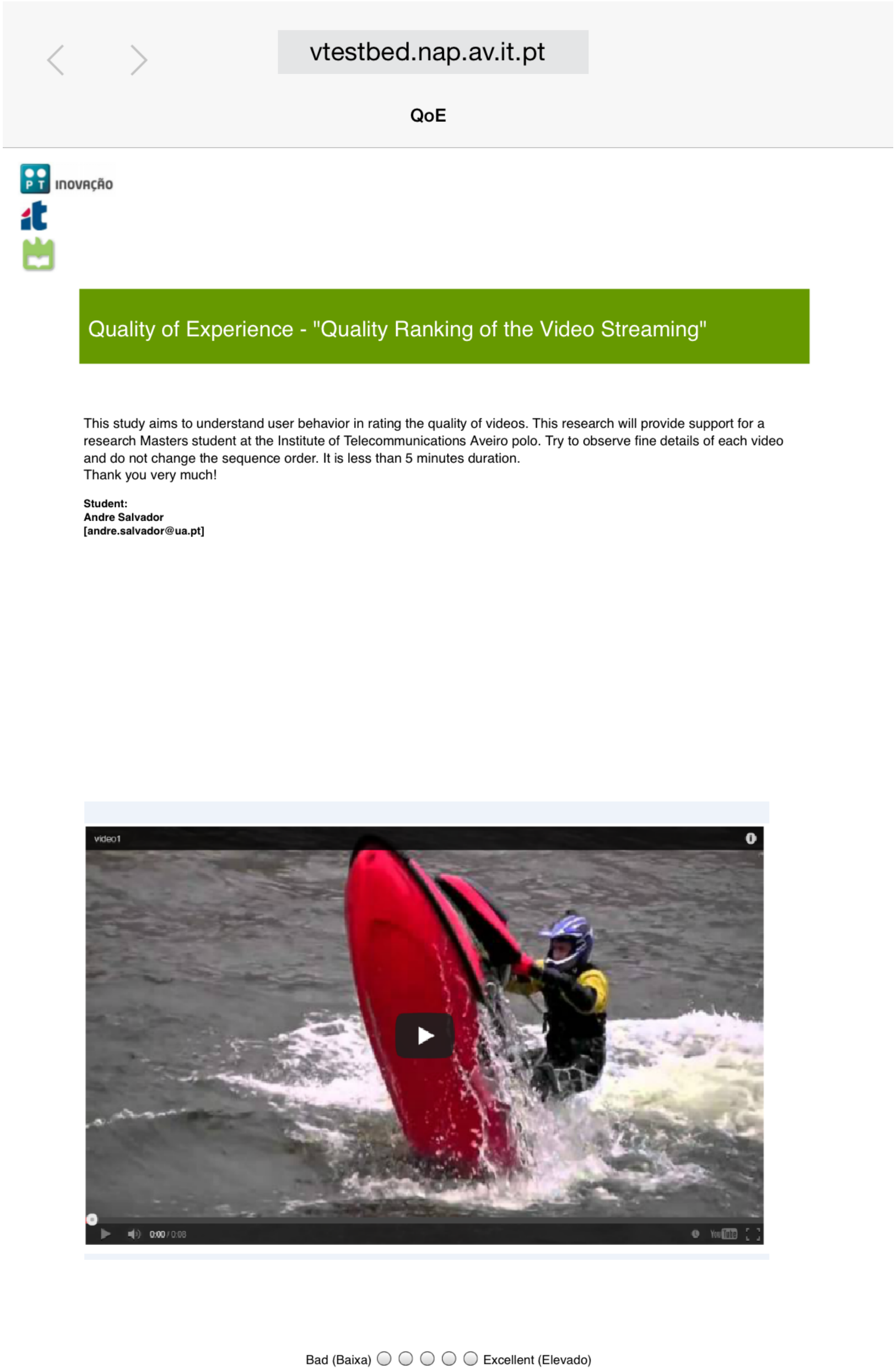

Figure 5.15: Survey - WEB Page.

5.4 Conclusions

In this chapter it was presented the implementation of the solution that was proposed for monitoring the transmission of adaptive video service in the Meo Go service. The possible implementation into their system requires an objective method. Thus the solution of implementation starts with the objective method, as mentioned in chapter 4. First, the scored chunk is assessed using the metrics bit-rate, FPS, re-buffering and the ratio of the window. The scored chunk is obtained based on a mathematical expression that depends on these metrics, so it was necessary to understand the behaviour of them on quality and to create a general expression.

The next phase is related to the calibration of the expression that was performed based on a foundation of trained data. The human memory interferes a lot on evaluating the quality of experience, so the value of the score chunk is updated based on a memory filter that was applied to simulate the reality of a streaming session. The objective of the trained method (with a database) was checked with a subjective method. The solution obtained was more reliable using this method. Thus, in a second phase, a survey of 20 test scenarios were conducted to prove the implemented method. The ITU recommends more than 50 people to obtain confident results, so, for the purpose of improving the way of acquiring the responses, we designed an on-line platform for anonymous users answer.

Chapter 6

Tests & Results

6.1 Introduction

This chapter contains the most relevant results acquired in this dissertation. After implementing the method for obtaining the quality of experience by the user, there is the need to test and verify the operation. The chapter contains also results related to the behaviour of the smooth streaming player. Thus, this chapter can be divided into two parts which will be shown below:

- **Section 6.2**

  The architecture of the player is the base of this method to get the quality of experience from the user. So in order to check the operation of the player, several tests were performed with videos in the networks with different conditions. These tests were expanded to VOD and VOL and their behaviour will be analysed. PT Comunicações operator aims to have communications services with lowest time zapping possible; then, in this section it is created testing scenarios to calculate the zapping time. The type of channel coding can provide better or worse compression ratios; then this behaviour will be analysed.

- **Section 6.3**

  After the method has been implemented, it is necessary to test and verify the expected values of quality of experience. This section will describe tests to demonstrate the operation of the method, both objective and subjective.

6.2 Scenarios in the operation of the player

6.2.1 Test Scenarios

In this section, it will be tested the behaviour of the player. Different testing scenarios are going to be created, with congestion, delay, limited bandwidth and losses for state validation of the heuristics.

6.2.1.1 Network Profiles

Table 6.1 contains profiles of networks which are subsequently used in the testing scenarios. Network Emulator for Windows Toolkit program was used to emulate networks with different characteristics. This program allows to create network profiles that allow to control the laptop interface with the real network. We note that, as was used a multimedia web server (requires internet access to the Internet network), the Ethernet connection to decrease the influence of the access network.

Profile	BW (Dw/Up) [bps]	Latency [ms]	Loss
56 K	56K/33.6K	Fixed (250)	10^{-06}
ADSL	512K/128K	Fixed (50)	10^{-06}
GPRS	35K/10K	Fixed (750)	~ 0.02
CDMA2000	2.4M/384K	Fixed (100)	~ 0.001
WCDMA	28M/11.5M	Fixed (60)	~ 0.001
IEEE802.11b	300K/300K	Fixed (10)	~ 0.001
IEEE802.11b	11M/11M	Fixed (10)	~ 0.001
IEEE802.11b	54M/54M	Fixed (10)	~ 0.001

Table 6.1: Network Profiles

6.2.1.2 Dedicated hosting service without traffic

The dedicated server is used to host the adaptive multimedia streaming of the providers. This server contains a 720p (HD) and the video file will be transmitted in various bit-rates that depend on network and device performances. As mentioned in the previous subsection, the network access interfere with the results, so a PING test was done to check the performance.

Video on Demand (VOD)

The server on Demand allows to access the storage video. In this case, the VOD server used has a *Big Buck Bunny* video in the http://video3.smoothhd.com.edgesuite.net/ server with the 35 ms of PING.

Bit-rate

We include tests with differents bit-rates to check the expected funding of the system. As predicted, a connection of 56K can not playback the stream of the video, because there is not enough bandwidth (Figure 6.1). Also, It is possible to observe in the Figure 6.1 that when the beginning of the play starts with the lowest quality, and subsequently, it rises at intervals. With 10 seconds, it is possible to have a high quality, because the these systems of transmissions have one fast adaptation to the network.

As the scenario presents no traffic, the networks achieve better performances and this results in quite good quality. The case of congestion will be treated in a Section 6.2.1.3 of this chapter.

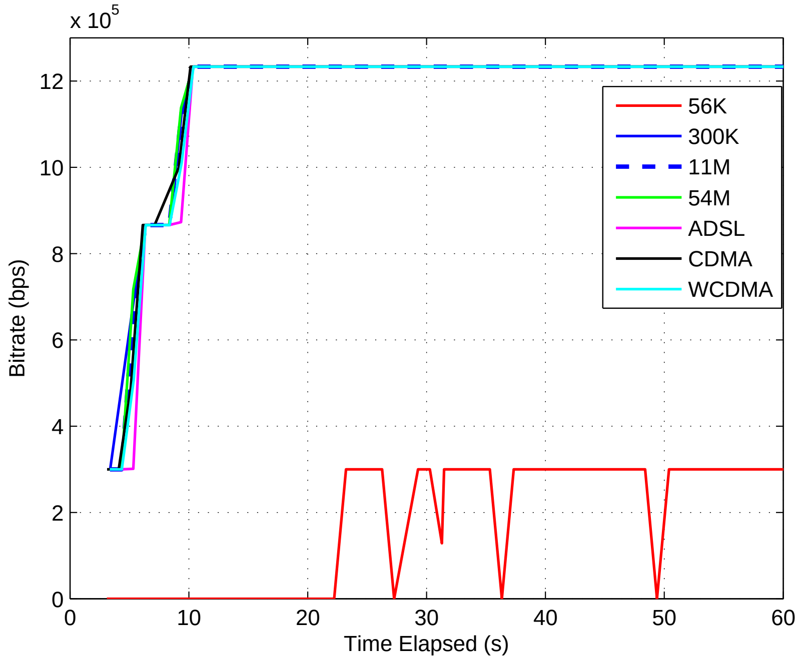

Figure 6.1: Bit-rate VOD.

Perceived Bandwidth

The player tries to realize the available bandwidth. This calculation, as mentioned in Chapter 3, is based on the time that takes between the requests and the size of the request. It can be seen that there are three groups in the graph (Figure 6.2): the 56K network is the weakest, the broadband mobile networks (CDMA, WCDMA) in the middle and finally the faster networks (IEEE 802.11g). In the case of mobile networks, the player perceives less bandwidth because the mobile networks exhibits more delay; then the player makes an estimate with lower bandwidth. This estimation of the player also depends heavily on the network status; in the case there is no traffic. The case scenarios uses a network that is almost ideal, because it use only the characteristics of the physical layer (without background traffic).

Latency

The latency is an important characteristic of the network, and if the value is high enough, it may compromise the network performance. Latency is a metric that can more to vary, it depends of the network congestion and performance. In this case, it is possible to observe in the Figure 6.3 that the mobile networks (CDMA & WCDMA) and ADSL are those with

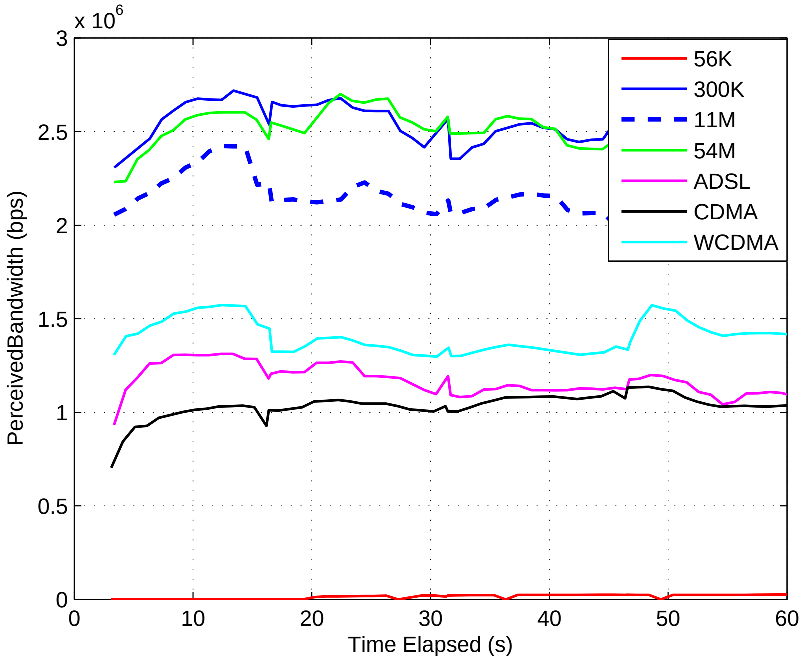

Figure 6.2: Perceived Bandwith VOD.

more latency. The 56k is not considered, because the value of latency was high in the scale of the graph. This latency is influenced by the 35ms ping to the server.

Video on Live (VOL)

Video on live is similar to live streaming, but have small differences, such as it will be possible to observe. This type of streaming is used by Meo Go service. The service allows live streaming and to see the oldest contents of the streaming video.

Bit-rate

The VOL requires the best performances, as it can be seen in Figure 6.4. Note that only two connection(profiles) can transmit streaming video with good quality. The others are trying to link up, but as there is more loss, delay and lower bandwidth it will decrease their bit-rate.

Perceived Bandwidth

The bit-rate has only two profiles that are able to provide performance. Here the same happens, as the player realizes that it is well suited to transmit high qualities. Figure 6.5 presents the increased bandwidth (54Mbps). This shows that the player does not know well the bandwidth, but it makes a good estimate, because when it is most needed by the network, it recognizes the best solutions in the access network.

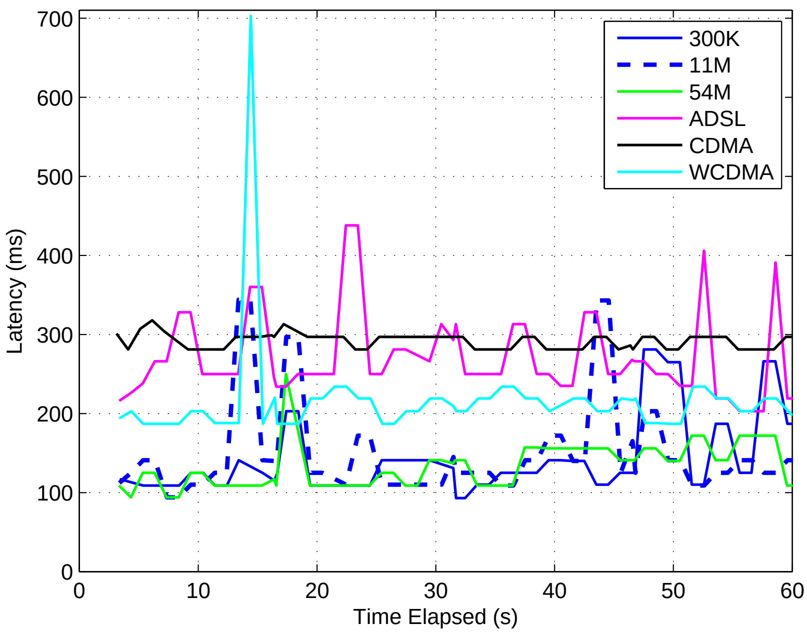

Figure 6.3: Latency - VOD.

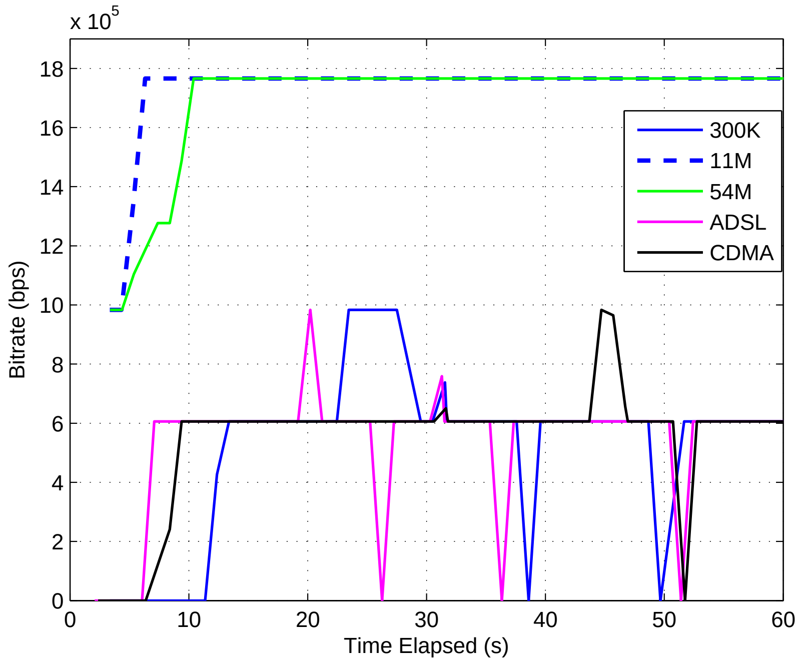

Figure 6.4: Bit-rate - VOL.

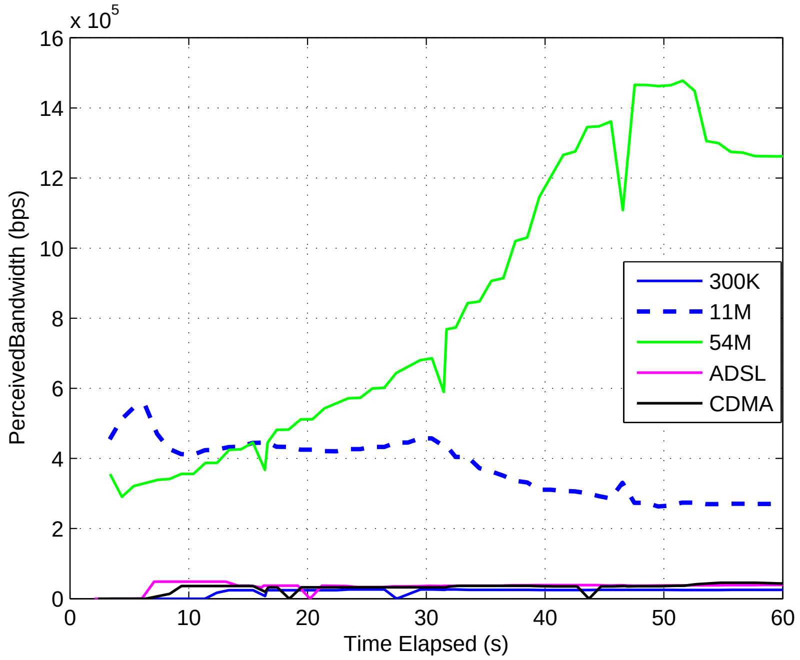

Figure 6.5: Perceived Bandwith - VOL.

Latency
Since the latency is affected by the network server and the status of the test computer, it is higher than expected. It is possible to observe in Figure 6.6, that again the two best networks continue to have the best values. We can observe a peak of 10 seconds because at that time the network is 11 Mbps with a peak bit-rate, as it is possible to observe in Figure 6.4.

Buffer Size
The buffer size in both VOD and VOL is shown to verify the differences in the sizes of the video buffers. As can be seen in the Figure 6.7, it is possible to verify that the VOD buffer fills up to avoid most impact in the network. But the VOL does not have a good behaviour, because the video files can not be ready in the server (Live Stream), so the data in the buffer is lower than the VOD. The buffer engine, with high capacity, allows to decrease the fails in the network. The management of the buffer is already prepared for the 2 scenarios. It is noteworthy that, in the service of Meo Go, the network is a factor that will affect the performance because we will consider a VOL.

6.2.1.3 Dedicated hosting service with traffic

This section considers the existence of background traffic. In the real network, the background traffic influences the network behaviour. In this case, the Figure 6.8 shows the test scenario with background traffic in the Smooth Streaming. The traffic was generated

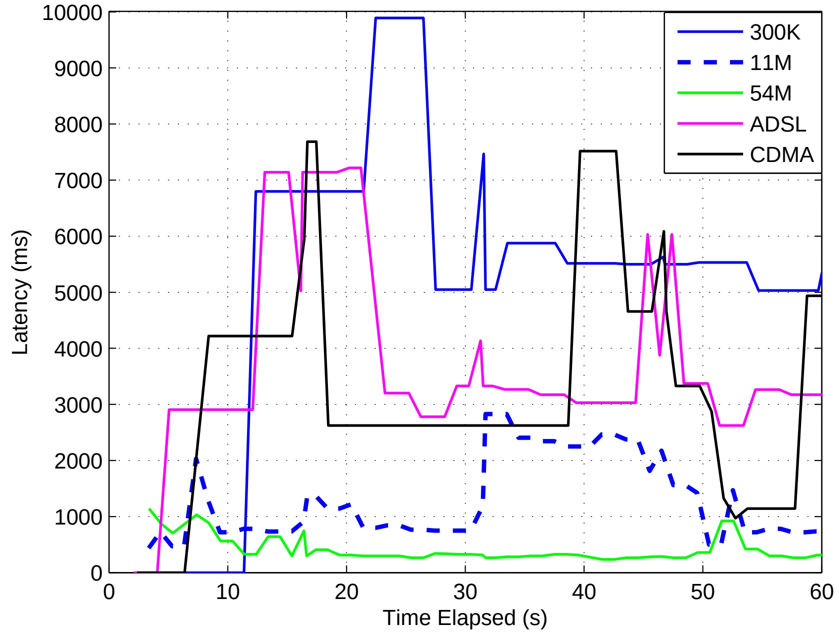

Figure 6.6: Latency - VOL.

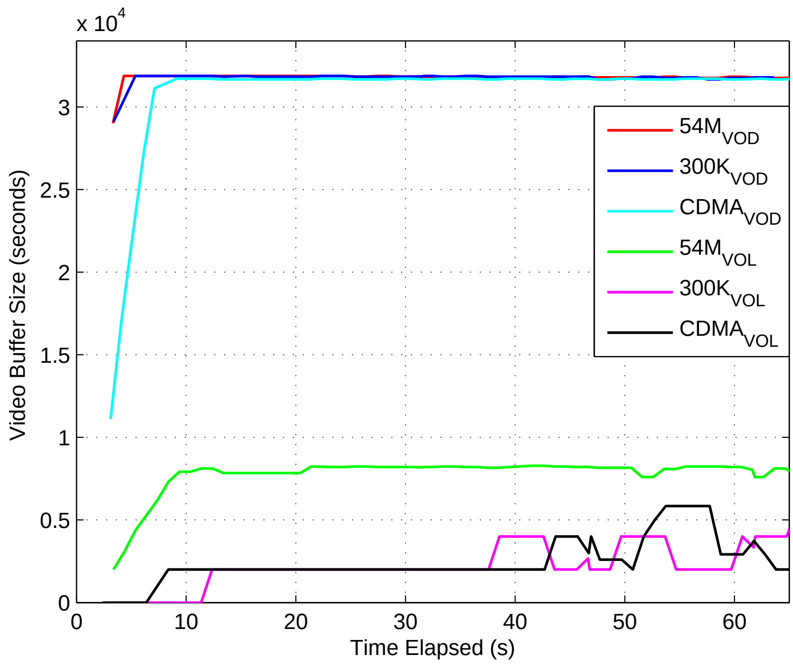

Figure 6.7: Video Buffer.

with the Network Emulator for Windows Toolkit and it is simulated with a *Pareto* distribution. This section presents two different scenarios. In the first one we used a network with a 54 Mbps bandwidth, we insert a background traffic in order to decrease the player performance, so we could verify the stability of the quality changes over time. In the second, we used a network with a 300 Kbps bandwidth, but in this cases it was simulated the quality performance with the traffic background between of the 250 Kbps and the 300 Kbps.

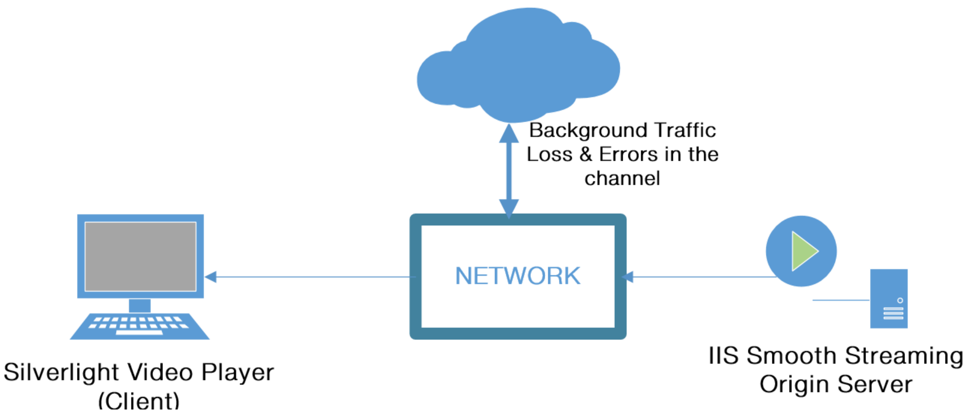

Figure 6.8: Scenario with traffic in the network.

Traffic Background

- 53.9Mbps of Traffic Background in network with 54Mbps of Bandwidth

  The traffic is simulated with *Pareto* distribution. As it is not fixed, then one can not say that there is a bandwidth of 100 Kbps. It is noted that the value of the traffic is not fixed, because their distribution varies and the network can have different levels of the occupation. In the Figure 6.9, the curve of the perceived bandwidth presents this effect with the maximum extremes of this value, in the 150 and 370 seconds. Then it is possible to observe that the bit-rate always tries to climb the minimum, but it can not do so. Sometimes it makes re-buffering, as can be seen near the 300 seconds. The buffer size varies with its network status. The bandwidth that aims to estimate varies because the network is not stable (Figure 6.9).

- Differences between 250 Kbps and 300 Kbps of Traffic Background in the network with 300 Kbps of Bandwidth

  In this phase, the tests are performed in a network with 300 Kbps, but with different traffics to verify the characteristics of the player with various congestions.

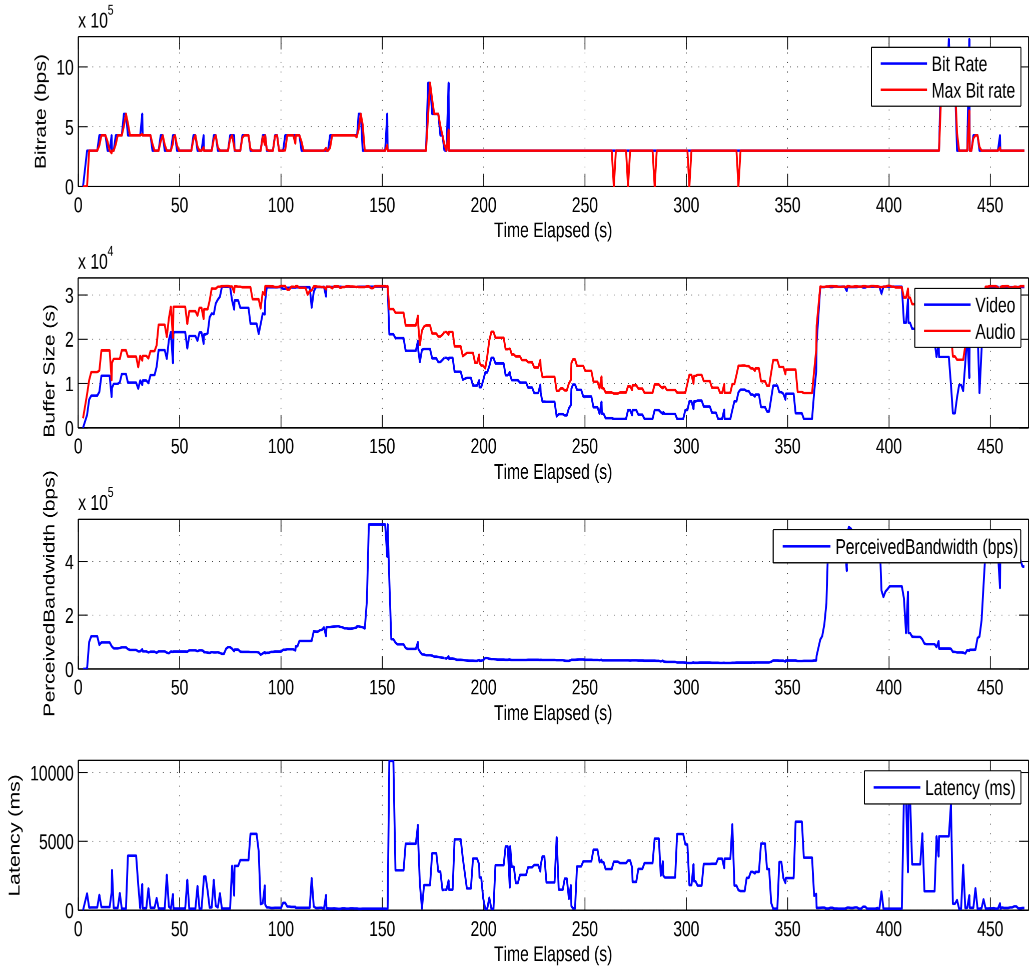

Figure 6.9: IEEE802.11g (54Mbps) with 53.9Mbps of Traffic.

In traffic of 250 kbps (Figure 6.10), it still tries the 600kbps bit-rate, but has no success and have to use 300kbps. After the saturated network (Figure 6.11), the player can try and reproduce the minimum quality and with some re-bufferings, because we want a value close to real traffic, so we used this distribution.
The buffer of Figure 6.10 grew initially as the bit-rate increased, during 20 seconds, to the peak with maximum bit-rate. Then, the buffer has a decrement, after there is a decrement in quality (the player lowered his quality). The player keeps the bit-rate of 608 Kbps more than 20 seconds.
After 50 seconds of time elapsed, the buffer grows the size of audio and video, because the video streaming contains a worst quality. The player has one metric (Switching Frequency), this metric suspends tracks that tries to increase the quality and it is

unsuccessful.

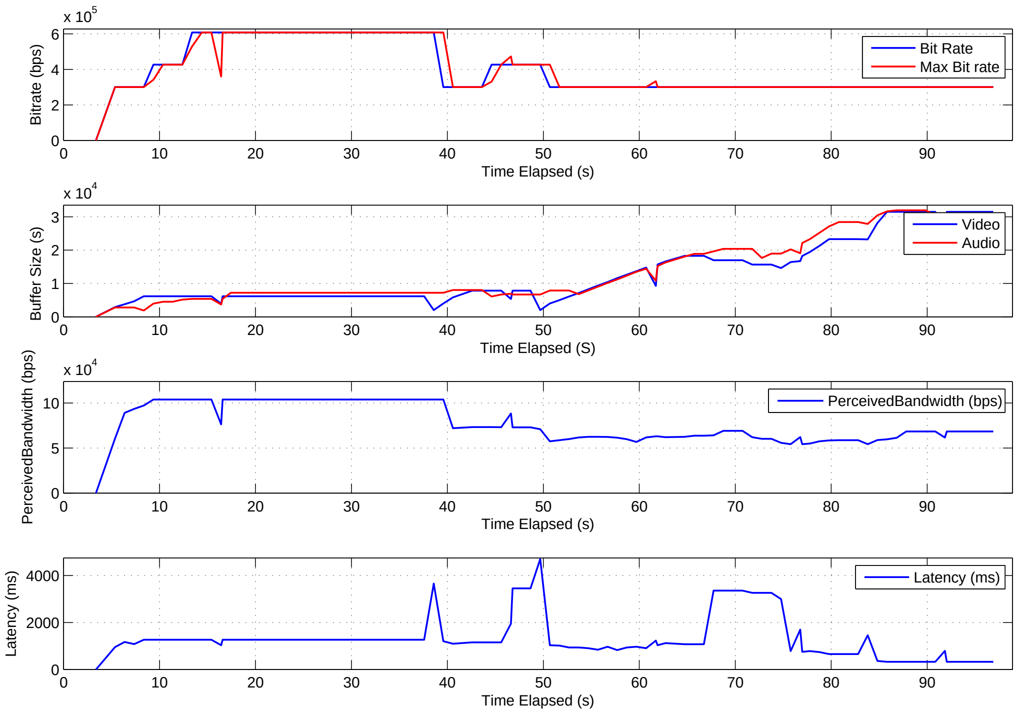

Figure 6.10: IEEE802.11g (300 Kbps) with 250Kbps of Traffic

6.2.1.4 Zapping time estimation

One factor that greatly affects quality of service is the time of zapping: when the consumer changes channels, aims at a rapid change and continue watching the next channel. Therefore, we include tests to analyse the zapping time. This time is calculated as the time of doing a request to the server to change the channel and until it is changed.

First Scenario

This scenario aims to observe the zapping time for multiple servers in the same section. Figure 6.12 presents the time to get the manifest to analyse its the zapping time, but what influences is the time to start playing. This may depend on the terminal where it is, because it is an issue of decompressing the video and start playing. The time has manifested this term because it is necessary to obtain and read the XML structure, so it is slower to get the chunk. The chunk, as is always the beginning with worse quality, makes the size the smallest possible, and therefore, it is be much faster obtaining the chunk.

Figure 6.13 shows more details about the aim to obtain various times zapping. In the case of source 2, it seems to be more unstable, since it attempts to increase the bit-rate

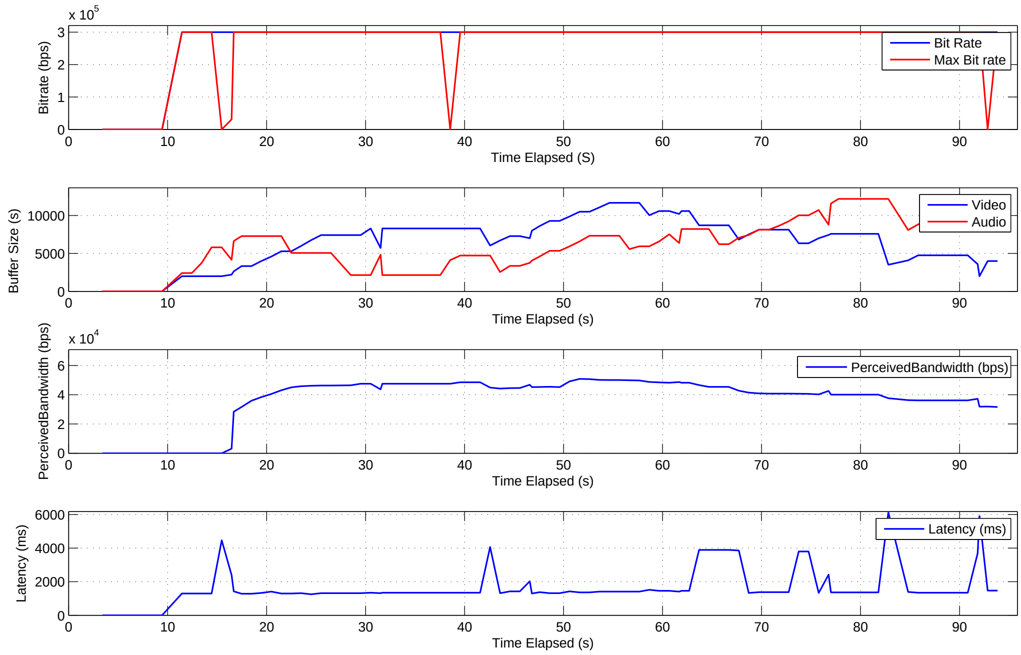

Figure 6.11: IEEE802.11g (300 Kbps) with 300Kbps of Traffic.

---

and it can not. The issue of buffer shows quite similar behaviour. When there is a change in the stream, it is necessary to re-buffering as it is possible to observe.

Regarding the bit-rate, the source 1 increases more easily than others, but the source 4 is there are with a a higher peak. This depends on the qualities that presents the manifest, since it can be scaled and take longer to increase, and on the other hand gives more metrics to take advantage of the best quality of service. To submit a zapping time course, it is necessary to have the minimum bit-rate to start, as it is observed in the source 3 and source 4. Later, it climbs as fast as possible to the peak, and therefore, it can not have too many steps because each step has a period at least 2 seconds.

Second Scenario - (VOD & VOL)

In a second test, we aim to calculate the time of zapping a streaming VOD and VOL. It is aimed to observe the state of the buffer and the performance level.

In the Figure 6.14 it can be observed that the zapping time doubles; it also shows the average variation of the zapping time, around 1.5 seconds. This is not very high, but the ideal value would be below 1 second. In the buffer size it is possible to distinguish the VOD & VOL, because the VOL requires better network performance, as can be seen in Figure 6.15. Also, the player estimates that the VOL bandwidth is smaller, but when playing a stream, we verify a higher bit-rate on VOD. When we have high rates of bit-rate

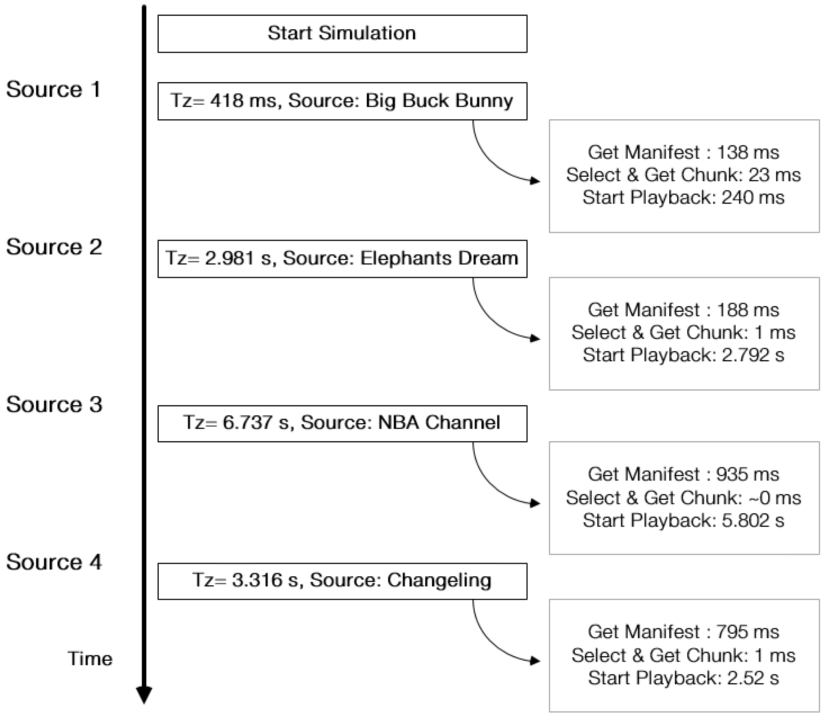

Figure 6.12: Scenario 1 - Zapping time

---

and the chunks (that are downloaded through TCP layer) have a high variation, there may increase the delay and this impacts on the bandwidth calculation. The Meo Go, by using the VOL service, requires the knowledge of the state of the network. This is important because, by knowing several network indicators, the provider (PT) can adjust the stream quality. So, it is necessary to improve the quality of service to increase the quality of experience.

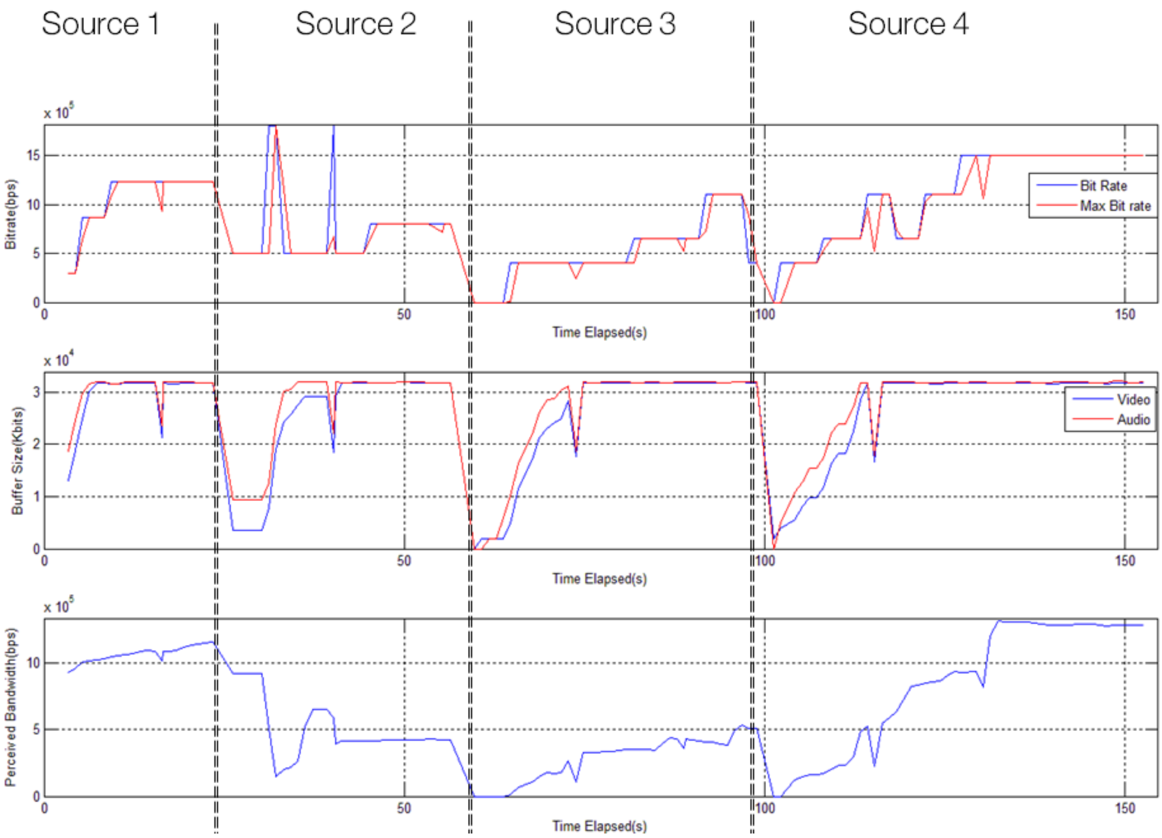

Figure 6.13: Scenario 1 - Zapping time (Metrics).

6.2.1.5 Chunk Size with the channel type

As mentioned in Chapter 4, when there is an assessment of QoE, the type of channel, as Sports, Information, Generalist, Kids, Movies and Documentaries, influence the compression of the videos. The H.264 (AVC) codec compresses the video by using the change of pixels between frames: if the video is slower than the average the compression is more efficient (smaller size [15]). The chunks size is influenced by the behaviour of several types of channels. It was tested this influence in the channels (47 channels ([11])) of the Meo Go server (*live-picoas.online.meo.pt*) during 48 hours.

The Table 6.2.1.5 contains the number of channels used in the experiments (the channels used were those served at Meo Go package without the premium channels).

In the test scenario, the chunks were obtained with the higher bit-rate (2.1 Mbps) to have the maximum size. The behaviour of sports channels are those with larger variations in size chunks. The movie and generalists channels have the lower variations and may be due to have similar characteristics of contents. The channel categories information, kids and documentary channels, have almost the same range of standard deviation. These tests were made at intervals of 2 hours, by getting an average value.

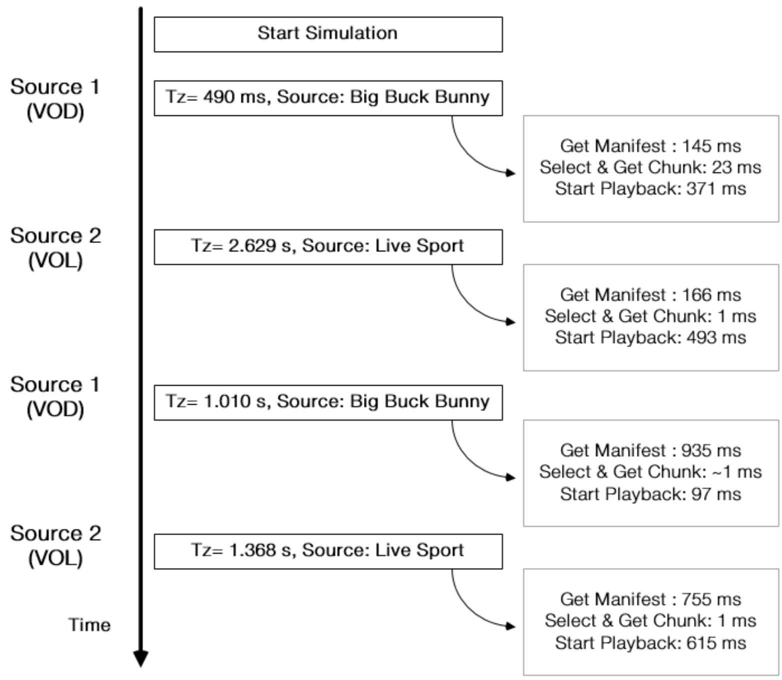

Figure 6.14: Scenario 2 - Zapping time(VOD & VOL).

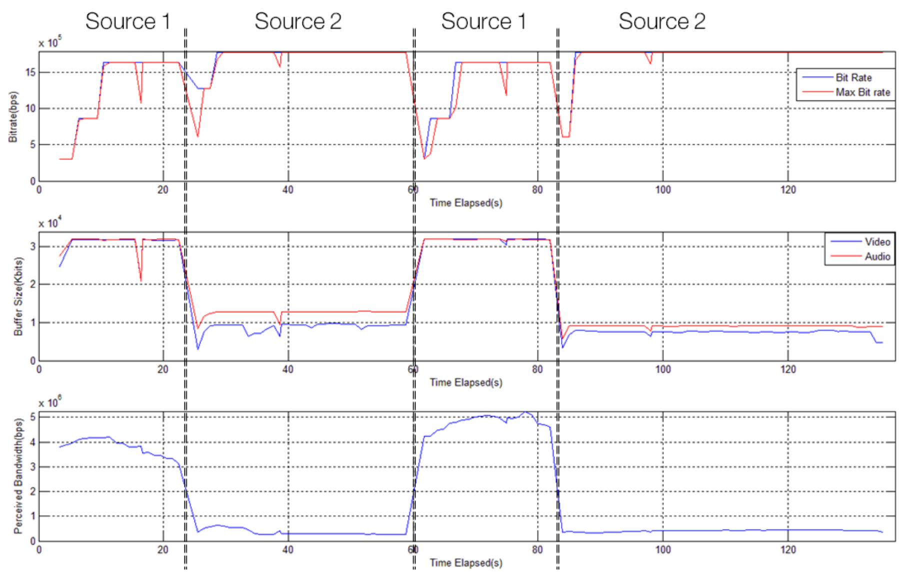

Figure 6.15: Scenario 2 - Zapping time(VOD & VOL)(Metrics).

It is possible to verify that the contents of the channels are important for assessing the quality of experience. The sports channels have greater realism, and therefore, a higher rate of frames per second. Information channels and kids have contents that with a lower rate of frames per second.

Number of Channels	Type
4	Generalist
7	Information
6	Sports
7	Kids
10	Movies
8	Documentaries
5	Music

Table 6.2: Number of Channels

This type of assessment can not be fully effective only by the size of the chunks, because the compression method may not be linear. This research aims to approximate the behaviour with measurements of long sessions, and the results are as expected. It is possible to relate the values of the chunk sizes (2 seconds at 2.1Mbps), with the type of the channel in the Figure 6.16.

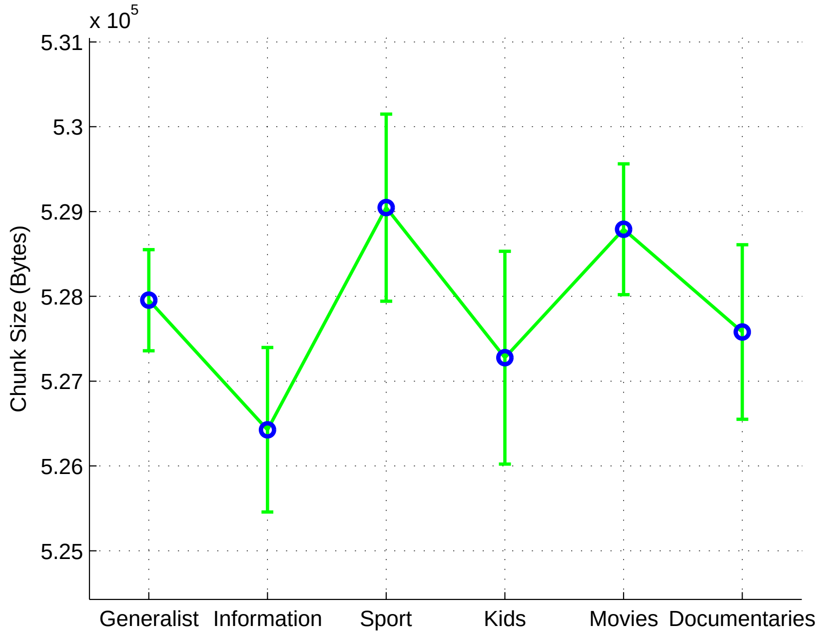

Figure 6.16: Chunk Size with the channel type.

6.3 Test Scenarios for the method of QoE

To analyse the proposed QoE evaluation method, we initially tests scenarios that were created in PEVq, where a section of video change bit-rate, FPS, Re-buffering and the ratio between the screen size and native resolution of the track. These scenarios are classified in PEVq software that are subsequently compared with the method implemented in this dissertation. These scenarios have to be close to reality because the method is planed to be implemented on the TV streaming service, that nowadays can be found away from home (ex.:riding a train, in the restaurant, among other places with a internet connection). Even at home, the conditions of the network may not be the best, because the Meo Go service is used usually away from the box (include hotspot WIFI).

6.3.1 Simulate Scenario

This scenario aims to verify if the method responds well to the worst cases scenarios. These simulated scenarios will be very important to observe the behaviour of the expected filter of human memory. This type of evaluation depends on the condition of the person, but on average it is expected that the filter is quite similar. In these scenarios, streaming sequences have to have many variations, so it is possible to test the majority of the effects of the filter of the human memory.

- Scenario 1 - Low quality of the network

  In this first scenario (Figure 6.17), we tested the method in very unfavourable conditions of the network. It features Re-bufferings and never increases from the minimum quality. This scenario was simulated on the PEVq and obtained 1.72 (MOS scale from 1 to 5). Using the function that was created with the method in MATLAB, this presented a result of 1.53. This scenario happens a lot when the network is heavily congested or the access network is weak.

- Scenario 2 - Increase quality of the network

  In this second scenario (Figure 6.18), we tested the method under conditions where the network fluctuates; in other words, it presents poor as well as excellent conditions in the network. It features Re-bufferings and once increases from the minimum quality. This scenario was simulated and obtained a 2.7 in the PEVq (MOS scale from 1 to 5). Using the function that was created with the method in MATLAB, this presented a result of 2.53. Compared to the previous case, one can observe a large increase in the MOS, because there was a large increase in the last 2 seconds of the bit-rate; although this increase provided great quality (Score chunk = 4.9), the quality of experience just have the value of 2.7: this shows that the memory of the consumer is a very relevant factor.

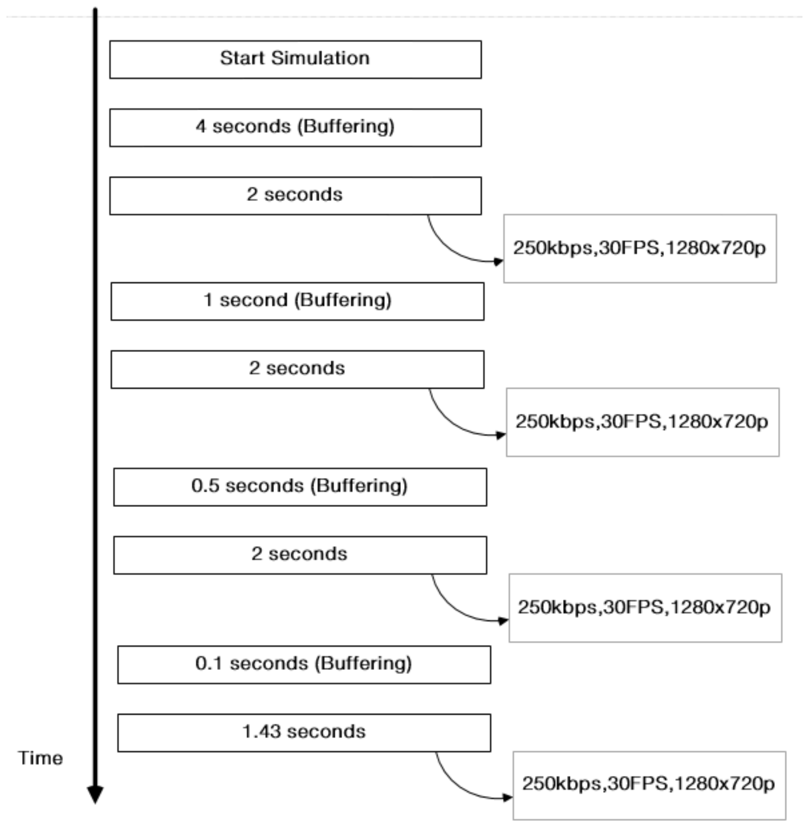

Figure 6.17: Scenario 1 - Low quality of the network.

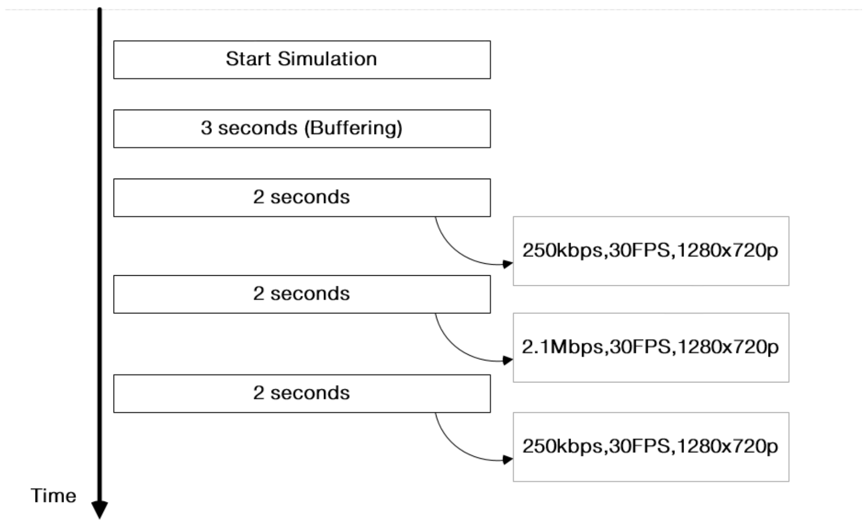

Figure 6.18: Scenario 2 - Low quality & High quality of the network.

- Scenario 3 - Video session with a larger size

  The scenario Figure 6.19 presents a longer session of a streaming, where the network conditions are not favourable. It performs well up to 1 Mbps peak, but then began to affect the network congestion and made the bit-rate reduced. Later, the streaming video is transmitted with low bit-rates and can make re-buffering. It was obtained a value of 1.1 PEVq, while the method gave a value of 1.2172. This type of scenario can occur, for example, when a user is watching TV on the tablet and plans to go to another room where the wireless network is poor [35].

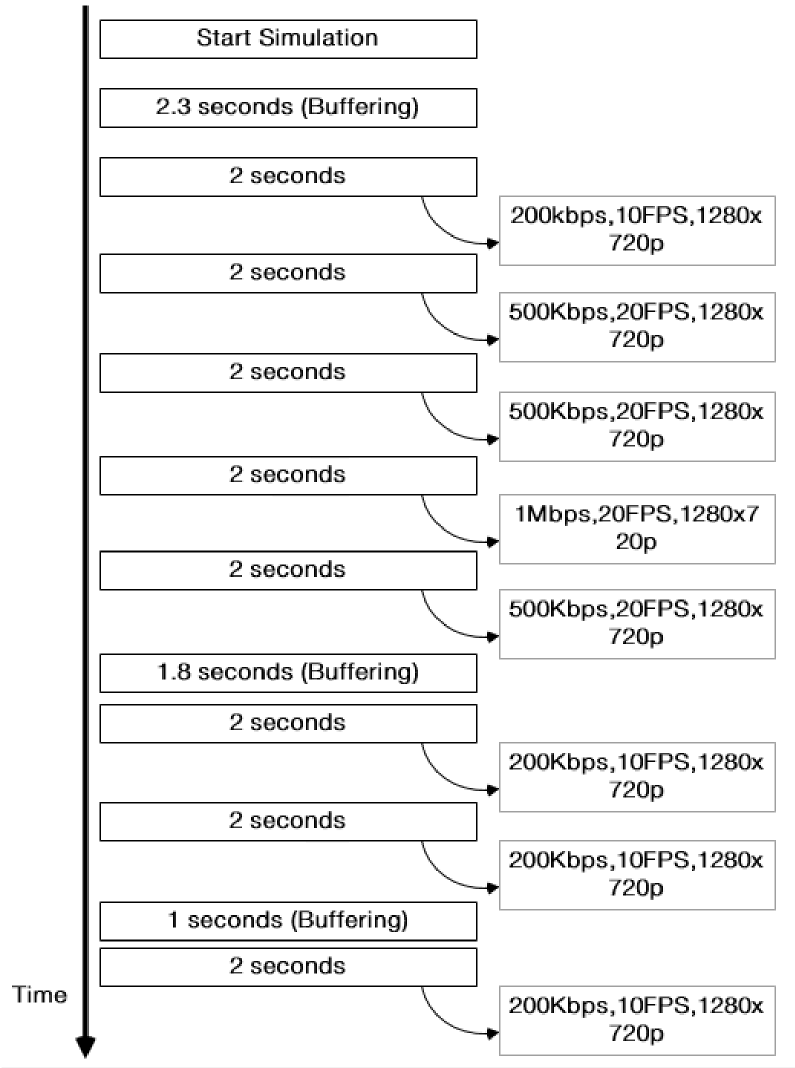

Figure 6.19: Scenario 3 - Video session with a larger size.

In the Figure 6.20, it is observed the convergence of the method that was used, over time. It can be observed that the method adjusts according to the quality of the conditions, by growing initially (which is good conditions). After the peak, it is possible to observe that the method declined and go to meet the final value (at 20 seconds) due to bad conditions (200kbps and 10 fps).

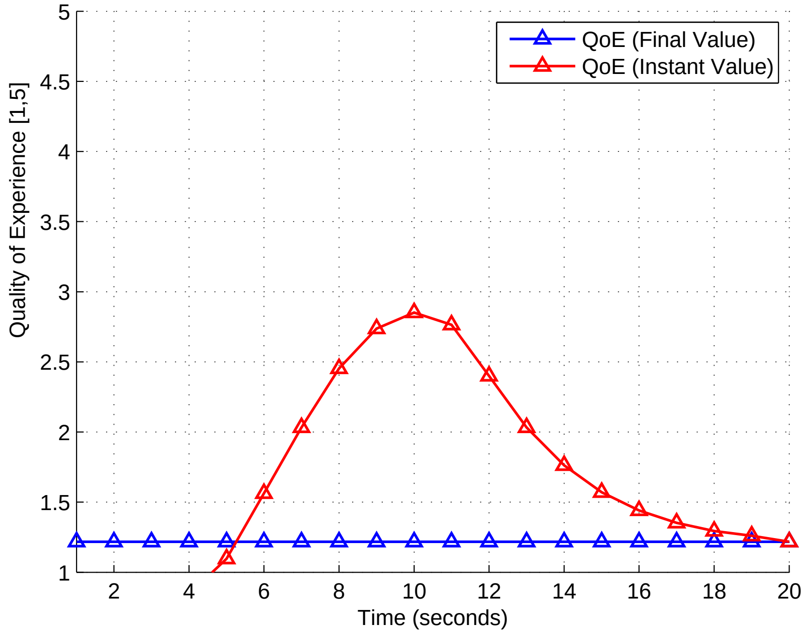

Figure 6.20: Scenario 3 - QoE value with the time (s).

6.3.2 Video Scenario

Stream	Video
1	Big Buck Bunny,1080p [22]
2	Skydive Dubai,2012,4K [21]
3	Jet Ski,720p [52]
4	Elysium Trailer,2013,4K [16]

Table 6.3: Video Streams

The implemented method was calibrated using a database of already trained sequences. At this stage, it is aimed to check whether the method has a good confidence interval (95%) for 4 examples of videos with different categories. The figure 6.21 shows a graph that contains the method and 4 videos (Table 6.3.2) in a range of FPS (5-30) at a fixed bit rate of 1Mbps. The confidence interval of the method is obtained in the implementation by Curve Fitting Tool (CFTOOL) function of MATLAB software: This range increases, on average, with increasing FPS, caused by the channel categories. For example, it is first necessary to realize that the bit-rate corresponds to an H.264 encoded video, and this bit-rate has a fixed value. In the decoding process, there is a maximum limit of FPS for the video, but the performance of the device can have a lower value of instant FPS. Therefore, if the video shows large movements in the content, then it creates defocus in the video, and it is decreased the quality of the experience. In the Graph, the Stream 1 (animation for kids) contains a high quality value at 20 FPS and the video has less defocus. However, in the Stream 4, the quality is lower than on the other streams (after 15 FPS), then this video has larger defocus. Clearly the results are related to the type of content of the video.

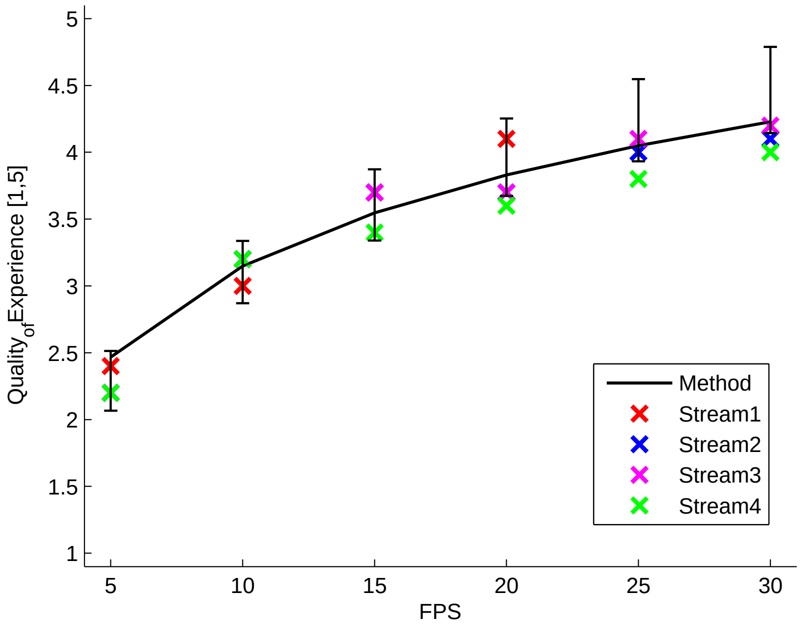

Figure 6.21: FPS variation @ Bit-rate = 1Mbps.

The point of view of the variation of bit-rate is also necessary to prove that the method has good confidence interval, thus the Figure 6.22 presents a test scenario of the method implemented, with four video streams, where the bit-rate variates and the FPS is fixed at 20 FPS. The implemented method has effective confidence intervals, since just two values are not in the range, and these values are related to an animation video. Such high ratio with quality / bit-rate allows the user to classify higher, and so these values are above the expected. However, the values are near to the confidence interval, so the method does not deviate from the expected. In this perspective the values of the videos stream are obtained by a calibrated data base, so they are not, until now, the real values. These real values are presented in the following section.

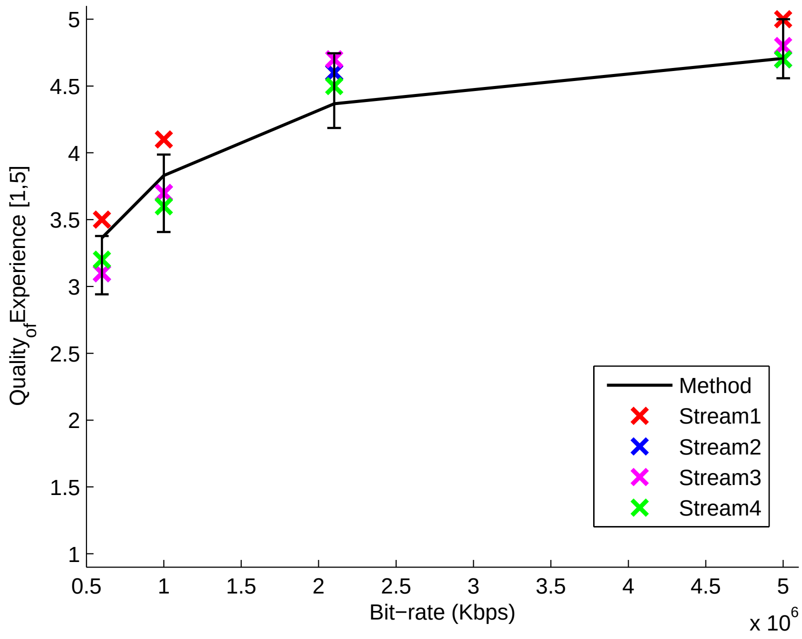

Figure 6.22: Bit-rate variation @ 20 FPS.

6.3.3 Real Scenarios

In the previous chapter, it was presented the subjective method. Following the ITU recommendations, the number responses in the survey must be at least 50 or higher, to achieve a larger convergence and thus increase the confidence level. This subjective method is quite crucial in verifying the implemented method, and in a perspective of subjective tests, it is difficult to get the average higher than 4.5 and less than 1.5, because not everyone gives 5 or 1, a characteristic of MOS tests.

The figure 6.23 shows the results of the video qualities questionnaire (subjective method), showing that the effect of ratings does not have values near the extremes (MOS = 5 or 1). The values obtained in the survey are closer to the video animation file (scenario 15-20 test). This effect may be due to having better compression than the compression made on the sport video (scenarios 1-15): if the bit-rate can be the same, but the quality is different, the sport video has more losses. The values obtained in the method do not deviate much from the measured ones in the questionnaire. In this perspective, it is possible to conclude that the method has a good behaviour when it uses the database trained, since the subjective method does not present results near the extremes.

Note a particular case, the scenario 1 and 11 are the same, but these two are separated by 10 scenarios. In this sense, the user when viewing the scenario 11, he does not remember the scenario 1, and with this perspective, the values of the evaluation may be different.

Note that the recommendation ITU indicates that it must be acquired at least 50 responses, and in this work we considered 64 responses.

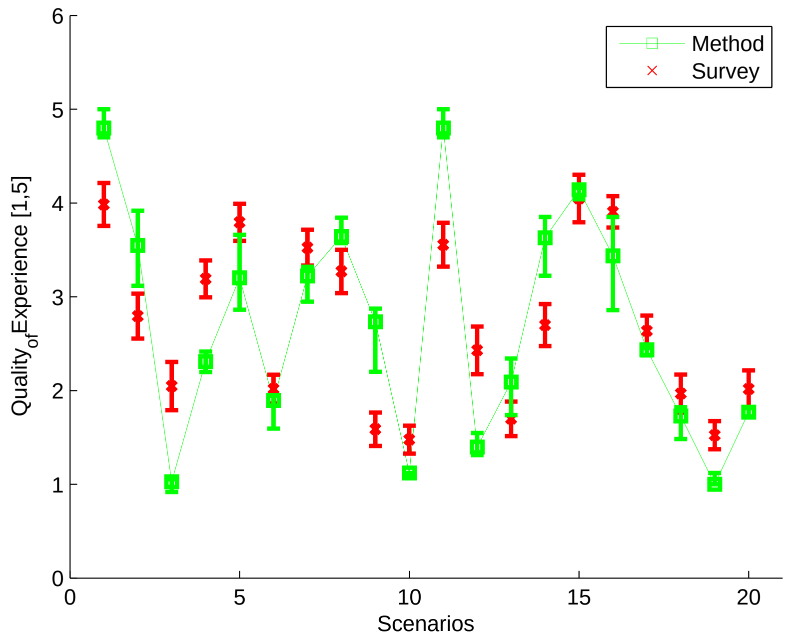

Figure 6.23: Survey with 20 Scores.

6.4 Conclusions

In this chapter, the results of the proposed solution were obtained. The first results are related with the study of the smooth streaming architecture: the test scenarios were built in order to validate the operation of the player. At this stage, the player has been tested with different network profiles, traffic, time zapping and chunk size by the type of channel. The results fall into the expected behaviour of the player described in Chapter 3. In the scenario without traffic, as we predicted, the selection of quality depends only on the profiles created, in particular the 56K profile is not able to get a continuous video streaming because it does not have enough bandwidth. In the case of mobile network profiles, it shows higher latency; this directly affects the player by presenting a lower bandwidth. This effect does not affect the quality of the video stream in the VOD service, but in the VOL service there is a decrease in the video stream quality.

The VOL and VOD services present different characteristics in the network performance. The tested scenarios show that these services have distinct buffer engines: the VOL buffer size is lower than the VOD buffer in order to minimize the latency in live streaming.

In real scenarios the network has background traffic, so we created test scenarios with traffic that affects the adaptative streaming. These results indicate the main characteristic of the smooth streaming player: the player suspends unstable qualities (track) to allow the maximum of the quality of experience, when the network status is affected by the traffic. For example, in the test scenario defined by 300 Kbps of bandwidth and 250 Kbps of traffic (Figure 6.10), after 52 seconds, the quality is limited to 300 kbps of the bit-rate that allows the continuous session of the video streaming.

In a second phase, we preseted results obtained by the objective and subjective methods. Initially, we created one test scenario to verify the influence of the different types of channels. In the sport video, the size of the chunk presents the larger size, but the influence is not significant due to the limitation of the encoded bit-rate. The method of QoE was tested with three different scenarios. The test scenarios were created to simulate real cases. The results allow to confirm the performance of the method implemented, because the test scenarios are characterized by many changes over the time.

Finally, we implemented the subjective method to confirm the behaviour in the real cases. The results of the survey show that the subjective evaluation gets similar results to the objective ones, and that the values present a larger error when they are near the limits (MOS = 5 and 1).

Chapter 7

Conclusions

7.1 Final Conclusions

This dissertation was proposed to develop a QoE monitoring system for the Meo Go streaming service. This method aims to assess the metrics obtained from the Smooth Streaming Player in order to create an ISP streaming QoE rating. This rating was established based not only on the network performance, but it also takes into account the device and the experience of the user.

Initially, the architecture of the player was studied to understand the implemented metrics and their analysis. Subsequently, test scenarios were performed to verify the operation of the player in various types of standard access networks.
In the second phase, it was necessary to understand which metrics translate the QoE of the user, as explained in the Chapter 4. With those metrics, it was possible to directly evaluate the first video analysis using a filter of human memory, that takes into account previous QoE analysis of the stream session that allows an adaptative rating over time.

To validate the performance of the developed method, both objective and subjective tests were performed, which allowed to evaluate, according to the chosen metrics, the quality of experience of the user.

In the objective test, it was performed a comparison between the implemented method and PEVq results over different network quality scenarios and metrics, such as Bit-rate, FPS, Re-buffering time and Screen Size. In this case, from all scenarios tested, the maximum deviation was 0.19 in the MOS scale (from 1 to 5). In addition, the metrics were fixed in order to test the influence of the video content in the method results, that led to conclude that the interval confidence is not exceeded in most of the cases. In a further phase of testing, the results of the objective method were obtained, which showed that the proposed method works even in the worst cases.

In the subjective assessment, a questionnaire was designed to create scenarios of tests in order to compare with the objective method. This questionnaire was created in an internet web page that allowed an automatic submission and an easy video hosting. In the animation video, the results were almost identical, but the sports video led to small discrepancies caused by the lack of identical submissions.

The proposed method has presented good results for evaluating the quality of experience, since it is influenced by features that a user considers important, such as bit-rate, FPS, the screen size and re-buffering. This will lead to better streaming quality and satisfaction from both parts: the provider that uses its infrastructure as efficiently as possible, and the user that gets the better streaming quality possible based on its resources. This matching is greatly improved by the QoE rating feedback that is given to the provider.

7.2 Future Work

This Dissertation presents some areas that are possible to develop and expand in the future. We envision several opportunities for future work:

- The operator can create an application to measure the QoE on mobile devices that allows to collect informations of the streaming sessions. The application will provide a non-intrusive measurement that can increase the quality and also provide more information about the users.
- The adaptative smooth streaming is used to test this architecture to obtain QoE, but there are a lot of other adaptative streaming platforms that can be used to test it; for example, the streaming technology by Apple, Adobe and Youtube.
- With the increase of the video quality, the implemented method will need new ranges of metrics to update QoE. For example, the ultra HD is a new technology that allows resolutions at 3840x2160p; the work in [7] presents experimental results from the analysis of subjective quality assessment on 4K video.
- The 3D streaming has new metrics of the performance of the device to evaluate the quality of experience, and its evaluation is yet only addressed in the work in [63] that introduces some possible methods to evaluate 3D qualities.

Bibliography

[1] 3GPP (Release 12). 3rd generation partnership project; technical specification group services and system aspects; transparent end-to-end packet-switched streaming service (pss). progressive download and dynamic adaptive streaming over http (3gpp-dash) (3gpp 26.247).

[2] 3GPP TS 26.234. Transparent end-to-end packet switched streaming service (pss), protocols and codecs. `http://www.3gpp.org/ftp/specs/html-info/26234.htm`, 2010. [Online; accessed 3-June-2014].

[3] 3GPP TS 26.244. Transparent end-to-end packet switched streaming service (pss), 3gpp file format (3gp). `http://www.3gpp.org/ftp/specs/html-info/26244.htm`, 2010. [Online; accessed 10-June-2014].

[4] 3GPP TS 26.247. Transparent end-to-end packet-switched streaming service (pss); progressive download and dynamic adaptive streaming over http (3gp-dash). `http://www.3gpp.org/DynaReport/26247.htm`, 2014. [Online; accessed 10-June-2014].

[5] N. Amram, Bo Fu, G. Kunzmann, T. Melia, D. Munaretto, S. Randriamasy, B. Sayadi, J. Widmer, and M. Zorzi. Qoe-based transport optimization for video delivery over next generation cellular networks. In *Computers and Communications (ISCC), 2011 IEEE Symposium on*, pages 19–24, June 2011.

[6] S. Arndt, J. Antons, R. Schleicher, S. Moller, and G. Curio. Perception of low-quality videos analyzed by means of electroencephalography. In *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*, pages 284–289, July 2012.

[7] Sung-Ho Bae, Jaeil Kim, Munchurl Kim, Sukhee Cho, and Jin Soo Choi. Assessments of subjective video quality on hevc-encoded 4k-uhd video for beyond-hdtv broadcasting services. *Broadcasting, IEEE Transactions on*, 59(2):209–222, June 2013.

[8] J. Billings. Smpte vc-1 [microsoft windows media]. In *IT to HD: Visions of Broadcasting in the 21st Century, The IEE 2-Day Seminar on (Ref. No. 2004/10760)*, pages 101–110, Nov 2004.

[9] Brandon. Silverlight smooth streaming: Wan optimization or content distribution network. http://brandonpotter.wordpress.com/2010/06/03/silverlight-smooth-streaming-wan-optimization-or-content-\discretionary{-}{}{}distribution-network/, 2010. [Online; accessed 12-June-2014].

[10] R. Bruno, M. Conti, and Enrico Gregori. Throughput analysis and measurements in ieee 802.11 wlans with tcp and udp traffic flows. *Mobile Computing, IEEE Transactions on*, 7(2):171–186, Feb 2008.

[11] PT Comunicações. Meo go. http://meogo.meo.pt/, 2014. [Online; accessed 10-June-2014].

[12] G. Cermak, M. Pinson, and S. Wolf. The relationship among video quality, screen resolution, and bit rate. *IEEE Transactions on Broadcasting*, 57:258–262, 2011.

[13] Chao Chen, Lark Kwon Choi, G de Veciana, C Caramanis, R W Heath, and A C Bovik. Modeling the Time;Varying Subjective Quality of HTTP Video Streams With Rate Adaptations, 2014.

[14] Chao Chen, Lark Kwon Choi, Gustavo De Veciana, Constantine Caramanis, Robert W. Heath, and Alan C. Bovik. A dynamic system model of time-varying subjective quality of video streams over HTTP. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pages 3602–3606, 2013.

[15] Chao Chung Cheng, Tian Sheuan Chang, and Kun B. Lee. An in-place architecture for the deblocking filter in H.264/AVC. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 53:530–534, 2006.

[16] CinemaSauce. Elysium trailer. https://www.youtube.com/watch?v=PcXOnoSNORE, 2013. [Online; accessed 12-June-2014].

[17] CISCO. Mobile multimedia services. http://www.cisco.com/c/en/us/solutions/service-provider/mobile-multimedia-services/index.html, 2014. [Online; accessed 1-June-2014].

[18] Apple company. Http live streaming overview. https://developer.apple.com/library/ios/documentation/NetworkingInternet/Conceptual/StreamingMediaGuide/StreamingMediaGuide.pdf, 2014. [Online; accessed 10-June-2014].

[19] John Crosby. Http dynamic streaming. http://pt.slideshare.net/jccrosby/http-dynamic-streaming-getting-started, 2011. [Online; accessed 10-June-2014].

[20] R. Dubin, O. Hadar, and A. Dvir. The effect of client buffer and mbr consideration on dash adaptation logic. In *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, pages 2178–2183, April 2013.

[21] Dubai Film. Skydive dubai 2012 - 4k. `https://www.youtube.com/watch?v=zFg_mlBFV2c`, 2013. [Online; accessed 4-June-2014].

[22] Blender Foundation. Big buck bunny. `www.bigbuckbunny.org`, 2008. [Online; accessed 6-June-2014].

[23] H. French, Jie Lin, Tung Phan, and A.C. Dalal. Real time video qoe analysis of rtmp streams. In *Performance Computing and Communications Conference (IPCCC), 2011 IEEE 30th International*, pages 1–2, Nov 2011.

[24] Moving Picture Experts Group. Recommendation p.1202 (10/12). `http://www.itu.int/rec/T-REC-P.1202-201210-I/en`, Approved in 2012-10. [Online; accessed 10-June-2014].

[25] Lingyan Hong, Shijue Zheng, Yan Zhang, and Li Tan. Parallel c-means algorithm in human memory meta database search mechanism. In *Web Information Systems and Mining, 2009. WISM 2009. International Conference on*, pages 133–136, Nov 2009.

[26] Pan Hong-jun, Yao Xiao-qiu, Qi Chang-song, and Chen Hong-tao. A category theory model for learning and memory of the human brain. In *Digital Manufacturing and Automation (ICDMA), 2010 International Conference on*, volume 1, pages 11–14, Dec 2010.

[27] Adobe Systems Incorporated. Http dynamic streaming / features. `http://www.adobe.com/products/hds-dynamic-streaming/features.html`, 2013. [Online; accessed 10-June-2014].

[28] IQinVision. Nine key considerations when selecting a video compression algorithm. `http://www.iqeye.com/iqeye/H.264_Considerations.pdf`, . [Online; accessed 3-June-2014].

[29] Dimitrios Miras John D. McCarthy, M. Angela Sasse. Sharp or smooth? comparing the effects of quantization vs. frame rate for streamed video. pages 2429, April 2004.

[30] Satu Jumisko-Pyykkö and Jukka Häkkinen. Evaluation of Subjective Video Quality of Mobile Devices. In *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*, pages 829–838, 2005.

[31] Dong Li and Meilin Cai. A video quality-estimation model for streaming media services based on human visual system. In *Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference on*, pages 1–4, Dec 2009.

[32] Zhe Li, M.K. Sbai, Y. Hadjadj-Aoul, A. Gravey, D. Alliez, J. Garnier, G. Madec, G. Simon, and K. Singh. Network friendly video distribution. In *Network of the Future (NOF), 2012 Third International Conference on the*, pages 1–8, Nov 2012.

[33] Yong Liu, Bin Du, Shuling Wang, Haibo Yang, and Xingwei Wang. Design and implementation of performance testing utility for rtsp streaming media server. In *Pervasive Computing Signal Processing and Applications (PCSPA), 2010 First International Conference on*, pages 193–196, Sept 2010.

[34] T. Lohmar, T. Einarsson, P. Frojdh, F. Gabin, and M. Kampmann. Dynamic adaptive http streaming of live content. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011 IEEE International Symposium on a*, pages 1–8, June 2011.

[35] Gough Lui, Thomas Gallagher, Binghao Li, Andrew G. Dempster, and Chris Rizos. Differences in RSSI readings made by different Wi-Fi chipsets: A limitation of WLAN localization. In *2011 International Conference on Localization and GNSS, ICL-GNSS 2011*, pages 53–57, 2011.

[36] K.J. Ma, R. Nair, and R. Bartos. Drm workflow analysis for over-the-top http segmented delivery. In *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, pages 1–4, July 2011.

[37] Nipan Maniar, Emily Bennett, Steve Hand, and George Allan. The Effect of Mobile Phone Screen Size on Video Based Learning, 2008.

[38] Mark A. Masry and Sheila S. Hemami. A metric for continuous quality evaluation of compressed video with severe distortions, signal processing. *Image Comm*, 19:133–146, 2004.

[39] Masataka Masuda and Takanori Hayashi. Non-intrusive quality monitoring method of voip speech based on network performance metrics. *IEICE Transactions*, pages 304–312, 2006.

[40] Microsoft. Microsoft windows media video (wmv) 9. `http://windows.microsoft.com/`, 2004. [Online; accessed 10-June-2014].

[41] Microsoft. Silverlight - smooth streaming. `http://www.microsoft.com/silverlight/`, 2013. [Online; accessed 10-June-2014].

[42] Mingfeiy. How does video streaming work? `http://mingfeiy.com`, 2013. [Online; accessed 5-June-2014].

[43] T.N. Minhas, M. Shahid, A. Rossholm, B. Lovstrom, H.-J. Zepernick, and M. Fiedler. Assessment of the rating performance of itu-t recommended video quality metrics in the context of video freezes. In *Telecommunication Networks and Applications Conference (ATNAC), 2013 Australasian*, pages 207–212, Nov 2013.

[44] A.K. Moorthy, Lark Kwon Choi, A.C. Bovik, and G. de Veciana. Video quality assessment on mobile devices: Subjective, behavioral and objective studies. *Selected Topics in Signal Processing, IEEE Journal of*, 6(6):652–671, Oct 2012.

[45] NETFLIX. Netflix long term view. `http://ir.netflix.com`, 2014. [Online; accessed 13-June-2014].

[46] LLC Neudesic. Experience iis smooth streaming. `http://www.iis.net/media/experiencesmoothstreaming`, 2014. [Online; accessed 15-June-2014].

[47] T. Oelbaum and K. Diepold. Building a reduced reference video quality metric with very low overhead using multivariate data analysis. 2007.

[48] Institute of Information Technology. Dynamic adaptive streaming over http (dash). `http://www-itec.aau.at/dash/`, 2012. [Online; accessed 10-June-2014].

[49] Society of Motion Picture and Television Engineers (SMPTE). Vc-1. `http://www.smpte.org/`, 2004. [Online; accessed 10-June-2014].

[50] Qualinet: European Network on Quality of Experience in Multimedia System and Services. Qualinet white paper on definitions of quality of experience. `qualinet.eu`, June 2012. [Online; accessed 2-June-2014].

[51] Ee Ping Ong, Mei Hwan Loke, Weisi Lin, Zhongkang Lu, and Susu Yao. Video quality metrics - an analysis for low bit rate videos. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, volume 1, pages I–889–I–892, April 2007.

[52] Germany OPTICOM GmbH. Pevq - advanced perceptual evalutoin of video quality. `http://www.pevq.com/`, 2005. [Online; accessed 4-June-2014].

[53] Ozgur Oyman and Sarabjot Singh. Quality of experience for HTTP adaptive streaming services. *IEEE Communications Magazine*, 50:20–27, 2012.

[54] Stéphane Péchard, Dominique Barba, and Patrick Le Callet. Video Quality Model based on a spatiotemporal features extraction for H.264-coded HDTV sequences. In *Picture Coding Symposium*, 2007.

[55] SA Portugal Telecom Inovação e Sistemas. Arqos. `http://www.ptinovacao.com.br/download.do?image=/pdf/ARQoS.pdf`. [Online; accessed 3-June-2014].

[56] Zhongshui Qu and Jianwei Wang. A color yuv image edge detection method based on histogram equalization transformation. In *Natural Computation (ICNC), 2010 Sixth International Conference on*, volume 7, pages 3546–3549, Aug 2010.

[57] Benjamin Rainer, Markus Waltl, Eva Cheng, Muawiyath Shujau, Christian Timmerer, Stephen Davis, Ian Burnett, Christian Ritz, and Hermann Hellwagner. Investigating the impact of sensory effects on the Quality of Experience and emotional response in web videos. In *2012 4th International Workshop on Quality of Multimedia Experience, QoMEX 2012*, pages 278–283, 2012.

[58] K. Sakamoto, S. Aoyama, S. Asahara, K. Yamashita, and A. Okada. Evaluation of viewing distance vs. tv size on visual fatigue in a home viewing environment. In *Consumer Electronics, 2009. ICCE '09. Digest of Technical Papers International Conference on*, pages 1–2, Jan 2009.

[59] Margaret H.Pinson Stephen Wolf. Application of the ntia general video quality metric (vqm) to hdtv quality monitoring, January 25-26, 2007. [Online; accessed 3-June-2014].

[60] K. T. Tan, M. Ghanbari, and Donald E. Pearson. An objective measurement tool for mpeg video quality. *Signal Processing*, (3):279–294.

[61] Tullis Tom and Albert Bill. Measuring the user experience. `http://www.measuringux.com/`, 2008. [Online; accessed 4-June-2014].

[62] U. Tos and T. Ayav. Adaptive rtp rate control method. In *Computer Software and Applications Conference Workshops (COMPSACW), 2011 IEEE 35th Annual*, pages 7–12, July 2011.

[63] A. S. Umar, R. M. Swash, and A.H. Sadka. Subjective quality assessment of 3d videos. In *AFRICON, 2011*, pages 1–6, Sept 2011.

[64] Internation Telecomunicao Union. Methodology for the subjective assessment of the quality of television pictures. 2012. [Online; accessed 10-June-2014].

[65] K. ur Rehman Laghari, O. Issa, F. Speranza, and T.H. Falk. Quality-of-experience perception for video streaming services: Preliminary subjective and objective results. In *Signal Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific*, pages 1–9, Dec 2012.

[66] Visual Networking Index (VNI). Cisco visual networking index: Global mobile data traffic forecast update, 2013 2018. `http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html`, February 5, 2014. [Online; accessed 10-June-2014].

[67] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13(4):600–612, April 2004.

[68] Jay Loomis; Mike Wasson. Vc-1 technical overview. `http://www.microsoft.com/windows/windowsmedia/howto/articles/vc1techoverview.aspx`, October 2007. [Online; accessed 10-June-2014].

[69] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra. Overview of the h.264/avc video coding standard. *Circuits and Systems for Video Technology, IEEE Transactions on*, 13(7):560–576, July 2003.

[70] Adrian Wills, Thomas Schon, Lennart Ljung, and Brett Ninness. Identification of hammerstein-wiener models. *Automatica*, 1(1):1–14, dec 2011.

[71] Hongfeng Xu, Zhen Chen, Rui Chen, and Junwei Cao. Live streaming with content centric networking. In *Networking and Distributed Computing (ICNDC), 2012 Third International Conference on*, pages 1–5, Oct 2012.

[72] Z. Yetgin and G. Seckin. Progressive download for multimedia broadcast multicast service. *MultiMedia, IEEE*, 16(2):76–85, April 2009.

[73] Junyong You, J. Korhonen, and A. Perkis. Attention modeling for video quality assessment: Balancing global quality and local quality. In *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, pages 914–919, July 2010.

[74] Alex Zambelli's. Microsoft iis smooth streaming technology. `http://ax-comp.com/microsoft-smooth-streaming.html`, 2013. [Online; accessed 10-June-2014].