

# MOBILE VIDEO STREAMING

Fábio Rodrigues (75528), Frederico Rodrigues (75533)  
Instituto Superior Técnico  
Av. Rovisco Pais, 1049-001 Lisboa, Portugal

## RESUMO

A constante partilha de vídeo quer em redes sociais como em diversas outras plataformas é uma realidade nos dias de hoje, sendo que o vídeo representa a maior fatia de tráfego na internet. As pessoas fazem-no em qualquer lugar, a qualquer hora, graças às redes e dispositivos móveis.

Em *mobile video streaming* utilizam-se arquiteturas como CDN, P2P e CDN-P2P, sendo a última a mais recente. O *streaming* de vídeo em dispositivos móveis pode ser feito em *true streaming*, *download and play*, *progressive download and play* e *streaming adaptativo*. Para a realização de *streaming*, os dispositivos móveis utilizam redes como 3G, 4G e Wifi. O codec de vídeo mais utilizado neste campo é o H.264/AVC para vídeo e MP3 e AAC para áudio. Para uma boa qualidade de experiência existem ferramentas como a resiliência a erros, codificação de canal e ocultação de erros, essenciais em *streaming* através de redes sem fios. O *mobile video streaming* é um campo ainda em desenvolvimento, existindo porém já muitos serviços e aplicações como a conhecida Netflix.

**Index terms** – Streaming, vídeo, áudio, móvel

## 1. INTRODUÇÃO

Atualmente os dispositivos móveis não são apenas ferramentas de comunicação. Com a evolução tecnológica, cada vez mais os utilizadores os utilizam como fonte de entretenimento, onde o *streaming* de áudio e vídeo é muito atrativo. O *streaming* é uma tecnologia que envia informações multimédia através da transferência de dados em redes de computadores. Em particular, o *mobile video streaming* é conteúdo comprimido, enviado continuamente através de sinais digitais sem fios, recebido por um utilizador num dispositivo móvel que o descodifica.

Com este artigo pretende-se conhecer a evolução histórica do *streaming* de vídeo desde o primeiro *stream* efetuado, até à forma como se conhece hoje em dia, com uma elevada qualidade de experiência. Abordam-se também as tecnologias associadas ao *streaming* de vídeo: a arquitetura, redes, protocolos, codecs e tecnologias usadas para robustez a erros. Faz-se ainda referência aos serviços e aplicações que existem hoje em dia e aos aspetos legais e sociológicos associados ao uso de *streaming* de vídeo. Por fim, abordam-se o modelo de negócio, a situação atual e perspetivas futuras do *mobile video streaming* em Portugal.

## 2. EVOUÇÃO HISTÓRICA

Através de plataformas de partilha de vídeos e músicas online como o youtube, hoje em dia o acesso a conteúdos audiovisuais está à distância de um clique. Mas nem sempre foi assim tão fácil. O *streaming media* surgiu na Internet em 1995, com o Real Áudio, da Progressive Networks. Esta tecnologia, muito rudimentar, funcionava apenas com som e a transmissão era pouco fiel ao

arquivo original. Apesar da baixa qualidade, a tecnologia foi aceite, uma vez que os utilizadores já não eram obrigados a fazer *downloads* longos. Em janeiro de 1996, transmitiu-se um fluxo de áudio ao vivo através de uma rede, com o Liv @nd in Concert, onde apresentava personalidades como Joey Ramone e Joan Jett. De seguida sentiu-se a necessidade de ir mais longe e a Progressive Networks apresentou o Real Vídeo, realizando o primeiro *streaming* de áudio e vídeo, um jogo de futebol americano entre Cleveland Indians e Seattle Mariners, em 9 de abril de 1997. A Microsoft fez posteriormente um acordo com a Progressive Networks, para desenvolver o seu próprio *software player*, o Netshow, que hoje tem o nome Windows Media Player [1].

Com a massificação dos dispositivos móveis, como os *smartphones* e *tablets*, com a evolução das tecnologias tanto a nível de *hardware* (memória, *routers* e *switches* cada vez mais capacitados) como em *software* (protocolos e codecs mais eficientes) e ainda com a introdução de redes sociais como o facebook e plataformas como o youtube, observou-se um aumento significativo do tráfego global, nomeadamente de *streaming* de vídeo. Atualmente, o *streaming* é responsável pela transmissão de música, conteúdos televisivos, vídeos de entretenimento, aulas (educação à distância) e outros.

## 3. MÉTODOS DE STREAMING

Existem diferentes tipos de *streaming* de vídeo, dos quais se destacam o *true streaming*, *download and play*, *progressive download and play* e o *adaptive streaming*.

Em *true streaming*, o sinal de vídeo chega em tempo real e é mostrado no dispositivo do utilizador de imediato, ou seja, o sinal de vídeo é guardado. Dois benefícios do *true streaming* é a visualização de conteúdos ao vivo e a redução de preço do dispositivo, pois não necessita de memória para o armazenamento.

Em *download and play*, o utilizador faz o *download* de todo o ficheiro que contém o áudio e vídeo, e só depois pode visualizar o conteúdo. O funcionamento é semelhante ao das páginas *web*, utilizando os mesmos protocolos, como o HTTP através de TCP, pelo que não é necessário nenhum *software* adicional.

O *progressive download and play* é um híbrido das duas tecnologias anteriores. O conteúdo de vídeo é dividido em pequenos segmentos, que são enviados para o *software* de leitura. Assim que o *download* de cada um destes segmentos termina, o conteúdo pode ser processado e mostrado ao utilizador, e enquanto isso, já o *download* do próximo segmento começou. Desde que o *download* do próximo segmento termine antes da visualização do primeiro, não haverá interrupção do vídeo. Neste tipo de *streaming*, o conteúdo é guardado temporariamente, pelo que a partir do momento em que o ficheiro chega, o utilizador pode avançar ou recuar na visualização. Uma desvantagem está no custo e complexidade acrescida. Um exemplo é a visualização de vídeos no youtube [2].

O *adaptive streaming* baseia-se no *download* progressivo, mas há a possibilidade de adaptar o ritmo de *download* à velocidade da internet do utilizador. Assim, a qualidade do vídeo depende da qualidade de conexão deste.

#### 4. TECNOLOGIA

A realização de *streaming* em dispositivos móveis implica a implementação de inúmeras tecnologias. Nesta secção, tratam-se algumas muito importantes como a arquitetura, redes móveis e wireless, resiliência a erros, codificação de canal, ocultação de erros, protocolos, codecs e controlo de acesso e pagamento de conteúdos não *free*.

##### 4.1. Arquitetura

O *streaming* de vídeo necessita de uma infraestrutura adequada, representada na figura 1. Antes de mais, o conteúdo necessita de ser preparado e processado passando pelas seguintes fases: captura, edição e pré-processamento. De seguida passa pelo processo de compressão, rotulagem e indexação, e por fim publicação.

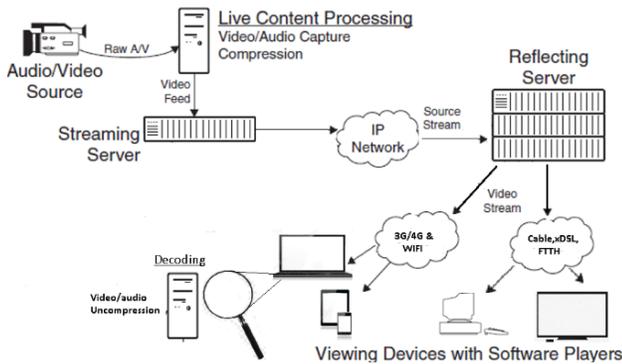


Figura 1-Arquitetura do streaming de vídeo, em rede IP.

A captura refere-se ao processo de capturar o conteúdo de vídeo e submetê-lo ao sistema de preparação. Aqui, o vídeo pode ser editado, dependendo do objetivo final do mesmo. É feita a organização do conteúdo, sincronizando o áudio, o texto sobreposto (por exemplo, legendas), e outros efeitos visuais e auditivos. De seguida, o vídeo é pré-processado, onde é feita a preparação para a compressão e eventuais correções de cor, remoção de ruído e ajuste do tamanho da imagem. No processo de compressão, o tamanho dos *streams* de áudio e vídeo é reduzido tirando partido das redundâncias temporal e espacial e da irrelevância do conteúdo. Se os dispositivos suportarem diferentes débitos binários, resoluções de vídeo, ou *software* de reprodução, então o processo de compressão é repetido para as diferentes combinações. Na rotulagem e indexação, são providenciadas descrições de conteúdo (meta dados), para que os utilizadores os consigam localizar. Por fim, o conteúdo é publicado, sendo transferido para o *streaming server* e são criadas páginas web que contêm links para os *streams* multimédia.

O *streaming server* é o responsável por distribuir os *streams* ao dispositivo de cada utilizador. Este *server* armazena o conteúdo multimédia e cria um *stream* para cada pedido do utilizador. Outra das responsabilidades do *streaming server* é criar cabeçalhos para os pacotes de áudio e vídeo enviados por rede IP, e assim chegarem ao endereço de destino correto. O *streaming server* pode ainda melhorar a segurança do ficheiro através da codificação dos

conteúdos. Porém, caso a rede fique congestionada, o *streaming server* deve começar a enviar o conteúdo gravado em baixa velocidade, sem que o utilizador se aperceba - *scaling a stream*.

De modo a suportar muitos utilizadores, o conteúdo de áudio e vídeo pode ser guardado em *reflecting servers*, onde são feitas muitas cópias, para entregar aos utilizadores em tempo real, em várias localizações[2]. A conexão de cada utilizador é feita de modo a que este receba o conteúdo do *reflecting server* mais próximo. É a CDN (Content Delivery Network) que engloba um conjunto de servidores localizados em diferentes localizações, que é responsável por ajudar na entrega de conteúdos, escolhendo qual o servidor que tem uma maior eficiência na entrega dos mesmos. O problema é a necessidade de muitos servidores e cada um deles poder ocupar uma largura de banda elevada, proporcionando ao cliente um conteúdo com má qualidade de transmissão [3].

Até agora discutiu-se acerca de *streaming* de vídeo de um servidor para um ou mais recetores. No entanto, a tecnologia peer-to-peer (P2P) video streaming, cuja estrutura se apresenta na figura 2, tem ganho especial interesse nos últimos anos, em que grandes quantidades de informação são pedidas aos servidores tradicionais, sobrecarregando-os. Nos sistemas P2P, todos os utilizadores funcionam como “servidores”. Cada utilizador obtém ficheiros de outros clientes, eliminando a necessidade de haverem servidores de ficheiros centralizados. A estrutura diz-se descentralizada. O sistema P2P utiliza uma codificação de vídeo escalável de modo a satisfazer as restrições impostas pela largura de banda disponível em qualquer ponto da rede móvel. Em P2P usa-se ainda Multiple Description Coding (MDC), que escolhe um caminho em vez de outro, por exemplo se este ficar indisponível, no caso de um cliente abandonar o sistema [4].

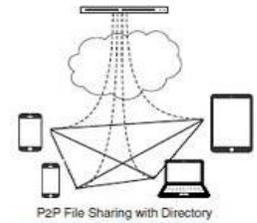


Figura 2-Arquitetura P2P.

A arquitetura CDN-P2P, representada na figura 3, é um sistema híbrido que tenta captar os benefícios das duas arquiteturas anteriormente referidas. Teve como objetivo alcançar a estabilidade dos sistemas P2P e os níveis de atraso e de taxa de transferência dos sistemas CDN [3].

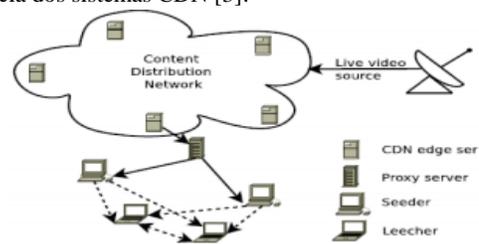


Figura 3-Arquitetura CDN-P2P [3].

Por fim, é o *software* de reprodução, presente no dispositivo móvel do utilizador, o responsável por aceitar os *streams*, descomprimi-los e convertê-los em imagens. É este *software*, que faz a descriptação do conteúdo. Caso os pacotes tenham sido separados em diferentes *streams* pelos protocolos utilizados (como RTP), então cabe ao *software* de reprodução ressincronizá-los.

##### 4.2. Redes móveis e wireless

A recente redução substancial de largura de banda necessária para transmissão de áudio e vídeo deve-se não só ao desenvolvimento

de algoritmos de compressão, mas também devido à massificação de tecnologias de comunicação como as redes móveis e wireless. Algumas das redes móveis mais comuns são as redes 3G e 4G/LTE.

A rede 3G (3ª geração que substituiu a rede 2G com uma maneira mais eficiente de se navegar na internet. A tecnologia móvel 3G suporta um número maior de clientes de voz e dados, além de maiores taxas de dados a um custo menor que a 2G. A rede 4G (4ª geração), também conhecida por LTE, sigla para Long Term Evolution, é o padrão mais recente, ainda em difusão no mundo e promete transmissões de dados em bandas ultra largas. É uma rede mais rápida e estável, pois prioriza o tráfego de dados em vez do tráfego de voz, como acontecia nas outras gerações [5]. Quer a rede 3G quer a 4G são muito utilizadas na *streaming* de vídeo. É também possível fazer *streaming* de vídeo a partir de um dispositivo móvel ligado a uma rede Wi-Fi, desde que o utilizador esteja na área de abrangência de um ponto de acesso (conhecido por *hotspot*) ou local público onde opere rede sem fios.

### 4.3. Resiliência a erros, codificação de canal e ocultação de erros

A partir do momento em que o vídeo é transmitido nas redes descritas na secção 4.2, o conteúdo está sujeito a sofrer perdas, e portanto há algumas considerações a ter em conta além da eficiência de compressão, nomeadamente a robustez a erros de canal. Existem três categorias de técnicas que tornam a ocorrência destes menos desastrosa, melhorando a qualidade de experiência do utilizador: resiliência a erros, codificação de canal e ocultação de erros. Porém há que ter sempre em conta o trade-off entre eficiência de compressão e resistência a erros. A resiliência a erros tem como objectivo tornar o fluxo de bits mais resistente a erros de canal e usa ferramentas como marcadores de resincronização, reversible variable length coding (RVLC), independent segment decoding e inserção de blocos ou tramas intra.

Os marcadores de resincronização são utilizados para codificação de comprimento variável, uma vez que não se sabe o início de cada palavra a priori pode causar erros catastróficos. O funcionamento baseia-se na inserção periódica de marcadores no fluxo de bits. Se for detetado um erro, o decodificador procura pelo próximo marcador para recomeçar a descodificação. Os dados entre a deteção do erro e o marcador, são obviamente perdidos. Desta forma, quanto menor for a distância entre os marcadores, maior será a resiliência a erros, mas a eficiência de compressão diminui. A ideia ilustra-se na figura 4.



Figura 4- Utilização dos marcadores de resincronização [2].

Mesmo com a utilização dos marcadores de resincronização, a perda de bits devido a um único bit errado, pode ser quase tão grande como a perda de informação no intervalo entre marcadores, dependendo da localização do bit errado. A técnica RVLC, ilustrada na figura 5, reduz a quantidade de informação perdida, criando palavras de código que podem ser descodificadas em ambas as direções, para a frente e para trás. Quando o erro é detetado, a descodificação começa no próximo marcador no sentido oposto. Havendo um bit errado entre os marcadores, será detetada uma palavra de código inválida e a descodificação pára

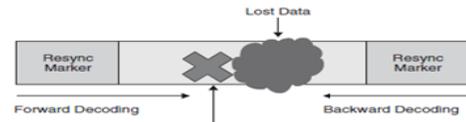


Figura 5- Técnica RVLC [2].

novamente, perdendo-se ainda alguns dados, embora em menor quantidade. Outra técnica utilizada é a chamada independent segment decoding. Consiste em dividir uma trama em regiões e apenas a informação da localização espacial correspondente da trama anterior pode ser usada para codificar a trama atual. Desta forma limita-se a propagação de erro. No entanto, todas as subsequentes tramas continuam corrompidas. A propagação de erro só termina totalmente quando uma trama é codificada em modo intra, pelo que faz sentido inserir blocos ou tramas intra, para melhorar a resiliência a erros.

A codificação de canal consiste em adicionar redundância (bits extra) no fluxo de bits comprimido para ajudar o decodificador de canal a detetar e corrigir erros de canal. A seleção da codificação de canal deve considerar as características do canal e a modulação. Também o tipo de modulação deve ter em conta as características do canal, nomeadamente em redes móveis tem de se ter cuidado com a taxa de erros de canal. Mesmo utilizando técnicas de resiliência a erros e codificação de canal, é impossível não perder alguma informação. O objetivo da ocultação de erros é estimar a informação perdida, e ocultá-la do utilizador. Existem duas técnicas: as intra e as inter. As intra ocultam uma parte da imagem, usando informações adjacentes na mesma trama. As inter utilizam informação de tramas anteriores àquela onde ocorreu erro. Neste tipo de técnica são muito usados os vetores de movimento com o objetivo de substituir um macrobloco na trama atual, pelo correspondente macrobloco com compensação de movimento da trama anterior. Se o vetor de movimento não for conhecido, este é estimado, por exemplo, analisando os vetores de movimento dos macroblocos vizinhos [2].

### 4.4. Protocolos

Os protocolos utilizados no *mobile video streaming* podem ser classificados em três categorias, consoante a sua função: protocolos de rede, de transporte e de sessão.

Os protocolos de rede são responsáveis pelo endereçamento e encaminhamento da informação. São exemplos o protocolo de rede IP (Internet Protocol), e o ATM (Asynchronous Transfer Mode).

Os protocolos de transporte entregam a informação. Os protocolos de transporte para redes IP são o TCP (Transmission Control Protocol), UDP (User Data Protocol), RTP (Real Transport Protocol) e RTCP (Real-time Transport Control Protocol).

O protocolo TCP é orientado à conexão e é responsável pela comunicação fiável entre dois intervenientes. O tamanho dos pacotes varia e é assegurado que chegam ao destino sem erros e completos. Quando um pacote se perde na rede ou é recebido com erros, procede-se à sua retransmissão. Devido à constante troca de mensagens e retransmissões de pacotes, a velocidade de transmissão do protocolo TCP diminui.

O protocolo UDP não possui mecanismos de correção de erros, possibilitando uma troca direta de pacotes através de uma rede IP. Os pacotes são enviados independentemente do estado dos anteriores, portanto é um protocolo não fiável. Há ainda a possibilidade de os pacotes chegarem fora de ordem ou duplicados.

Os pacotes devem portanto estar bem identificados e as aplicações têm de ser capazes de detetar e corrigir estas anomalias. Há um menor consumo de recursos, sendo a velocidade de transmissão mais elevada relativamente ao TCP e não é necessário uma ligação entre o emissor e cada recetor.

Os protocolos RTP e RTCP são implementados no topo dos protocolos TCP/UDP, fornecendo apenas um conjunto de serviços que suportam o transporte em tempo real de áudio e vídeo. O RTP tem algumas funcionalidades como *time-stamping*, para a sincronização do áudio e vídeo, atribuição de um número a cada pacote facilitando a deteção de perdas no recetor e identificação da fonte. O RTCP tem como principal objetivo fornecer feedback de QoS aos participantes de uma sessão RTP.

Os protocolos de sessão definem as mensagens e procedimentos para estabelecer e controlar a transmissão do conteúdo audiovisual entre o cliente e servidor. Um exemplo é o RTSP (Real Time Streaming Protocol). O RTSP é um protocolo ao nível da aplicação que fornece controlo sobre a entrega de dados em tempo-real. Este protocolo tem por objetivo controlar múltiplas sessões de transmissão de dados, possibilitar a escolha do protocolo de transporte, tal como UDP, multicast UDP ou TCP. As principais funções do protocolo RTSP são o controlo remoto de servidores multimédia, e a inicialização e manutenção de sessões [7]. Algumas das mensagens trocadas estão representadas na figura 6. Este protocolo não tem interferência no transporte dos dados, deixando essas funções para outros protocolos, normalmente o RTP. Uma grande desvantagem deste protocolo é o facto de as suas transmissões serem muitas vezes bloqueadas por *routers* e *firewalls* [6].

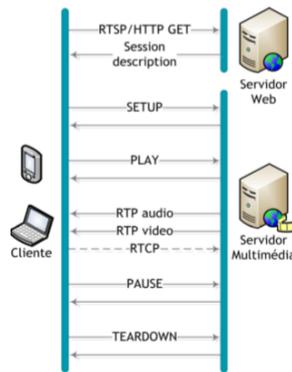


Figura 6- Comunicação servidor-bloqueadas por *routers* e utilizador utilizando RTSP [7].

A figura 7 sintetiza a relação entre as diferentes categorias de protocolos. Num sistema de *stream* de vídeo, o fluxo de bits é empacotado pela camada RTP enquanto que as camadas RTCP e RTSP fornecem a informação de controlo. Depois o fluxo de bits segue para a camada UDP/TCP e depois para a camada IP. No recetor, o procedimento é o inverso.

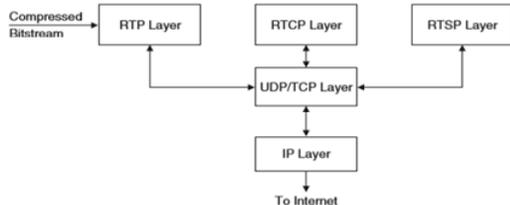


Figura 7- Estrutura dos protocolos utilizados.

Porém os protocolos anteriormente descritos são já considerados os clássicos protocolos de *streaming*. A nova tendência no *streaming* é utilizar tecnologias de *streaming* adaptativo, que utilizam o protocolo HTTP. O HTTP é um protocolo *stateless*, utiliza um débito binário adaptativo, ao contrário do que acontecia com os protocolos clássicos em que a velocidade de *stream* era a velocidade de codificação. Também a

entrega baseada em HTTP não tem quaisquer problemas de *firewalls* ou *routers*, nem requer *proxies*, *caches* ou servidores especiais. É o programa de reprodução do cliente que controla o débito binário dos fragmentos a serem pedidos ao servidor, que depende da largura de banda e capacidade de processador do cliente. Algumas destas tecnologias, que se focam na ideia de utilizar a mesma infraestrutura que a implementada na internet são o HLS (HTTP Live Streaming), da Apple, HDS (HTTP Dynamic Streaming), da Adobe e Smooth Streaming, da Microsoft.

O HLS permite enviar transmissões ao vivo ou pré-gravadas através de HTTP, a partir de um servidor web comum para dispositivos iOS. O conteúdo é codificado (o vídeo com H.264 e o áudio com MP3 ou AAC) e segmentado no servidor e a distribuição é feita através de um servidor web ou sistema de cache web, tal como representado na figura 8. O servidor pode conter várias versões dos segmentos de vídeo em diferentes formatos. Assim, os terminais móveis com maior largura de banda disponível podem utilizar versões com qualidade superior relativamente aos terminais com menor largura de banda disponível, por exemplo o Wi-Fi. Há portanto um ajuste dinâmico da qualidade do vídeo que assegura uma visualização contínua do conteúdo [8].

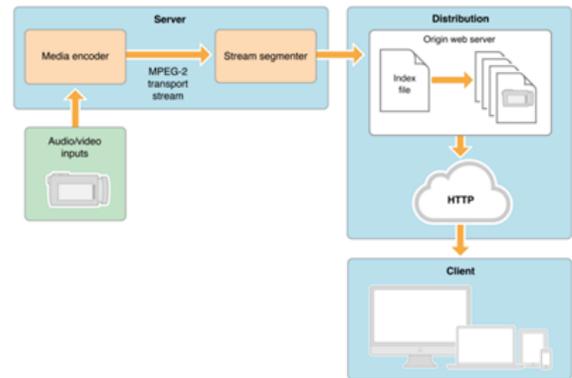


Figura 8- Configuração do streaming HLS [8].

O HDS permite a entrega de vídeo pré-gravado ou ao vivo com um débito binário adaptativo, através de conexões HTTP regulares. Aproveita as infraestruturas de cache existentes [9]. A figura 9 demonstra o processo desde a preparação do vídeo até ao

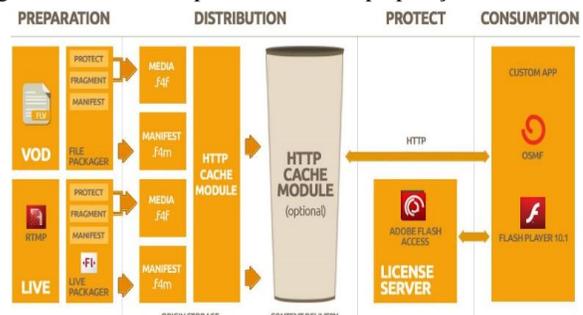


Figura 9- Preparação, distribuição, proteção e consumo de vídeo através de HDS [3].

consumo num dispositivo móvel. Tal como no protocolo RTSP/RTP também no HDS existe comunicação servidor/cliente. Quer o cliente quer o servidor sabem a largura de banda do cliente, pelo que ambos podem adaptar o vídeo a transmitir.

O Smooth Streaming, cuja arquitetura se apresenta na figura 10, tem uma implementação baseada na entrega HTTP adaptativa.

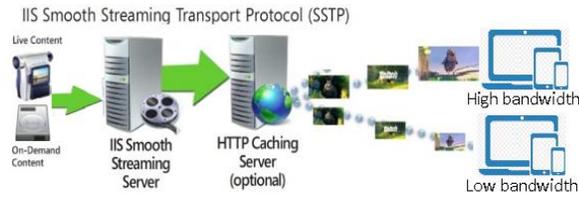


Figura 10–Arquitetura do smooth streaming [10].

Ao contrário de outras soluções, esta não divide os ficheiros originais fisicamente, ou seja, no servidor, em vez de estarem guardados pequenos ficheiros para cada conteúdo, estão guardados os ficheiros completos. A divisão é feita virtualmente consoante a largura de banda e processador do utilizador, melhorando significativamente o processo de gestão de ficheiros. Existem ainda ficheiros com diversas qualidades para que haja adaptação da qualidade do vídeo ao estado da rede móvel. Usa a plataforma IIS, Internet Information Services, também esta baseada em http [10].

O MPEG-Dash surgiu da necessidade de haver interoperabilidade entre os servidores e clientes de diferentes fornecedores. O conteúdo existe no servidor em duas partes: Media Presentation Description (MPD) e os segmentos que incluem os diferentes débitos binários do conteúdo, sob a forma de pedaços. Para reproduzir o conteúdo, primeiro o cliente recebe o MPD, ficando a saber todas as características do conteúdo, como por exemplo a resolução do vídeo. Com estas informações o cliente DASH começa a transmissão de conteúdo usando solicitações HTTP GET, monitorando variações da largura de banda disponível, tirando do servidor segmentos com diferentes débitos [11]. A arquitetura apresenta-se esquematizada na figura 11.

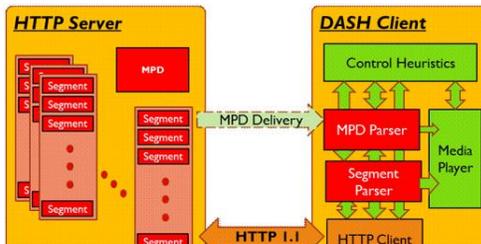


Figura 11–Arquitetura MPEG-DASH [11].

## 4.5. Codecs

Na distribuição de vídeo em *stream*, muitos codecs têm vindo a ser utilizados e substituídos por outros mais eficientes. Embora existam alguns concorrentes como o VP8 da Google, VC-1 da Microsoft ou o Theora da Xiph.org Foundation, hoje o codec de vídeo mais utilizado, nomeadamente em *streaming* em dispositivos móveis, é o H.264/AVC, razão pela qual será tratado neste artigo detalhadamente. Também serão tratados os codecs de áudio MP3 e AAC pois são os utilizados em conjunto com o codec H.264/AVC. Já o codec de áudio Vorbis é utilizado em conjunto com os codecs de vídeo VP8 e Theora.

O codec H.264/AVC foi desenvolvido juntamente pela ITU-T e ISO/IEC MPEG, mas ao contrário de codecs anteriores, como o MPEG-2 Vídeo, não surgiu para uma aplicação específica. O codec H.264/AVC surgiu por volta de 2003 e substituiu as normas anteriores, pois possui uma codificação melhorada, nomeadamente cerca de 50% menos débito para a mesma qualidade subjetiva, em comparação com as normas já existentes (MPEG-2 Vídeo). Esta melhoria foi conseguida sobretudo com um aumento de

complexidade de codificação. Para além disto, tem um bom desempenho de resiliência a erros em ambientes móveis e internet via *wireless*. A arquitetura de codificação do H.264/AVC, representada na figura 12, é muito semelhante à arquitetura de normas anteriores, mas os blocos são mais complexos.

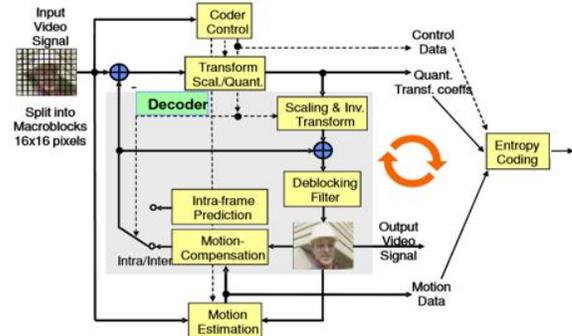


Figura 12- Arquitetura do codec H.264/AVC [12].

O ganho de eficiência está essencialmente ligado à compensação de movimento com blocos de tamanho variável (e menores), múltiplas tramas de referência, transformada hierárquica com blocos menores (4x4), filtro de efeito de bloco no *loop* de predição e codificação entrópica melhorada (Context Adaptive Variable Length Coding-CAVLC e Context Based Adaptive Binary Arithmetic Coding-CABAC) [12]. Este codec, relativamente a outros anteriores, é mais eficiente em redes móveis pois reduz o consumo de tráfego, assim como suporta larguras de banda inferiores. Para além disto, destaca-se também por possuir uma boa resiliência a erros, utilizando várias ferramentas essenciais em *mobile video streaming*, algumas delas já usadas em normas anteriores e referidas na secção 4.3. O codec H.264/AVC utiliza ferramentas como segmentação da imagem em *slices*, inserção de *slices* intra e tramas de refrescamento IDR, *reference picture selection*, partição de dados, ordenação flexível de macroblocos (FMO) e *slices* redundantes.

A inserção de *slices* intra continua a ser uma ferramenta essencial para evitar a propagação de erros. Existem também as tramas de refrescamento (IDR) que faz com que em figuras seguintes não possam usar figuras anteriores às IDR na compensação de movimento.

A ferramenta *Reference Picture Selection* consiste em o decodificador, caso exista feedback, sinalizar o codificador quando a imagem recebida foi danificada e portanto este não a utilizará no *loop* de predição.

A partição de dados consiste na partição de um *slice* em 3 partições A, B e C, permitindo diferentes proteções a erros de acordo com a importância dos elementos.

A ordenação flexível de macroblocos consiste em agrupar os macroblocos, através de um mapeamento. Por exemplo, a figura 13 possui dois *slices*: o de macroblocos em branco e de macroblocos em cinza. Este mapeamento facilita a ocultação ou disfarce de erros, porém diminui a eficiência de compressão pois quebra os mecanismos de predição dentro da própria figura.



Figura 13- Ordenação flexível dos macroblocos (FMO) [4].

A ferramenta *slices* redundantes permite a codificação de *slices* usando menos bits e só são utilizados se o *slice* primário, de melhor qualidade, não estiver disponível no decodificador[4].

Para gerir o elevado número de ferramentas de codificação da norma H.264/AVC e a máxima complexidade do decodificador aceitável para determinada aplicação há diversos níveis/perfis que podem ser utilizados, figura 14.

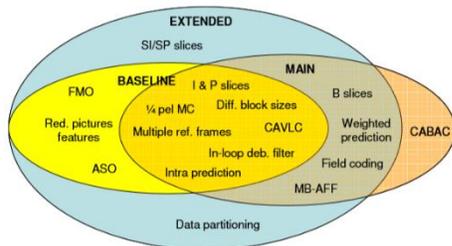


Figura 14- Perfis/Níveis da norma H.264/AVC [12].

Os perfis *Baseline* e *Extended* são aqueles que mais ferramentas incluem no que diz respeito à resiliência a erros, por isso mais indicados quando se codifica vídeo para ser transmitido em canais com taxas de erro elevadas, como em ambientes móveis.

Para além disto, para resolver a grande heterogeneidade atual em termos de dispositivos, criou-se a extensão SVC do H.264/AVC, que oferece uma stream contruída por camadas. Cada cliente pode utilizar tantas camadas quanto a capacidade do seu dispositivo móvel. Uma maior proteção da camada base é também uma forma de garantir uma qualidade mínima aceitável.

O codec MP3, de MPEG-1 Audio Layer 3, é um sistema de compactação de áudio que frequentemente acompanha o vídeo compactado em H.264/AVC. Trata-se de um formato bastante popular devido à sua taxa de compactação, que pode chegar a 12 vezes o tamanho do ficheiro original, não havendo perdas significativas e perceptíveis na qualidade de som. O método de compressão consiste em retirar do áudio amplitudes ou frequências não perceptíveis ao ser humano (irrelevância), e em escala menor, tira vantagem de alguma redundância existente. O seu débito binário é da ordem dos kbps, sendo 128 kbps a taxa-padrão. Porém, o MP3 tem desvantagens como o seu atraso e a relação de compactação comparando com outros codecs de streaming.

O codec Advanced Audio Coding (AAC), que também frequentemente acompanha o vídeo compactado em H.264/AVC foi projetado para substituir o codec MP3, conseguindo melhor qualidade de som para débitos binários idênticos. Também a sua taxa de compactação é superior face à do MP3. À medida que o débito binário aumenta, a eficiência do formato de áudio torna-se menos importante em relação à implementação do codificador, e a vantagem do AAC em relação ao MP3 passa a ter uma diferença menos significativa. Quanto ao número de canais, o AAC é mais abrangente que o MP3, possuindo 48 canais contra os 5.1 no MP3. O AAC é o formato padrão para o iPhone, o iPod e o iTunes da Apple, entre outros dispositivos móveis. O AAC, embora tecnicamente superior, ainda não teve toda a sua capacidade de implementação explorada [3]

#### 4.6. Controlo de acesso e pagamento para conteúdos não free

A proteção dos conteúdos cujo acesso não é universal é também algo a ter em conta na *mobile video streaming*. Existem

plataformas de vídeo online que servem este propósito. Um exemplo é a plataforma Ooyala, que oferece uma solução completa que permite aos clientes proteger o seu conteúdo com a flexibilidade necessária para apoiar uma variedade de modelos de negócios. Neste tipo de plataforma destacam-se a tecnologia autenticação e autorização de reprodução (*playback authorization*). A autenticação é aquela que permite identificar cada cliente como utente único, em que é necessário um sistema de gestão de assinantes. Já a autorização de reprodução garante que apenas os clientes autorizados possam ver os conteúdos. Por exemplo, os utilizadores estão impedidos de partilhar conteúdos com amigos que não tenham feito subscrição a conteúdos *premium*.

## 5. MODELO DE NEGÓCIO

Uma das grandes aplicações do *streaming* de vídeo é a videovigilância, em que atualmente o cliente pode ter acesso em tempo real às câmaras de segurança através da internet no seu dispositivo. Já o *streaming* de áudio e vídeo oferece outro tipo de utilizações. Duas essenciais são a comunicação através de uma câmara entre pessoas em pontos diferentes do mundo e a transmissão ao vivo de eventos de grande escala por algumas empresas. A enorme utilização deste tipo de plataformas que permitem transmitir e assistir vídeo, tornam a inserção de publicidade e *sponsored* links bastante aliciente monetariamente.

Um modelo que se tem tornado bastante comum em aplicações e sites de *streaming* de vídeo é o Anonymous Ad-Supported. Este modelo fornece conteúdos de graça, mas restringe o acesso a recursos *premium*. É utilizado essencialmente como ferramenta de marketing, em que os conteúdos podem ser antigos, limitados em termos de duração temporal e on-demand. Um exemplo é a aplicação NBC Olympics sem autenticação. Quanto aos conteúdos Authenticated Ad-supported, são aqueles que só podem ser vistos através de autenticação, com publicidade antes, durante e depois do vídeo. Muitas vezes os conteúdos em tempo real estão associados a este modelo de negócio. Um exemplo é a aplicação NBC Olympics com autenticação requisitada. Existe também outro modelo de negócio, a subscrição, em que se tem como exemplo a Netflix. Neste tipo de modelo, os clientes já têm por exemplo acesso aos conteúdos completos, e a possibilidade de os ver em HD. Por último tem-se o modelo transação, em que um exemplo é o iTunes. Há três tipos de transação: compra, aluguer ou vídeo pré-gravado(VOD) e pagamento-por-visualização (PPV) [13]. Quer a subscrição quer a transação podem ter como benefícios a remoção de publicidade.

Um modelo de negócio muito comum consiste em não cobrar a primeira mensalidade num serviço ou aplicação. Outra forma é disponibilizar aos utentes um serviço durante um período de experimentação (trial).

## 6. ASPECTOS LEGAIS E SOCIOLÓGICOS

Desde há alguns anos, a Internet é utilizada de forma intensiva para a violação de direitos de autor. Um exemplo de violação de direitos de autor são as transmissões em tempo real de jogos de futebol, cuja exclusividade normalmente pertence a canais de sinal fechado, como a Sport TV. A visualização destes conteúdos seria apenas possível mediante pagamento de mensalidade.

No combate à violação de direitos de autor em Portugal, temos por exemplo o fecho recente do site WarezTuga pelo U.S Copyright Group. O WarezTuga é um site de *streaming* de

conteúdos como filmes, séries, documentários entre outros, que operava em Alistar-security.ro, serviço de alojamento sedado na Roménia. Em fevereiro de 2015 pertenceu ao top 20 de sites mais populares em Portugal [14].

Temos ainda, como exemplo, os operadores de serviços de telecomunicações que são obrigados a bloquear o acesso a sites com conteúdos considerados ilegais enviados pela Inspeção-Geral das Actividades Culturais (IGAC) à Meo, NOS, Vodafone e Cabovisão. A IGAC afirma que nos sites em causa "eram distribuídas obras protegidas à margem de qualquer autorização dos titulares de direitos, de forma ilegal" [15].

Em termos sociológicos, o *mobile video streaming* tem inúmeras aplicações como por exemplo na componente educacional. Segundo o estudo IAB – Interactive Advertising Bureau Mobile phone Video Diaries, conduzido pela empresa OnDevice Research, vídeos de notícias e do estilo “how to” estão na lista dos mais vistos em dispositivos móveis [16].

Através do *true streaming*, a distância tornou-se relativa. O contacto com vídeo entre amigos ou familiares a longas distâncias tornou-se possível, assim como a realização de reuniões e conferências por parte das empresas, que de outra forma, exigiria presença física. O crescente uso de *streaming* em termos de tecnologia móvel é também uma mais-valia. Com a diferença horária entre os vários pontos do mundo, a liberdade do utilizador é muito importante. O utente pode efetuar chamadas de vídeo em qualquer lugar para qualquer lugar do mundo, que de outra forma teria que ser feita num escritório.

Por outro lado, o *mobile video streaming* levou também a um isolamento social. Hoje em dia é muito comum que cada pessoa tenha o seu dispositivo móvel, onde tem acesso aos mais variados conteúdos. Com isto, hábitos como ver televisão em família ou ida ao cinema com amigos são hoje menos comuns. Como consequência, temos o encerramento de inúmeras salas de cinemas, por exemplo da empresa portuguesa Castello Lopes Cinemas.

## 7. SERVIÇOS E APLICAÇÕES

O *mobile video streaming* é utilizado para vários fins, nomeadamente educativo, entretenimento e empresarial. Abordam-se de seguida serviços e aplicações como a Netflix, First Row, o Zero-Foot-Print Live Mobile Streaming e ustream.

A Netflix é um distribuidor internacional de *streaming* de conteúdos. É possível assistir aos filmes e séries no televisor através de Smart TVs, de consolas de jogos, através de um leitor Blu-ray ou sistema cinema em casa. Mas este serviço pode ser mesmo utilizado em qualquer lugar, em *tablet*, telefone, portátil, sendo apenas necessário ter a aplicação. Atualmente a utilização do serviço varia sensivelmente entre os 8 e 12 euros por mês, consoante a qualidade de imagem e o número de ecrãs que se pode ver em simultâneo, sempre com séries e filmes ilimitados [17]. Particularmente, em Portugal a aplicação Netflix é uma das mais descarregadas, segundo o director executivo Reed Hastings.

O First Row é um serviço que disponibiliza *streams* em tempo real na internet, dos mais variados desportos, gratuitamente. Para isso, basta aceder ao site através de um dispositivo com conexão à internet e seleccionar o jogo pretendido. É a forma através da qual muitas pessoas, que não querem pagar os canais onde os jogos são transmitidos, acompanham os desportos. Contudo, o cliente tem de se sujeitar muitas vezes a publicidade e pop-ups indesejados e a uma fraca qualidade de imagem.

A Bring-Your-Own-Device introduziu um serviço com aplicações no espaço de trabalho, o Zero-Foot-Print Live Mobile Streaming, que tem a capacidade de transmitir vídeo de alta qualidade em tempo real para qualquer dispositivo dentro e fora da organização. Esta tecnologia permite que um membro assista a uma apresentação no dispositivo móvel, sem ter de estar presente, ou que um médico veja uma cirurgia noutra parte do edifício.

O ustream é um serviço de vídeo online que permite transmitir facilmente conteúdos ao vivo a partir de um computador pessoal ou dispositivo móvel, para todo o mundo. A aplicação é compatível com redes 3G e Wifi, pelo que a sua utilização em dispositivos móveis é enorme.

## 8. SITUAÇÃO EM PORTUGAL

Segundo o estudo do IAB (Interactive Advertising Bureau), “Mobile Video 2015:A global perspective”, o consumo dos portugueses em dispositivos móveis apresenta um crescimento de 27%, face ao ano 2014. No entanto, Portugal ainda se encontra atrás de países como EUA (50%), Canada (42%), Nova Zelândia (42%), África do Sul (42%) e Reino Unido (40%). Outra das principais conclusões do estudo, relativamente a Portugal, refere-se à recetividade a publicidade como forma de ter acesso a conteúdos gratuitos: 86% dos inquiridos portugueses afirmaram preferir ter acesso a vídeos gratuitos com publicidade como opção, a ter que pagar para aceder aos conteúdos vídeo. Portugal também está no topo dos países onde a descoberta de vídeos através da publicidade é mais elevada (atinge os 19%), ficando apenas atrás dos EUA (com 22%) e França (com 20%). Quanto aos tipos de conteúdo mais vistos, os vídeos musicais e os vídeos divertidos/virais estão no topo da lista, empatados com 49% das respostas em termos globais, seguidos dos *trailers* de filmes com 41%. Em Portugal existem algumas diferenças, como os vídeos musicais a alcançarem apenas 25% de respostas, enquanto os filmes divertidos/virais chegam aos 50% e os trailers aos 48% de respostas. Os portugueses também se destacam quanto à partilha de conteúdos *mobile*, com 75% dos inquiridos a partilharem vídeos com outras pessoas, um dos valores mais elevados [16].

## 9. PERSPECTIVAS DE EVOLUÇÃO FUTURA

O constante aumento de dispositivos móveis conectados à internet, das resoluções dos ecrãs e qualidade de imagem, e da velocidade de conexão, culmina claramente num aumento do *mobile video streaming*. O gráfico da figura 15 mostra o crescimento de tráfego móvel, prevendo um crescimento de 13 vezes de tráfego de vídeo em dispositivos móveis de 2014 para 2020 e que 60% de todo o tráfego móvel será de vídeo [18].

Em termos de codecs há uma tendência para o aumento de complexidade e eficiência de compressão. Naturalmente esta tendência tem de ser acompanhada pelo desenvolvimento das baterias dos dispositivos móveis. Prevê-se, nomeadamente, que o futuro esteja na utilização da recente norma H.265 (High Efficiency Video Codec-HEVC) ou da norma VP9. Apesar de serem normas bastante mais complexas apresentam ganhos na eficiência de compressão. Espera-se que a tecnologia MPEG-DASH, que combina os protocolos HLS da Apple, HDS da Adobe e Smooth Streaming da Microsoft, venha a ser adotada globalmente. Num momento em que as redes 4G começam a ser o padrão, alguns fabricantes já trabalham no 4.5G e também no 5G.

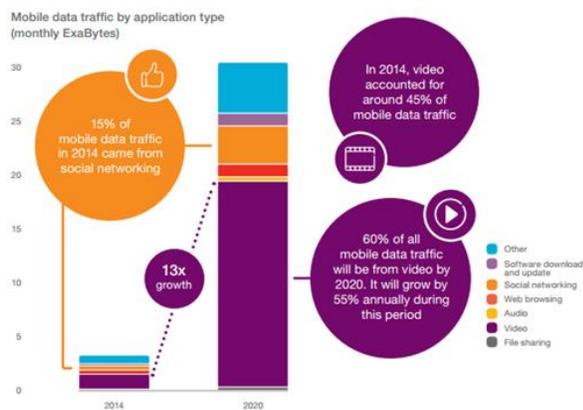


Figura 15-Tráfego móvel em 2014 e 2020 [18].

Porém as tecnologias de quinta-geração só deverão ser implementadas nos operadores em 2020.

## 10. CONCLUSÕES

Atualmente, assiste-se a uma explosão do mercado móvel e a requisição por vídeos de alta qualidade nunca foi tão grande. Muitas empresas implementam os seus próprios modelos de negócio, disponibilizando serviços para a realização de streaming de vídeo. O streaming revolucionou os hábitos das pessoas e trouxe muitos aspectos positivos. Porém, há que saber impor limites na sua utilização para que o impacto destas tecnologias não leve à perda de valores sociais importantes, ou ao aumento de criminalidade online, por exemplo a violação dos direitos de autor.

O *streaming* adaptativo, através de HTTP é a nova tendência de realização de streaming. Devido às questões de largura de banda, a compressão de vídeo é essencial para o *streaming* de vídeo. Em conexões móveis os canais são muito críticos, pelo que as tecnologias de resiliência a erros, codificação de canal e ocultação de erros são fundamentais, mesmo que baixem a eficiência de compressão. Apesar de outros codecs de vídeo terem sido desenvolvidos já depois do H.264/AVC, é ainda o H.264/AVC o codec de vídeo mais popular atualmente, utilizado em conjunto com os codecs de áudio MP3 e AAC.

## REFERÊNCIAS

- [1] Interrogação Digital, “O que é o Streaming”, <http://www.interrogacaodigital.com/central/o-que-e-streaming/>, acessado a 10/10/15.
- [2] SIMPSON, Wes-Video Over IP.2ª Edição. United States of America: Elsevier, 2008, ISBN 978-0-240-81084-3.
- [3] Ana Gomes; Rodrigo Lourenço, “Internet Live Streaming”, [http://www.img.lx.it.pt/~fp/cav/ano2011\\_2012/Trabalhos\\_MEEC\\_2012/Artigo25/Streaming\\_OTT\\_CAV\\_Grupo25\\_Site/OTT.html](http://www.img.lx.it.pt/~fp/cav/ano2011_2012/Trabalhos_MEEC_2012/Artigo25/Streaming_OTT_CAV_Grupo25_Site/OTT.html), acessado a 10/10/15.
- [4] Wang ,Haohong et al. – 4G Wireless Video Communications. 1ª edição. United States of America: John Wiley & Sons, Ltd, 2009, ISBN 978-0-470-77307-9.
- [5] Rafael Romer, “GSM, 3G, EDGE, HPSA, 4G e LTE: entenda as siglas de conexão mobile”, <http://canaltech.com.br/o-que-e/telecom/gsm-edge-hpsa-lte-entenda-as-siglas-de-conexao-mobile/>, acessado a 08/12/15.
- [6] João Ascensão, Codificação escalável de elevada

Granularidade, [http://www.img.lx.it.pt/publications/Theses/JoaoAscenso\\_MSc.pdf](http://www.img.lx.it.pt/publications/Theses/JoaoAscenso_MSc.pdf), acessado a 18/10/15.

- [7] José dos Santos, “Mobile Streaming”, <https://repositorio-aberto.up.pt/bitstream/10216/57907/1/000143374.pdf>, acessado a
- [8] Apple, “HTTP Streaming Overview”, <https://developer.apple.com/library/ios/documentation/NetworkingInternet/Conceptual/StreamingMediaGuide/Introduction/Introduction.html>, acessado a 29/10/15.
- [9] Adobe, “HTTP Dynamic Streaming”, <http://www.adobe.com/products/hds-dynamic-streaming.html>, acessado a 29/10/15.
- [10] Microsoft, “IIS Smooth Streaming Technical Overview”, [http://jvmv.vse.cz/wp-content/uploads/2011/05/t\\_IIS\\_Smooth\\_Streaming\\_Technical\\_Overview.pdf](http://jvmv.vse.cz/wp-content/uploads/2011/05/t_IIS_Smooth_Streaming_Technical_Overview.pdf), acessado a 29/10/15.
- [11] Iraj Sodagar, Microsoft Corporation, “MPEG-DASH: The Standard for Multimedia Streaming Over Internet”, [http://www.bogotobogo.com/VideoStreaming/images/mpeg\\_dash/DASH-IEEE-multimedia-preprint.pdf](http://www.bogotobogo.com/VideoStreaming/images/mpeg_dash/DASH-IEEE-multimedia-preprint.pdf), acessado a 14/11/15.
- [12] Fernando Pereira, “Digital Television: Second Generation”, [http://www.img.lx.it.pt/~fp/cav/ano2014\\_2015/Slides%202014-2015/CAV\\_7\\_Digital\\_TV\\_2nd\\_Generation\\_2014\\_2015\\_Web.pdf](http://www.img.lx.it.pt/~fp/cav/ano2014_2015/Slides%202014-2015/CAV_7_Digital_TV_2nd_Generation_2014_2015_Web.pdf), acessado a 15/11/15.
- [13] Ooyala, “Protecting and Monetizing Premium Content” disponível em: <http://community.ooyala.com/t5/Ooyala-Documentation/Whitepaper-Content-Protection-Landscape-for-Video/td-p/2450>, acessado a 26/11/15.
- [14] Andy, “U.S. COPYRIGHT GROUP” SHUTS DOWN PORTUGAL’S LARGEST PIRATE SITE disponível em: <https://torrentfreak.com/u-s-copyright-group-shuts-down-portugals-largest-pirate-site-150709/>, acessado a 01/12/15.
- [15] Cláudia Bancalero, “Mais 39 sites pirata com ordem de bloqueio pela IGAC”, disponível em: <http://www.publico.pt/tecnologia/noticia/mais-39-sites-com-ordem-de-bloqueio-pela-igac-1713289>, acessado a 01/12/15.
- [16] Sara Ribeiro, “Estudo: 61% dos portugueses vê vídeos em dispositivos móveis enquanto assiste TV”, disponível em: [http://www.jornaldenegocios.pt/empresas/detalhe/\\_estudo\\_61\\_dos\\_portugueses\\_ve\\_videos\\_em\\_dispositivos\\_moveis\\_enquanto\\_assiste\\_tv.html](http://www.jornaldenegocios.pt/empresas/detalhe/_estudo_61_dos_portugueses_ve_videos_em_dispositivos_moveis_enquanto_assiste_tv.html), acessado a 02/12/15.
- [17] Netflix, <https://www.netflix.com/pt/>, acessado a 03/12/15.
- [18] Ericsson, “Ericsson Mobility Report”, disponível em <http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf>, acessado a 03/12/15.



Fábio Freire Rodrigues nasceu a 25 de dezembro de 1994 em Coimbra. Frequenta neste momento o 4º ano do Mestrado Integrado de Engenharia Electrotécnica e de Computadores, no Instituto Superior Técnico, em Lisboa. A sua área principal é telecomunicações, com formação secundária em computadores. É atualmente colaborador nas Jornadas de Engenharia Electrotécnica e Computadores.



Frederico Freire Rodrigues nasceu a 25 de dezembro de 1994 em Coimbra. Frequenta neste momento o 4º ano do Mestrado Integrado de Engenharia Electrotécnica e de Computadores, no Instituto Superior Técnico, em Lisboa. A sua área principal é telecomunicações, com formação secundária em computadores. É atualmente colaborador nas Jornadas de Engenharia Electrotécnica e de Computadores.