Balancing Quality and Energy Efficiency for Mobile Rich Media Service Delivery in a Future Wireless Heterogeneous Network Environment

John Monks

B.Eng in Electronic Engineering

A dissertation submitted in fulfilment
of the requirements for the award of
Doctor of Philosophy (Ph.D.)

School of Electronic Engineering
Dublin City University

Supervisor: Dr. Gabriel-Miro Muntean

January, 2020

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: ____________________ ID No.: _______________ Date: __________

Acknowledgements

First of all I would like to thank my dearest parents, Thérèse and Noel. There is not a day that goes by that I do not think about you. You always wanted me to do the best I can and to get a good education, I can only hope I have done you proud! I would not be where I am today without the opportunities and support you gave me growing up. There are no words to describe how grateful I am to be your son. Love you always.

An extra special thanks to my dearest fiancée Aida Olaru. You have been my biggest supporter and motivation over these past few years, I would not have made it without you. Everything that has happened, everything that we have gone through, you have been there for me every step of the way. Thank you for your unconditional love and support, you have truly made me the man I am today. I am delighted we can finally move forward and start building our life together.

An extra special acknowledgement goes to my amazing supervisor, Dr. Gabriel-Miro Muntean, or Gabi as I have so fondly come to know him. Gabi has been there for me every step of the way with never-ending support, patience, advice and ideas even in the lowest moments. Never have I met a more enthusiastic, energetic and caring supervisor and I am truly grateful to have worked with you. The knowledge and experience I have gained from our chats and time together will stay with me forever. Thank you for guiding me to a successful PhD and for being not only an extraordinary supervisor, but also a great friend.

Special thanks to the friends and colleagues I have made over the years in DCU

PEL. Anderson, Ian and Fabio I believe that each of you have become close friends and I am grateful for all our chats together, even if it was procrastination! They helped me to get through some difficult moments. Fabio, you are next! Best of luck, I know you can do it!

Thank you to all of the other colleagues that I have met and worked with during my time in DCU, just to name a few: Irina, Ting-bi, Longhao, Mohammed, Dan, Diana, Abid, Bharat and Joan. Thank you for our chats and laughter throughout the years.

I am so grateful to Aida's family for being such great, wonderful and supportive people. They have truly welcomed me into their family with open arms and have helped me in every way that they could. Liliana and Petre, you have taken me in like a son of your own, I can never express how grateful I am for everything you have done for me. Andreea, you have been like a sister to me, thank you for all of your support. Adriana, Piero and Thomas, thank you to each of you, our times in Italy together are some of my best memories. And last but certainly not least, Elena or Granny. Thank you for everything, you are an amazing woman and I hope to one day have a conversation with you!

A massive thank you to the Loftus family for always being there for me. Mary, you and John have been like second parents to me and I will always be grateful to you for that. Caroline, Richard and Michael, you have always treated me like a brother and I see each of you in the same way. You have always been there for me and helped me through so much in my life. Thank you all, there is no way I can ever repay you for everything you have done for me.

A special thank you to Shane Carroll. You have been a dear friend to Aida and I since our early undergrad days and have always been there to help and support us. Thank you for everything, I couldn't have chosen a better bestman. Of course a massive thank you to my family, my late grandparents and all my aunts, uncles and cousins. While we may not be as close as we should be, you were all there for me when I needed you. Thank you all so much.

I would like to extend a special thank you to the technical and academic staff of the School of Electronic Engineering in DCU for all the support over the years.

I am truly grateful to Prof. Vlad Popescu, Dr. Marissa Condon and Prof. Patrick McNally for the time they committed to organising and participating in my PhD defence.

This work has been supported by the Irish Research Council Entrerprise Partnership Scheme and Dublin City University, grant number EPSPG/2015/111 and in part by Science Foundation Ireland grant 13/RC/2094 to Lero - the Irish Software Research Centre (http://www.lero.ie) and the European Union's Horizon 2020 Research and Innovation programme under grant no.688503 for the NEWTON project (http://newtonproject.eu).

Dublin, January 2020

John Monks

List of Publications

Journal

- J.Monks, GM. Muntean, 'Quality and Energy-aware Cooperative Distributed Architecture for Mobile Rich Media Delivery in Wireless Heterogeneous Environments', *Submitted for review in IEEE Transactions on Multimedia*, 2019.

Conference

- J. Monks, A. Olaru, I. Tal, GM. Muntean, 'Quality of experience assessment of 3D video synchronised with multisensorial media components,' *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Cagliari, 2017.

- J. Monks, GM. Muntean, 'A Distributed Energy-Aware Cooperative Multimedia Delivery Solution,' *2017 IEEE 42nd Conference on Local Computer Networks Workshops (LCN Workshops)*, Singapore, 2017.

- J. Monks, A.Olaru, GM. Muntean, 'Buffer-Aware Dynamic Adaptive Streaming over Content Centric Networks,' *2019 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, Jeju, 2019.

- J. Monks, CH. Muntean, GM. Muntean, 'A Mobile Quality-oriented Cooperative Multimedia Delivery Solution,' *15th International Wireless Communica-*

tions and Mobile Computing Conference (IWCMC), Tangier, 2019.

Table of Contents

List of Figures

List of Tables

List Of Acronyms

3GPP Third Generation Partnership Project.

3GPP2 Third Generation Partnership Project 2.

aMPD area MPD.

AP Access Point.

ATSC Advanced Television Systems Committee.

CAGR Compound Annual Growth Rate.

D2D Device to Device.

DASH Dynamic Adaptive Streaming over HTTP.

DMB Digital Media Broadcasting.

DTB Direct Television Broadcast.

DTMB Digital Terrestrial Multimedia Broadcast.

DVB Digital Video Broadcasting.

DVB-T DVB Terrestrial.

EDGE Enhanced Data Rates for GSM Evolution.

ENCO ENergy aware COoperative multimedia delivery.

FR Full Reference.

GPRS General Packet Radio Service.

GSM Global System for Mobile Communications.

HAS HTTP Adaptive Streaming.

HDS HTTP Dynamic Streaming.

HLS HTTP Live Streaming.

HSPA High Speed Packet Access.

HSPA+ Evolved High Speed Packet Access.

ICT Information and Communications Technology.

IEEE Institute of Electrical and Electronics Engineers.

IoT Internet of Things.

ISDB Integrated Services Digital Broadcasting.

ISDB-T ISDB Terrestrial.

ISM Industrial, Scientific and Medical.

ITU-T International Telecommunication Union Telecommunication Standardization Sector.

LAN Local Area Network.

LTE Long-Term Evolution.

LTE-A LTE Advanced (LTE-A).

MAC Media Access Control.

MADM Multi Attribute Decision Making.

MENCO A Mobile Quality-oriented Cooperative Multimedia Delivery Solution.

MEW Multiplicative Exponentially Weighted.

MIMO Multiple Input and Multiple Output.

MPD Media Presentation Description.

NR No Reference.

NTSC National Television System Committee.

P2P Peer-to-Peer.

PAL Phase Alternating Line.

PHY Physical Layer.

PSNR Peak Signal to Noise.

QoE Quality of Experience.

QoS Quality of Service.

RR Reduced Reference.

RTCP RTP Control Protocol.

RTP Real-time Transport Protocol.

RTSP Real Time Streaming Protocol.

SECAM Séquentiel couleur à mémoire (Sequential colour with memory).

SSIM Structural Similarity.

SVC Scalable Video Coding.

TCP Transmission Control Protocol.

UDP User Datagram Protocol.

UMTS Universal Mobile Telecommunications Service.

URL Uniform Resource Locator.

VIF Visual Information Fidelity.

VoD Video on Demand.

VQM Video Quality Metric.

VSNR Visual Signal-to-Noise Ratio.

WiMAX Worldwide Interoperability for Microwave Access.

WLAN Wireless Local Area Network.

WMAN Wireless Metropolitan Network.

WPAN Wireless Personal Area Network.

Abstract

Mobile devices have become a popular, almost essential technology in people's lives. They are capable of a large range of services including multimedia streaming and gaming. The performance of these devices is rapidly growing with new generations being released every year. A major limitation to mobile devices is the battery-life with modern devices struggling to survive a full day.

Furthermore, the popularity of social networks and streaming services such as Facebook and YouTube have contributed to a massive global IP traffic growth with mobile devices being one of the main contributors. Advancing performance in devices further stimulates this growth with higher data rates required to meet user and technology demands. Current network solutions are not capable of sustaining such a growth in traffic alone.

This thesis proposes a smart distributed cooperative approach for rich media delivery to heterogeneous mobile users in wireless environments. A cooperative distributed approach promotes reuse of data, takes traffic away from the network and allows extendibility. Achieving the vital balance between performance and energy is challenging in such an environment. The solutions proposed in this thesis to solve the problems above are:

1. **The ENergy aware COoperative multimedia delivery solution (ENCO)** monitors device performance in order to dynamically adapt requested quality to balance the energy cost and quality level in a cooperative scenario. Host

selection is performed based on distance and load of the hosts.

2. **The Mobile aware quality-oriented COoperative multimedia delivery solution (MENCO)** proposes a novel distributed host selection mechanism that considers any mobility and selects the best host based on the velocity of travel. Hosts are selected so that they are within range for the required amount of time to successfully deliver the data.

Chapter 1

Introduction

The Information and Communications Technology (ICT) industry is rapidly evolving with mobile devices becoming increasingly powerful and more services available than ever before. Some of these services (e.g. YouTube, Netflix, Facebook etc.) provide video sharing which can put a strain on existing networks and devices. These services are providing so much content that current networks are experiencing a massive growth in traffic. Modern delivery solutions need to be capable of delivering the large amounts of data with minimal degradation in quality. Furthermore, devices have to process this content in an efficient manner that will not impact the overall experience of the user. This is especially important for mobile devices which are growing more powerful while battery technologies are lagging behind. This growth in performance and limitation in battery life as well as the large strain being put on networks are major problems. The research community has been actively pursuing solutions to help with these issues. However, there are a number of issues and challenges that still require much attention. This thesis focuses on solutions that will allow the explosion of traffic to be supported by available technologies while maintaining a high user experience in heterogeneous mobile wireless environments. The motivation and goals behind the work are introduced in this chapter followed by an outline of the main contributions and the structure of the thesis.

1.1 Research Motivation

The latest forecast by Cisco [1] predicts that the annual global IP traffic will reach 3.3 zettabytes by 2021. This will be an increase of 24% Compound Annual Growth Rate (CAGR) between 2016 and 2021. Of this, 82% is expected to be IP video traffic according to the Cisco technical report [2]. Furthermore, it is expected that there will be 3.5 networked devices per capita by 2021 with mobile traffic growing at twice the rate of fixed traffic. As such, the number of mobile subscriptions is expected to reach 9 billion by 2022 according to the 2017 Ericsson Mobility Report [3]. This clearly shows the pressure being put on the currently available network technologies. This large demand for connectivity has pushed the industry to accelerate the 5G R&D efforts.

With this rapid growth in traffic, and the slow advancement of networking solutions, it is vital to find new, more efficient approaches involving current network solutions. The current network architectures are not capable of supporting such an explosion in traffic and the proposed specifications of 5G are still not enough to support emerging technologies such as AR/VR on a large scale. Furthermore, while the capability of mobile devices is rapidly growing, following a trend similar to Moore's law, battery technologies are growing slowly with no recent major advancements. The energy density of Lithium-Ion batteries has only been increasing annually by an average of 5% [4] as reported in 2016. While the capabilities offered by the state-of-the-art Lithium-Ion batteries has exceeded what was originally expected, the technology will soon reach its limit [5]. This slow advancement and limited capacity is leading to a bottleneck forming with regards to performance and energy for mobile devices.

In this regards, there is a need for a solution that is capable of alleviating the strain being experienced by network infrastructures. Furthermore, user experience is essential and must be considered. It is not enough to give the highest quality video to a user, devices need to be able to perform beyond the needs of a single video. There is a further need for smart balancing between quality and energy during multimedia streaming on battery powered devices.

1.2 Problem Statement

With the massive growth in traffic comes many issues for network technologies and providers. It is not feasible for any single network or provider to support the vast amount of video content being exchanged. Going forward, there is a need for smart solutions that allow cooperation between various technologies to enable stable, secure and **high quality data transfer**. Furthermore, it is necessary for seamless transitions between networks (e.g. a local WiFi network to Long-Term Evolution (LTE)) to allow users to always be connected to their favourite services, wherever they may be. Not only do the networks need to co-exist, it is important to leverage all available resources. With the staggering amount of personal devices capable of wireless communication, and the growth of Device to Device (D2D) communications, it is essential that **cooperation** is leveraged. Not only will this help to alleviate the pressure on the current network infrastructures, it will also expand the reach of the current services and provide more options to users. Of course, cooperation introduces additional challenges and it is a largely overlooked avenue with considerable potential. One considerable challenge that must be considered is the really dynamic movement of mobile devices which makes their performance hard to predict and services they may support potentially unreliable.

The research community largely focuses on improving the experience of the user by increasing the quality of the multimedia content and the rate at which it can be received. However, more focus needs to be given to the **energy** concerns of the user devices. Many users today are consuming content using some form of mobile wireless connected devices. The nature of these devices entails a battery for operation which naturally leads to energy constraints. Unfortunately, while the performance of these devices is increasing, battery technologies are rapidly lagging behind. Furthermore, this rise in performance leads to higher energy costs in the form of CPU processing, displays, network connectivity etc. Short of significantly increasing the footprint of these devices to house larger batteries, other avenues need to be explored in order to minimise the energy costs and prolong the battery life. In this thesis the network load and user experience are considered. More precisely, cooperation is leveraged to offload load from the network while the underlying devices are capable of self management and maintain a balance between energy and quality. Metrics such as load, distance, energy and throughput are utilised to develop novel solutions for multimedia adaptation and cooperative host selection.

1.3 Thesis Contributions

Focusing on the problem of achieving high quality video delivery in a context with node mobility and energy constraints, this thesis proposes two novel solutions. Their contributions to extending the state of the art are described in the context of the solutions as follows:

- A **Distributed Energy-Aware Cooperative Delivery (ENCO)** solution is proposed for adaptive multimedia delivery in a **cooperative** distributed environment:
  - A new adaptive mechanism is introduced which achieves a balance between quality and energy consumption by user device. This has not

been done before in a cooperative distributed scenario between multiple mobile devices.

  - An innovative neighbour discovery and host selection method is deployed to allow user devices to find and communicate with other users in order to request and share content. Based on prior knowledge, the best host is selected by considering the distance and load experienced in order to improve performance and fairness.
  - Testing involving video delivery shows that ENCO outperforms comparative studies by maintaining a high quality of experience for the user while achieving a balance between energy and quality

- A **Mobile Quality-oriented Cooperative Multimedia Delivery (MENCO)** solution is proposed to bring novel **mobility awareness** to a cooperative video delivery while balancing energy and quality.
  - A novel mobility-aware cooperative adaptive multimedia delivery solution which balances quality and energy consumption
  - The novel mobility awareness method introduced observes the velocity of the user in relation to potential hosts and considers the available time for communication between devices
  - Testing shows that this introduced awareness reduces delay and loss when requesting content and in comparison with other solutions achieves improved quality-energy balance

1.4 Report Structure

- Chapter 1 introduces the research topic and motivations as well as the expected goals of the work. It also presents the major thesis contributions

- Chapter 2 discusses the relevant background technological information required to understand the work discussed in this thesis.
- Chapter 3 performs an in-depth review of the latest relevant works in the literature and identifies their strengths and limitations.
- Chapter 4 presents a detailed description of the proposed solutions, including the architecture and all of its components and the novel ENCO and MENCO algorithms.
- Chapter 5 describes the testing employed and discusses the results in the context of the state of the art.
- Chapter 6 concludes the thesis with a summary of the work and directions for future work.

Chapter 2

Technical Background

In this chapter the technical background for this thesis is presented. A view of wireless networks is first discussed, looking at the various 802 groups established followed by the history of the popular 802.11 standard. The history of cellular is also presented alongside a brief discussion of broadcasting and satellite technologies. Next, the concept of adaptive video delivery is discussed following an overview of the history of video delivery. Video quality assessment methods are then introduced and the different methods available.

2.1 Wireless Networking

2.1.1 Wireless Broadband Networks

Three main standards have been defined by the IEEE which are categorised as broadband networks, 802.11 WLAN, 802.15 WPAN and 802.16 WMAN/WWAN. This thesis will primarily focus on Wireless Local Area Network (WLAN), however the other technologies will be briefly touched upon here.

2.1.1.1 Wireless Personal Area Network

The IEEE established the 802.15 group to define a set of standards to specify the Physical Layer (PHY) and Media Access Control (MAC) requirements for developing Wireless Personal Area Network (WPAN) networks. A WPAN network is a network established when different devices in a local area can communicate with each other without the use of an infrastructure. These networks typically only cover a few meters at the most and are usually found with personal devices such as headphones, watches, remote controllers etc.

Two main applications have been established, ZigBee [6] and Bluetooth [7]. The former technology uses radio waves to communicate on a line of sight basis and can reach distances of 10-20m. The communication is kept simple with little set-up required while also being secure. Furthermore, ZigBee achieves low operational costs and low power consumption for bit rates of up to 250Kbps. This makes it ideal for applications which require occasional bursts of data such as a light switch or sensor.

Bluetooth on the other hand is developed for more personal applications such as wireless headphones, speakers and game controllers. It can reach greater distances of up to 100m with higher energy consumption if necessary and does not require line of sight. Bit rates of up to 2Mbps can be achieved using the latest Bluetooth 5 technology. Future work is showing lower energy costs and even wider ranges being proposed. Additionally, higher rates are being achieved for Bluetooth technologies.

2.1.1.2 Wireless Metropolitan Area Network

The IEEE 802.16 group was created in 1999 to define the standards based on the development of Wireless Metropolitan Network (WMAN) networks which were first published in 2002. These networks are much larger than WPAN and even WLAN networks. They involve connecting a large geographical area and typically incorporate numerous local networks into one WMAN. Some examples of this would include a campus scenario where each building has a Local Area Network

(LAN) or WLAN and these buildings are connected under a campus network or WMAN using a network backbone. Worldwide Interoperability for Microwave Access (WiMAX) [8] is an established implementation which showed great promise when it was first released. It has achieved very wide ranges with rates of 30-40 Mbps initially and recent mobile rates of up to 1 Gbps. This solution was a candidate for the 4G proposal and was surpassed by LTE.

2.1.1.3 Wireless Local Area Network

The IEEE established the 802.11 group to define a set of specifications for PHY and MAC for the creation of WLAN and simpler implementation when compared to the costly creation of a LAN which requires potentially large amounts of cables to establish an infrastructure. The first specification was released in 1997 [9] but did not initially gain popularity due to the low bit rates when compared to Ethernet [10]. The first amendment came in 1999 with the release of 802.11b [11] which gained rapid success from the industry. With rates increasing from 1-2Mbps up to 11Mbps, the cost of set-up became viable and was widely adopted.

The 802.11 group has continued releasing amendments and updates. Five distinct frequency ranges have been predominantly used across the releases, 2.4, 3.6, 4.9, 5 and 5.9 GHz. Two recent releases have moved to new frequency bands, 802.11ad [12] has moved to the 60GHz band while 802.11ah [13] has moved down to the sub 1 GHz range. To be more specific, 802.11ad achieves very high rates of up to 7 Gbps in exchange for different propagation characteristics that see the waves reflecting from more surfaces due to the higher frequency. On the other hand, 802.11ah is aimed at the growing Internet of Things (IoT) trend and uses the 900 MHz band. This decision allows greater penetration through materials with lower power consumption which makes it ideal for sensors which can be located in many places and require little maintenance.

Table 2.1: IEEE 802.11 Major Releases

Release	**Year**	**Frequency (GHz)**	**Upper Data Rate (Mbps)**
a [14]	1999	5	54
b [11]	1999	2.4	11
g [15]	2003	2.4	54
n [16]	2009	2.4/5	600
ad [12]	2012	60	6757
ac [17]	2013	5	3467
ah [13]	2016	0.9	8.67
ax [18]	2019	2.4/5/6	10530
ay [19]	2020	60	20000

Two soon to be released amendments are 802.11ax and 802.11ay which will be successors to 802.11ac and 802.11ad respectively. 802.11ax has a goal of improving throughput performance of ac 4-fold and is designed to operate across all Industrial, Scientific and Medical (ISM) bands between 1 and 6 GHz. 802.11ay on the other hand is being developed to succeed the high frequency 802.11ad release to further improve the high throughputs achieved by ad. These rates come with cost of low penetration due to the characteristics of high frequency waves, just like 802.11ad. However, 802.11ay could potentially replace Ethernet for indoor uses with theoretical rates of up to 20Gbps.

These are just a few of the latest amendments. There have been many releases since the group was first established with 802.11 a/b/g/n/ac [11], [14]–[17] being some of the main ones. Each amendment introduces a change that may be a fix on a previous release or a new feature. There have also been a number of full releases, most recently in 2016, that bring all of these amendments together into the newest version of the specification. A table summarising the main features of each of the amendments can be seen in Table 2.1.

Two main modes are available under 802.11 WiFi standard, infrastructure and ad-hoc. In infrastructure mode, devices in the network communicate with each other through previously established Access Point (AP) as seen in Figure 2.1. These

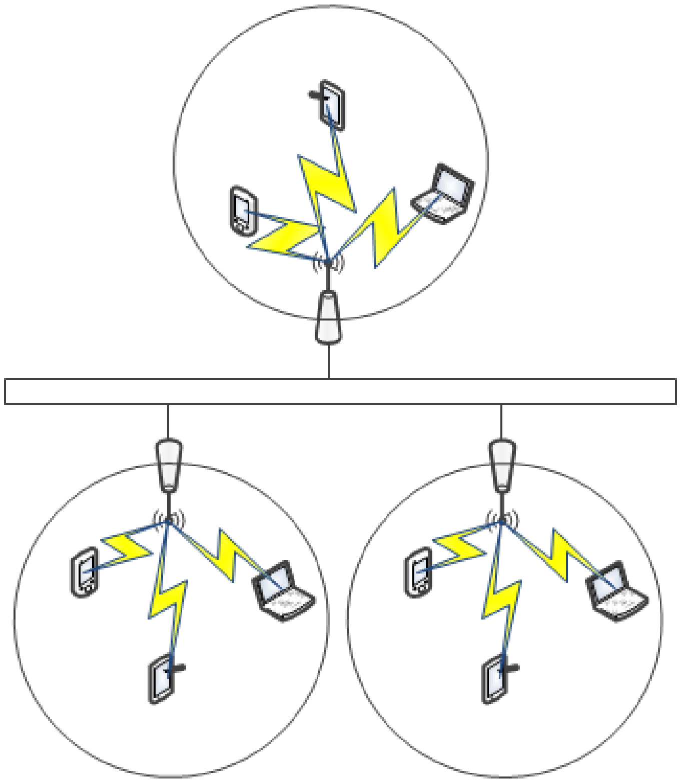

Figure 2.1: 802.11 Infrastructure Mode

points are crucial in infrastructure mode as they allow the devices to access the network. This comes with the added benefit of a wider network and more stable performance. However, it requires a potentially large cost to set up and is usually static and cannot be easily moved once created.

Ad-hoc mode allows devices to communicate with each other directly without the use of an intermediate AP as shown in Figure 2.2. This allows a network to be rapidly created among neighbouring devices and easily managed. Furthermore, this can allow for a network to scale easily and the network can be moved. However, a number of issues are also involved with this mode. Due to the lower power of the transmitters located on personal devices, the communication range is shorter. If a device wants to communicate with another device outside of its own range, multiple "hops" will be necessary to send the message between neighbour-

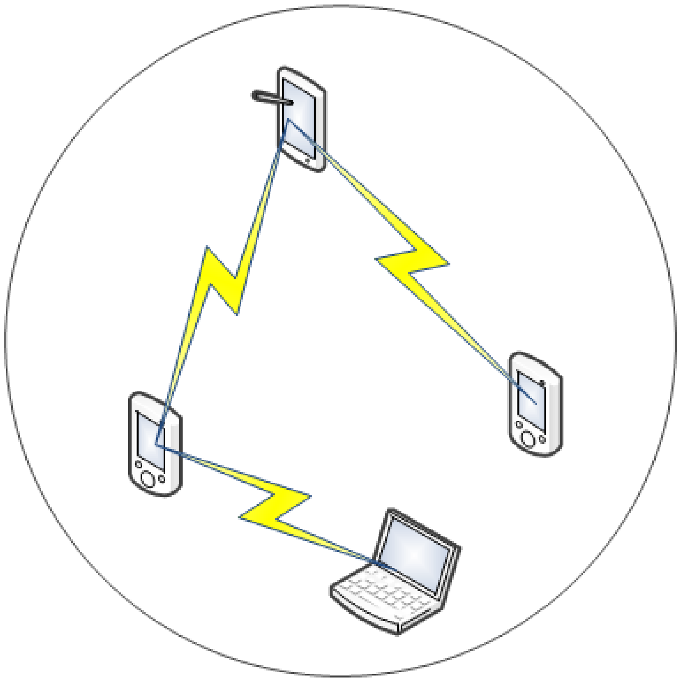

Figure 2.2: 802.11 Ad-Hoc Mode

ing devices, which is significantly slower than through an AP. Furthermore, the energy of the personal devices suffers due to the additional transmissions. Both modes have advantages and disadvantages which can be leveraged to provide innovative solutions that can potentially answer all demands.

2.1.2 Broadcast Networks

Broadcast networking focuses on the distribution of media to a large audience. In the past, analogue broadcasting technologies have been used for distributing television and radio media with three main technologies dominating the industry, National Television System Committee (NTSC) [20], Phase Alternating Line (PAL) [21] and Séquentiel couleur à mémoire (Sequential colour with memory) (SECAM) [22]. Recent years have seen the birth and growth of a new technology, Direct Television Broadcast (DTB). Four main solutions have emerged from this technology, Digital Video Broadcasting (DVB) [23], Advanced Television Systems Committee (ATSC) [24], Integrated Services Digital Broadcasting (ISDB) [25] and Digital Terrestrial Multimedia Broadcast (DTMB) [26]. Each of these solutions strives to achieve high quality and efficient delivery of the media but take different, non-compatible, approaches that have led to separate groups forming around each solution. DVB has become one of the most widely adopted solutions worldwide with early deployment in Europe, Singapore, New Zealand and Australia. Meanwhile, the United States and Canada have stuck with ATSC as their primary broadcasting technology with the FCC declining a move to DVB Terrestrial (DVB-T) [27], DVBs initial terrestrial standard. ISDB has been primarily used in Japan and the Philippines. However, a recent update, ISDB Terrestrial (ISDB-T) International is gaining popularity in South American countries as well as Portuguese speaking African countries. DTMB and a related solution, Digital Media Broadcasting (DMB) [28], only have a presence in China and South Korea respectively.

2.1.3 Cellular Networks

Cellular communications have become a vital part of networking. They allow mobile devices to keep connected even when away from an established network or moving at high speeds. However, the technologies we take for granted today have grown from very simpler concepts over the past decades. This section will discuss this history and touch on the achievements of each generation.

2.1.3.1 First Generation (1G)

The first cellular generation emerged in the 1980s and used analogue signals for sending and receiving data. The so called "1G" technology was created to achieve automated mobile voice communication to replace the manual solution previously in place. Some examples of deployed 1G solutions include the Nordic Mobile Telephone (NMT) [29] mainly in Nordic countries, Advanced Mobile Phone System

(AMPS) [30] originally in America and the Total Access Communications System (TACS) [31] in the United Kingdom. The transmission of voice data was not encrypted which posed a major issue. Anybody with access to a scanner device could easily intercept a conversation. As a result of this, when the second generation emerged with encrypted transmission and more, the 1G technologies where largely replaced.

2.1.3.2 Second Generation (2G)

The second generation of cellular technology was a major change from 1G. Signals were no longer analogue, but were changed to digital which opened a large range of possibilities and advantages. Communications were now encrypted and more services were available beyond voice including text, picture and multimedia messages. Two major standards emerged under 2G, Global System for Mobile Communications (GSM) [32] and cdmaOne [33]. GSM was adopted nearly everywhere but the US. At this early stage, circuit switching was the primary means of routing in GSM which was very inefficient. Each user was allocated to a particular circuit which was very limiting. Due in part to this inefficiency, GSM was originally intended for duplex telephony only. However, two updates to the standard extended the potential of GSM going towards the third generation (3G).

General Packet Radio Service (GPRS) [34] was the first major update to GSM which was flagged as 2.5G. Routing was updated to a packet switching solution which allowed for packets to be individually routed across the network. This eliminated the inefficiency involved with circuit switching. As such, GPRS was able to achieve significantly better performance than the base GSM.

The second update to be introduced in 2G was the Enhanced Data Rates for GSM Evolution (EDGE) [35]. A new encoding technique, 8PSK, was introduced in this update which tripled the capacity of transmissions. This improvement made

EDGE nearly comparable to the 3G specifications at the time while still allowing backwards compatibility with GSM. Rates of up to 200 Kbps which officially suited the requirements of the 3G specifications. However, due to the limited bandwidth of GSM, higher rates proved to be to difficult to achieve leading to new technologies being proposed for the new generation.

2.1.3.3 Third Generation (3G)

In 1998, two separate groups where formed to develop 3G solutions based on the 2G technologies. The Third Generation Partnership Project (3GPP)[1] group was established to develop a 3G technology based on GSM. A second, completely separate group, the Third Generation Partnership Project 2 (3GPP2)[2] was created to develop upon cdmaOne in the US. From these groups, two main 3G technologies emerged, Universal Mobile Telecommunications Service (UMTS) [36] and CDMA2000 [37]. Only UMTS will be discussed further in this thesis.

With a theoretical rate of 384 Kbps, the original UMTS proposal exceeded the required rate of 200 Kbps stated by the 3G specification requirements. One of the main contributing factors was the change to the WCDMA air interface which allows for higher bit rates to be carried. As with 2G, two additional updates where released for the 3G technology. The first was called High Speed Packet Access (HSPA) [38] with theoretical speeds of up to 337 Mbps. However, in practice HSPA typically reaches rates of 14 Mbps. The second release was an update to HSPA called Evolved High Speed Packet Access (HSPA+) [39]. With this release, 3G speeds of up to 42.4 Mbps were achievable. HSPA+ is seen as a 3.75G release and allowed companies to bridge the gap between 3.5G and 4G without immediate investments. Furthermore, HSPA+ saw the introduction of new concepts such as beam-forming and Multiple Input and Multiple Output (MIMO), both of which

[1] http://www.3gpp.org/
[2] http://www.3gpp2.org/

would go on to be included in 4G solutions.

2.1.3.4 Fourth Generation (4G)

The current generation is 4G. When the specifications were being created, there were two main contenders for the dominant 4G system, LTE [40] and WiMAX [8]. LTE was developed by the 3GPP group and, despite not meeting the 4G specifications set out, was chosen as the main 4G technology. This choice was made due to industry pressure for better solutions. With a peak rate of 100 Mbps, LTE falls significantly short of the required rate of 1000 Mbps or 1 Gbps for low mobile users. However, an update came in the form of LTE Advanced (LTE-A) (LTE-A) [41] which managed to achieve the specified bit rate for the 4G standard.

2.1.3.5 Fifth Generation (5G)

The requirements outlined by International Telecommunication Union Telecommunication Standardization Sector (ITU-T) in [42] describe what is expected from the final 5G standard when it is due to be completed in 2020. Some of the components have already been finalised with the 5G New Radio (NR) specification released in 2017. The explosion of data traffic being experienced due to the growth in social media and content quality led to 5G NR being rushed and the industry is now working on implementing it. While originally the specification for 5G was to have a speed of 20 Gbps, 3GPPs proposal only slightly improves upon the performance of the existing 4G solutions.

2.1.4 Satellite

Satellite communications [43] are used for transmitting data across very large distances, even between countries. Three components are required as part of the network, a satellite in orbit, a gateway for relay data from the ground to the satellite and a dish capable of receiving the data from the satellite at the users location. Due to the line of sight nature of satellite communications, this technology is not suited to city environments with large structures and/or vehicles which may block satellite dishes. Any form of blockage would cause the connection to be interrupted and lost between the dish and the satellite. Therefore, this form of communication is useful for isolated locations with little connectivity and few potential blockages.

2.2 Video Delivery Solutions

Streaming is a method of receiving media content over a network. It allows the receiver to start playing the content immediately before all of the content is received. This is an alternative to the more traditional method of downloading media fully and playing it back at a later time. The quality of a stream is dependent on a number of contributing factors such as the state of the network and the condition of the devices involved in the transaction. Here the history of streaming will be briefly described, from the initial non-adaptive functionality to the more recent adaptive paradigm.

2.2.1 Non-Adaptive Solutions

Non-adaptive video delivery means that the content is transmitted at a constant quality and will be received in a potentially degraded condition based on the state of the network. The original video delivery solutions worked in this manner and simply tried to deliver the content. The video was stored on a server device and encoded before being sent across a network to a user. StarWorks [44] was the first streaming product, released in 1992. It worked over a corporate Ethernet network and allowed users on the network to stream video encoded using MPEG 1 [45] on demand.

Various protocols have been used over the years for delivering video. User Datagram Protocol (UDP) was the most simple solution to implement as it allowed the video to be sent quickly allowing the network to achieve the necessary rates to maintain playback. However, due to the "fire and forget" nature of UDP, it was inherently unreliable. Packets could be lost without any notice, and the order of delivery was not guaranteed meaning that additional complexity was required at the client side. Transmission Control Protocol (TCP) on the other hand was very reliable as packets were sure to be delivered. However, as a result of this reliability, it is much slower than UDP. If a packet happens to be lost, the server will re-transmit after a given time, causing potentially significant delays and pauses in playback.

An additional protocol, Real Time Streaming Protocol (RTSP) [46] was introduced that brought together the advantages of UDP and TCP. RTSP was developed by combining two other proposed protocols to achieve fast and reliable delivery. The Real-time Transport Protocol (RTP) [47] was proposed to enable fast transport of video content over UDP. Secondly, the RTP Control Protocol (RTCP) [48] was further proposed to monitor the delivery of content over the RTP protocol. RTCP does not deliver the content itself, but instead provides statistical feedback to the server, enabling TCP like behaviour over a UDP network. RTSP brings together the functionality of RTP and RTCP to provide reliable and fast performance.

Two main forms of streaming emerged as more capable solutions became available, live streaming and Video on Demand (VoD) streaming. Live streaming involves the video content being streamed to the viewer as it is being recorded. In this case, a capturing device (e.g. camera, webcam etc.) is needed to capture the video. As it is captured, it is streamed via a computer across the network to any and all viewers. It is optional whether the source device records it or not. Twitch [3]

[3] https://www.twitch.tv/

is a modern example of a live streaming service which is primarily used for streaming video game content.

The more traditional form of streaming is known as Video-on-Demand (VoD). In this situation, the video content is stored on a server device and is made available to users to stream. A client can choose to view the video whenever they want and can skip forward, rewind and re-watch the content as much as desired. This form of streaming has become very common with many websites offering VoD services including YouTube [4], Facebook [5], Netflix [6] and Amazon [7].

Due to the nature of video streaming, it is very susceptible to delays. Playback starts while the content is being received by the user. This leads to the possibility of pauses occurring during playback because of losses or delays on the network. Buffering has been introduced to help counteract this issue by allowing a buffer to build up extra content before starting playback. This allows player to capable of compensating for occasional losses and delays. However, this is still far from ideal. In order to sustain high quality video playback, a stable bit rate is required which is not always possible. As a result, adaptive solutions have been introduced to try to solve these issues.

2.2.2 Adaptive Solutions

Adaptive video delivery is vital for delivering the best Quality of Experience (QoE) to the user. Many different approaches have and still are been investigated to achieve the optimal experience in video delivery. These solutions typically monitor various attributes of the video delivery system, including throughput, buffer capacity, packet loss and device conditions. Based on the measured values, the stream can be adjusted to maintain a high user quality. Here, two main approaches

[4] https://www.youtube.com/
[5] https://www.facebook.com/
[6] https://www.netflix.com/
[7] https://www.amazon.com/

are introduced, Scalable Video Coding (SVC) [49] and HTTP Adaptive Streaming (HAS) [50].

SVC is an extension of the latest encoding standard, H.264 [51]. SVC encodes the source video by using a layering technique where the base layer is the lowest quality and each subsequent layer adds more packets to achieve higher quality. Each layer is associated to a particular stream that the client can request. These streams allow the client to request the content at different quality by requesting more or less layers.

HAS is a different approach to adaptive streaming that has been gaining popularity. With HAS, the video content is encoded at different quality levels and segmented into sections of equal temporal length. These segments are stored at the server side and described in a XML manifest. The client can then request the video content, segment by segment, depending on the desired quality level. The MPEG group has defined a standard called Dynamic Adaptive Streaming over HTTP (DASH) [52] which is the first adaptive streaming standard. DASH is agnostic to the encoding standard used for the clips and does not define a specific adaptation mechanism for the client device. Instead, it describes the format for the manifest and best practices for handling the segments.

Besides DASH, a number of companies have released their own interpretation of HAS. Apple has released HTTP Live Streaming (HLS) [53], Microsoft has Smooth Streaming [54] while Adobe has developed HTTP Dynamic Streaming (HDS) [55]. Each of these proprietary technologies belongs solely to the corresponding company and can change drastically depending on that company. Alternatively, the DASH standard has been created by the MPEG group with support from various companies including Netflix, Microsoft and Apple. While the features of DASH have not yet been fully explored by the research community or industry, it is gaining widespread support with companies including Netflix and Youtube integrating it into their video players.

2.3 Video Quality Assessment

In order to deliver the best experience to the viewer, it is necessary to measure the impact of various factors on the users enjoyment levels. As a result of this, various methods and metrics have been developed over time to both accurately and quickly measure enjoyment levels during multimedia experiences. Two types of assessment are commonly practised, subjective and objective, each having its own advantages and disadvantages. These will be explored further in this section.

2.3.1 Subjective Assessment

Subjective quality assessment involves gauging how a user experiences a video. To perform a subjective test is time consuming and can be costly. It involves setting up a set of test clips with different qualities aimed at studying the impact of a certain attribute (such as packet loss). These are then presented to a test group of people in a controlled environment. Feedback is gathered from all the participants and the results are analysed. Based on this analysis, a model can be determined that shows the impact of the attribute of interest on the participants experience. Obviously, this is very time consuming. Furthermore, if the test group is insufficient, the results may be skewed. As a result of this, the ITU-T has published a number of standards and recommendations around the best practices required when performing subjective tests. ITU-T P.910 [56] and P.911 [57] describe the best practices for the non-interactive subjective assessment of multimedia quality. ITU-T P.913 [58] recommends methods for performing non-interactive subjective assessment for audio, visual or audiovisual quality in various environments such as internet video.

Despite the large cost, if performed properly, a subjective test can provide valuable data. Not only are the results directly from human opinions, accurate models can be established to map these opinions to the underlying attributes that caused the results. Because of the large cost of these tests they can not be quickly and/or reliably repeated for different input parameters. However, if properly performed, a model can be established which can be repeatedly used to estimate human opinions for differing inputs. This idea has led to the emergence of Quality of Experience (QoE) [59] which is the measure of a user perceived enjoyment.

2.3.2 Objective Assessment

An alternative approach to subjective quality assessment, objective assessment depends on mathematical models to determine the experience of a user. These measurements are much quicker to perform and will always give the same output for any particular input. Therefore, objective metrics can be implemented directly into dynamic adaptive mechanisms. The Quality of Service (QoS) [60] was originally associated to objective metrics. QoS is the measurement of a users experience to a particular service based on the conditions of the system, and not the user themselves.

There are three main approaches taken when considering objective metrics, full reference, reduced reference and no reference. Full Reference (FR) involves comparing the received media to the source content frame by frame. This is a very accurate approach, but is very inefficient. Reduced Reference (RR) is more efficient but not as accurate. RR only considers some frames from the source and received content. Finally, No Reference (NR) requires no source material. As a result of having no comparative, it is the least accurate. Instead, it analyses the received content directly. One method involves checking the media, pixel by pixel, for any errors or artefacts. Another introduces headers that provide information without the need to decode the media. Hybrid approaches are also taken, all of which are more efficient than NR or RR but less accurate.

Some metrics which were originally used for picture quality assessment have been adapted for video assessment. Peak Signal to Noise (PSNR) [61] is one of the most well known objective metrics which was originally a FR model. It compares, image by image, the expected signal to the measured error occurred by the received image. Peak Signal to Noise (PSNR) can be easy to implement and can give a reasonable understanding of the state of playback, however it is poor at estimating the users experience. Structural Similarity (SSIM) [62] is another model which is a FR metric. It compares the reference and received images section by section where each section is a window of a predefined size. While SSIM is more accurate than PSNR, it is also more intensive to perform. Additional metrics that can be considered include Visual Signal-to-Noise Ratio (VSNR) [63], Video Quality Metric (VQM) [64] and Visual Information Fidelity (VIF) [65]. These metrics are not currently considered in the scope of this research due to their intensive nature.

2.4 Chapter Summary

This chapter presents the evolution of wireless communications including broadcasting, broadband, cellular and satellite. The generations and versions within each category are discussed with emphasis given to the most popular solutions. Following from this, video delivery and how it has changed is discussed. Video quality assessment methods are then introduced and discussed with various standards and metrics highlighted.

Chapter 3

Literature Review

This chapter presents a review of the latest research on the following topics: multimedia adaptation solutions and content delivery mechanisms. The existing solutions and metrics are discussed and compared. Challenges and open areas of research are highlighted with regards to energy efficiency for heterogeneous mobile devices in wireless environments.

3.1 Multimedia Adaptation

Multimedia delivery over a network is sensitive to the ever changing conditions experienced by the network over time. Furthermore, different users will have different requirements based on their personal preferences, device characteristics, location and network access. From these requirements, adaptive solutions are widely utilised across research and industry in order to deliver the best experience to the user. These solutions take different approaches to adapting the content. The remainder of this section is divided into the following topics based on awareness: performance and energy.

3.1.1 Performance Aware Solutions

In this section, the solutions consider the performance of the device as part of their adaptive mechanism. The solutions include: buffer length management that strives to maintain a healthy buffer by adjusting the rate of the requested content to meet network conditions; bandwidth and throughput monitoring which reduce the requested quality if there is not enough bandwidth or throughput to receive the desired rate; hybrid solutions which combine multiple approaches such as buffer and throughput monitoring; and device aware solutions which consider the characteristics of the users device.

Beben et al. [66] propose an Adaptation & Buffer Management Algorithm (ABMA+). HTTP adaptive streaming is used as part of this solution. The authors use the probability of re-buffering to estimate the download time of future segments. This probability is used in conjunction with pre-computed buffer maps to determine the best video quality to be requested. The selected quality is the highest quality that can be downloaded in the estimated time therefore avoiding freezes. By using pre-computed maps for analysing the buffer, good performance is achieved with little overhead. This work does not consider a mobile scenario or energy consumption.

Spiteri et al. [67] proposed a bitrate adaptation algorithm alongside an online control algorithm called BOLA for video streaming. The bitrate algorithm is a utility maximisation problem that considers rebuffering and quality. An increase in rebuffering time leads to a lower utility while an increase in quality is reflected by a higher utility. Furthermore, the online adaptation algorithm called BOLA uses Lyapunov optimisation to minimise buffering and maximise quality. A near optimal utility can be achieved with the online BOLA controller. This solution relies on the online controller and does not consider the energy cost.

Vergados et al. [68] propose a fuzzy logic rate adaptation scheme for MPEG-

DASH. FDASH aims to efficiently adjust video rate to suit network conditions while avoiding buffer underflow and frequent quality changes. These conditions have been shown to be two of the major degradations to user QoE. Results show that the solution not only provides optimal video rates, but also avoids underflow and unnecessary changes. Despite the objective improvement, no proper QoE metrics are considered as part of this work.

Yin et al. [69] proposed a control theoretic based solution to DASH adaptation. The solution combines two different adaptive approaches, rate control and buffer capacity. These two approaches are widely disputed among the adaptive research community for their advantages and disadvantages. By combining the benefits of both approaches, into a optimised Model Predictive Control (MPC) approach, significant performance improvements are achieved over state-of-the-art algorithms and industry standards. This work only considers basic throughput prediction via a harmonic mean approach. Furthermore, only a single player is focused on with no consideration given to fairness.

Jiang et al. [70] presented a study of bitrate adaptation. From their findings, a framework for creating adaptation mechanisms is proposed. Following their own framework, the FESTIVE algorithm is proposed. A range of issues are improved upon including fairness, efficiency and stability. FESTIVE is designed following the proposed logic to achieve a trade-off between the mentioned issues. This solutions outperforms other methods with improvements of at least 40% for fairness, 50% for stability and 10% for efficiency. This work does not consider heterogeneous environments (different players or even traffic types), mobility or energy.

Li et al. [71]proposed a client side rate control algorithm for HAS called "Probe and Adapt" or PANDA. The authors aim to solve the fair share issue that emerges when numerous HAS streams are sharing the same channel, causing a bottleneck. PANDA uses an application layer solution to probe the network to judge the clients fair-share bandwidth and adapts the requesting bit rates accordingly. This solution effectively addresses the problem of bitrate oscillation often seen in rate control solutions. This work however does not consider mobile scenarios where probing may not be a viable option due to the ever-changing nature of the network.

De Cicco et al. [72] propose ELASTIC, a control theoretic approach to adaptive streaming. ELASTIC is a client side controller which avoids the problematic on-off pattern of HAS and aims to fairly share the bandwidth bottleneck with other streams. Other approaches typically use two separate controllers to control the video level and buffer length. The authors design a single feedback controller to throttle the video level in order to drive the buffer length to a certain point. By using one controller the idle time between segments is reduced and the available bandwidth is matched by the video level when the set buffer length is achieved. This solution outperforms FESTIVE and PANDA by obtaining a fair share of channel resources in all scenarios. Mobility and energy are not considered in this work.

Juluri et al. [73] propose the segment aware rate adaptation (SARA) algorithm. This novel approach considers the different sizes of segments between quality levels when estimating download time for future segments. Higher quality segments will require more data than lower levels which is rarely considered in adaptation. The authors show good performance with higher video quality and less quality changes. However, they do not consider suitable metrics or simulation models. Furthermore, they only consider quality in their algorithm. Despite the limited analysis performed by the authors, the segment aware approach is of particular interest for its novelty.

Mok et al. [74] proposed an early solution, QDASH. A two phased system for DASH that is QoE aware was proposed in this work. A measurement proxy is introduced which "hijacks" the data streams going from server to client. The proxy then transmits the packets according to the QDASH probing technique. Furthermore, the proxy can provide measurements to the clients so they can adapt their requests. The quality of the video is gradually adjusted based on the network conditions. Intermediate quality levels are introduced to provide a gradual change between levels. Through subjective testing, this gradual shift was shown to improve user QoE, when compared to abrupt changes.

Various metrics have been proposed throughout the literature in order to best measure the quality experienced during playback. These metrics are evolving with the growth of the QoE paradigm from the more traditional QoS metrics which only consider objective measurements. Three main types of metric exist in the literature, Full Reference (FR), Reduced Reference (RR) and No Reference (NR). As implied by their names, each type has a certain amount of dependence on a reference model. See Section 2 for further explanation.

Fang et al. [75] proposed an NR image quality metric. The metric combines the blurriness, noisiness and blockiness of an image in a non-linear manner. The authors observe that for any image, the difference between two adjacent pixels follows a Laplace distribution. The behaviour of the distribution changes based on the type of artefact. They used these characteristics to define the three metrics and mapped them to human perception scores. A close correlation with perceptual tests is achieved along with smaller computational times than previous metrics. The solution requires the generation of a codebook in order to map the metric to human perception. This is a QoE metric that can estimate a humans' response to an image based on objective metrics.

Rodriguez et al. [76] propose a no-reference video quality metric for streaming using MPEG-DASH. The novelty of this proposal is not only that it is a no-reference model, but also the impact factors considered. The authors measure the initial delay, temporal interruptions, interruption locations, and resolution changes. More broadly, the solution considers the pauses and resolution changes, two of the major influences on user QoE. This metric has low cost and low complexity as it only considers application layer parameters. The metric achieved good performance with a high correlation coefficient.

Xue et al. [77] propose a no reference quality metric that considers the impact of frame freezing due to packet loss or delay. A neural network is trained using features chosen to capture the impact of frame freezing on perceived quality. Some of these features include number of freezes, freeze duration and the distance between freezes, and have been chosen based on the results of subjective studies. The resulting metric has been shown to have low complexity with a high accuracy based on the data sets. However, the high accuracy of the metric is achieved based on a small set of data, the videos used for training the metric and a separate 14 video test dataset.

Utility metrics are commonly used throughout adaptive solutions for assigning comparable scores to measurements. Multi-Attribute Decision Making (MADM) methods make use of these utilities to combine and aggregate the various metrics into one overall value for further decision making. Trestian et al.[78] perform a comparison of the most popular MADM methods. They compare the performance of GRA, MEW, SAW and TOPSIS for network selection in a multimedia wireless network environment. The authors mainly consider energy and user perceived qualities gathered from real testing. Results from the study show that the Multiplicative Exponentially Weighted (MEW) method achieves the fairest trade-off between the attributes.

Zhao et al. [79] perform a survey of the recent works in the area of Quality of Experience. The various influencing factors related to QoE are surveyed and categorised. A point of interest that is highlighted in this work is the need for QoE modelling to give further consideration to context and human influencing factors. It is clear that QoE is essential in improving the video transmission system and will be essential with the wide uptake of DASH. A number of open areas are highlighted including integrating these latest influence factors into real-time QoE monitoring and identifying the most important factors.

Seufert et al. [80] present a survey analysing the latest subjective studies covering the QoE aspects of adaptive solutions. HAS influence factors and QoE models are identified along with a number of open issues. Some interesting points highlighted from this work include the fact that stalling is the worst degradation in adaptation, HAS outperforms classical streaming, frequent or large quality changes are negative, multi-dimensional adaptation is a future direction (consider picture quality, rate and resolution!), and finally a holistic QoE model is drastically needed to combine all the various aspects of QoE.

Despite the significant work that has already been completed in this area, there is still a dire need for a comprehensive QoE metric. Various solutions have introduced NR models that assess certain contributing factors, such as [75] and [76]. There is still much to learn about the exact influence these various factors have on the users' perceived experience. Furthermore, few of the discussed works even consider the energy cost involved during the decision making process. This will be crucial going forward with mobile devices becoming the dominant device across networks.

3.1.2 Energy Aware Solutions

This section discusses the solutions that perform multimedia adaptation while considering the energy costs on the device. These methods include: battery monitoring and energy prediction to determine the remaining life of a device; content aware solutions that consider the energy cost of particular multimedia rates; and solutions that aim to balance energy and quality.

McMullin et al. [81] propose a power saving adaptive multimedia delivery mechanism in this work. PS-AMy aims to extend the amount of time the user can view a stream on a battery powered device. The quality of the stream is adapted based on the loss and battery life of the device. The energy of the device dictates the maximum quality rate while the loss adjusts the quality between the lowest and the highest. The solution extends the battery life of the device while keeping losses to a minimum therefore maintaining quality. Despite the lower energy costs, PS-AMy consistently delivers a lower quality which is not desirable.

Kennedy et al. [82] propose a cross-layer solution to achieve power savings while preserving user QoS levels. BaseAmy considers the remaining battery, remaining video duration and packet loss rate to decide if video adaptation is required in order to achieve power savings while maintaining user perceived QoS levels. The authors consider five quality levels based on video bitrates ranging from 400kbps up to 2 Mbps. Based on the measured values, the mechanism adapts the requested quality level down when the remaining battery life is low, and up when the battery life is high. This allows the solution to achieve energy savings when required but also provide the user the best experience when possible. Testing showed battery life was extended by up to 18% when compared to a non adaptive solution with quality increased by up to 4%. A considerable drawback to this solution is that it does not consider additional traffic or mobile environments.

Jalal et al. [83] propose a content aware solution for power saving educational multimedia adaptation on mobile devices. A minimum quality level is determined as the point at which the fragment is still suitable for learning purposes. This level is determined and used as a constraint for the adaptation. Priority is assigned to fragments based on their content. If the visual contents of a fragment are quality sensitive, then a higher priority will be given over a fragment with mainly oral content. Through this fragment by fragment adaptation, the battery efficiency is maximised in this solution. Such a solution would not be effective in a non education scenario where visual content is equal if not more important than audio.

Zou et al. proposed a series of solutions [84]–[86] based on balancing energy and performance. eDOAS [84] proposes an energy aware device oriented adaptive multimedia solution. This solution adapts the content when offloading from a cellular network to a Wi-Fi network to suit the network conditions as well as the device energy consumption. This was one of the first works to consider the characteristics of the user device during adaptation. When compared to other state of the art solutions, eDOAS achieved a performance improvement of up to 20dB and a reduction in energy consumption of 11%. Only QoS metrics are used in the adaptation solution within this work.

E2DOAS [85] extends upon eDOAS by introducing a cloud based crowdsourcing assessment tool. New users are invited to join the assessment tool and perform a brief subjective test using their device. The web-based tool then models the QoE metrics based on the results drawn from the various users. By making use of this, an adaptive multimedia mechanism on the client device adjusts the quality of the requested video considering the characteristics of the user device and the network. An optimal trade-off between QoE and energy savings is achieved. The web-based CQTS tool seems quite cumbersome, having to make the users complete registration and subjective tests in order to benefit from its advantages.

The final extension of the eDOAS solution is described in [86]. This paper introduces a coalition game-based rate allocation strategy to balance perceived quality and mobile energy savings and achieve system fairness in a heterogeneous wireless environment. With adaptive multimedia, fair network resource allocation is a major issue which needs to be addressed. The two-stage coalition oriented game increases system fairness by up to 20% when compared to other solutions. Furthermore, E3DOAS achieves an optimal balance between energy and quality and outperforms other solutions from the literature in performance.

Trestian et al. [87] perform a study on the relationship between energy and quality for multimedia over mobile devices. More specifically, they analyse the relationship between user preferences, content type, quality and energy consumption. Subjective tests are used to measure the quality of the content. From their studies, they made a number of interesting observations. User perception of quality is different based on the content type and the video quality can be dropped in certain scenarios without a negative impact of user QoE. In fact, energy savings of up to 54% can be achieved by sacrificing high quality. Only a single device was considered during this work.

Hoque et al. [88] present a survey to compare solutions that improve energy-efficiency in multimedia streaming between hand-held devices. What makes this particular survey unique is the fact that it considers the consumption of the wireless interfaces across the various layers of the internet stack. By splitting the solutions by the various layers, this survey is very beneficial. Cross layer solutions are highlighted as dominant solutions while numerous 802.11 approaches are highlighted and examined.

3.2 Multimedia Delivery

In this section, solutions for multimedia delivery are investigated. This topic focuses on the various methods proposed in the literature for delivering multimedia content to the user. Various approaches are taken in order to maximise efficiency and minimise loss and congestion. Again, this section will consider the solutions under two main topics of awareness: performance and energy. As the amount of data being transferred across the internet continues to rapidly grow, putting a substantial strain on the existing infrastructures, new approaches need to be considered. One direction that is regularly returned to is cooperative distribution.

Such approaches have existed for years including torrenting clients such as BitTorrent, uTorrent. However, the concept has been repeatedly dismissed due to the complexity and individual resource costs involved as well as few advancements being made (due to lack of attention). This section provides an extensive review of the recent work with regards to multimedia delivery.

3.2.1 Performance Aware Solutions

Performance aware solutions take into consideration the performance of the involved devices and/or the network as a whole. They consider the bandwidth of the network and fair sharing of the resources; the location of multimedia content and which hosts are most suited to request from; and neighbour load/capacity is considered as part of their mechanisms in order to reduce overburdening individual devices and fairly sharing the workload.

Rovcanin et al. [89] proposed a DASH-Based distribution system which considers the network conditions and the device characteristics. This solution dynamically updates the content MPD file based on which devices are currently holding the segments. Powerful peers such as laptops are prioritised as local hosts so that weaker peers are not overburdened. While performance is improved, this work gives no consideration to the energy cost of the quality levels and only considers a LAN environment. Also, the solution relies on the presence of a central gateway device for host selection.

Lederer et al. [90] propose a peer-assisted DASH solution. This was one of the earliest works to consider cooperative segment distribution using DASH. The authors compare the solution against conventional DASH approaches. A modified Media Presentation Description (MPD) is introduced in this work to provide additional Uniform Resource Locator (URL)s for segment downloads. The customised version still complies with the DASH standard. The client can choose to download from the infrastructure URL or from a peer URL. When compared to conventional client/server solutions, pDASH manages to reduce the traffic to the server by up to 25%. This work provides early proof that Peer-to-Peer (P2P) is a viable option for managing the large data explosion being experienced. However, this work only considers generic clients device awareness including energy consumption. Furthermore, no load balancing is considered as a part of this work, meaning certain clients potentially do much more than others.

Tian et al. [91] propose an incentive based P2P dynamic streaming solution. In P2P dynamic streaming, swarms are generated for each representation of the video which can then internally share the content. This solution designs an incentive mechanism using cooperative game theory to encourage peer devices to contribute to segment distribution. Peers are assigned to swarms and, based on their upload contribution and willingness to pay, can request segments of the desired quality from other peers within the swarm. From simulations, this solution is shown to be efficient. Using such an incentive solution, rate improvements of up to 30% can be seen. This work does not consider MPEG-DASH or device characteristics such as energy consumption.

Natali et al. [92] propose a P2P-DASH architecture which leverages the advantages provided by each technology. Several overlays are used, one per DASH representation, in order to distribute the segments cooperatively among client peers. A novel rate algorithm is proposed to steer peers between layers based on the performance of the overall architecture and the peer. This mechanism is decentralised and each peer is responsible for its own adaptation. Through simulations, the solution is shown to be effective with users guaranteed a very good viewing experience. Furthermore, a reference model is compared against via an Integer Linear Programming model. A near optimum number of peers are shown to receive their desired rate in good conditions and only suboptimal in the worse scenarios.

An extension of their previous P2P-DASH control strategy work, **Natali et al. [93]** explore the system behaviour in order to improve performance. A number of features are introduced to improve the switching delay experienced when going from one representation to another. A silent neighbour mechanism is proposed which allows peers to be aware of a portion of peers in adjacent overlays. Buffer reuse allows fully received segments to be reused after transitioning, while priority queuing helps new peers. Finally, weaker peers are allocated to lower rates. All of these additions together are shown to achieve excellent performance with reduced delays and a high delivery ratio. The authors do not consider mobility or energy in this proposed work.

3.2.2 Energy Aware Solutions

The solutions in this section focus on achieving energy efficient delivery of multimedia content. Approaches include clustering and multicasting to reduce repeated transmissions of similar or identical content; energy efficient multipath transmission to collect content from multiple sources simultaneously; heuristic solutions based on consumption modelling to match objective solutions to real world values; and load balancing within a network for energy efficient multimedia delivery.

Zhang et al. [94] propose an energy efficient scheme for multimedia transmission over heterogeneous cellular networks. From previous works, it was observed that some stream requests by users in similar geographical locations are similar for a period of time. The authors exploit this user behaviour in order to optimize energy efficiency for multimedia stream delivery. Broadcasting/multicasting is employed to serve users with requests that are similar or identical, therefore reducing the amount of repeat transmissions. If a user requests content late, they will be added to an existing multicast and any content they have missed is unicast.

Base station cooperation enables this work to optimise the energy efficiency while maintaining user QoS. This work does not consider user QoE or other network interfaces.

Zeid et al. [95] investigate the potential of location awareness in green media delivery. By considering the location and mobility of a user, long term planning can be achieved. It has been shown that users regularly follow similar paths with a predictability of up to 88%. As a result, the authors demonstrate the improvement that can be achieved for both QoE and energy savings. Three main issues are highlighted as a part of this study. User behaviour prediction is one of the main points of focus for not just location, but also content prediction. Uncertainty is a major concern for prediction as it will never be possible to be fully certain of a users behaviour. Therefore, there needs to be an effective means of handling this (e.g. if a user unexpectedly changes direction). Finally, the implementation of location aware solutions is an open area with NFV and SDN being considered as potential approaches for dynamic network configuration.

Wu et al. [96] proposes an energy-video aware multipath transport protocol. A mathematical model of the frame level tradeoff between energy and quality is developed for multipath wireless video transport. The authors propose a scheduling algorithm for prioritised frame scheduling and path selection to achieve the target quality with minimal energy cost. A loss protection scheme is also introduced to minimise the total distortion on the video. The overall protocol EVIS, achieves excellent performance improvements.

Hoque et al. [97] propose an energy efficient multimedia delivery scheme in this work. The authors model the energy consumption based on bursty streaming traffic over TCP. The energy efficient delivery scheme EStreamer is proposed based on this modelling. The traffic is shaped based on a heuristic defined using the energy consumption models in order to maximise energy efficiency while maintaining quality. The solution is agnostic to the wireless network interface and is shown to perform well over WiFi, 4G and 3G. Four different smartphones are used to test the solution with varying results. While savings are achieved overall, the hardware differences clearly have an impact on the performance. Furthermore, signalling overhead is investigated for various configurations with best practices suggested.

Trestian et al. [98] propose an energy aware network selection algorithm. E-PoFaNS considers energy, quality, cost and mobility in order to select a network. These metrics are weighed against each other using a Multiplicative Exponentially Weighted (MEW) metric with four separate weight values which are related to the user preferences. These allow the users to choose the metrics they are most interested in (e.g. one user may be more interested in higher quality and is willing to sacrifice battery life). The solution determines the best target network based on this performance and will trigger the handover process when necessary. When compared to other recent solutions, their results show energy savings of up to 30% while maintaining the perceived quality.

Singhal et al. [99] propose a cross-layer optimisation solution for improving the quality and energy of multimedia for wireless broadcast receivers. The users are grouped by considering the device capabilities and network conditions. The adaptive content is then broadcast to the groups using SVC. The encoding parameters for the multimedia content are optimised using cooperative game theory. Layer aware time slicing enables a optimum trade off between QoE and energy savings to be achieved based on the users preferences. Improved performance is shown compared to existing works with energy savings of up to 17% achieved with 3.5% quality improvements.

Anedda et al. [100] proposes E-ARMANS, a load balancing energy aware video delivery solution. The authors aim to achieve a balance between energy savings and network traffic load while maintaining high video quality. The solution monitors the network conditions as well as the device characteristics of the user in order to provide the best connectivity for the user. Through simulation based testing, E-ARMANS achieves an increase of up to 47% in battery life delivering to a 1080p screen. Only packet loss is considered within this work which is simple and does not give an accurate representation of user QoE levels.

Al-Kanj et. al. [101] present a survey that studies the main energy aware cooperative solutions in the literature. They highlight the potential advantages of short range over long range communications. In a geographically close scenario, short range communications can achieve higher data rates with low energy costs than long range.

3.3 Chapter Summary

This chapter presents the related work from the area of multimedia adaptation solutions and content delivery schemes. The emphasis of the discussion focuses on the performance and energy aware solutions within these topics. Much work has been carried out to solve the various highlighted issues including the traffic explosion and mobile energy limitations. However, there is no clear framework or solution that seamlessly brings together energy awareness, quality balancing, load balancing, high QoE and mobile delivery. Many research issues remain open and are investigated within this thesis.

Chapter 4

Proposed Architecture and Algorithms

This chapter presents the system architecture and detailed descriptions of the two main solutions proposed in this thesis: 1) A distributed Energy-Aware Cooperative delivery solution (ENCO) establishes the distributed architecture and introduces energy-aware adaptive multimedia streaming to a cooperative environment; 2) A Mobile Quality-oriented Cooperative Multimedia Delivery Solution (MENCO) brings mobility awareness to the distributed cooperative environment which expands the scope of the architecture and allows more intelligent host selection for the devices.

4.1 Mobile Multimedia Delivery: Current Experience

The popularity of video is growing rapidly with the use of embedded videos in social media and popular streaming platforms such as YouTube, Twitch and Netflix. Furthermore, mobile video traffic is expected to grow 9-fold by 2021 accounting for 79% of global mobile data traffic [1]. This massive growth will put a significant strain on the existing infrastructure with new solutions needed to meet the demand. Indeed the current growth of traffic is causing a bandwidth bottleneck on the Internet [102]. Further contributing to this problem, and bringing problems of its own, is the growth in mobile device capability and popularity. Companies are rapidly releasing new more powerful devices each year. The user is therefore expecting more from their devices, especially with services improving just as quickly. High definition video has become the normal standard for many users with new technologies rapidly becoming more accessible including 4k, VR/AR and eventually 8k video. These new technologies not only increase the resource demand on the user device, but also cause a rapid growth in network traffic.

Two main issues arise out of the above, 1) current and future network infrastructures struggle to keep up with the growing traffic demand; 2) the user wants the best experience at all times while also expecting their device to survive a long time without the need for charging. The approach considered in this thesis to help alleviate the traffic problem is the introduction of a cooperative architecture which is self aware and makes use of existing resources. Such an approach takes repeated requests away from the network and takes advantage of the abundance of devices available. Moreover, each device within the architecture is capable of adapting to the network conditions in order to maintain a high quality of experience for the user. By utilising adaptive multimedia methods, a balance can be achieved between quality and energy. While this is not new, introducing it into such a cooperative distributed environment poses its own set of challenges which have yet to be solved.

This thesis proposes a distributed cooperative solution that will enable mobile users to actively cooperate with each other to retrieve and share multimedia content within a heterogeneous mobile wireless environment. By making use of the local device resources, network load is reduced due to less requests being sent to the remote servers. Furthermore, such a distributed solution allows for the range of the network to be extended without the need for additional infrastructure. Adaptive multimedia is introduced to the cooperative paradigm to allow each user to adapt their desired content based on their current performance and the state of the network. This adaptation allows for a balance between quality and energy to be achieved if the user desires.

Two major contributions are proposed in this thesis in order to achieve the goals described: 1) A distributed Energy-Aware Cooperative delivery solution (ENCO) introduces adaptive streaming into a distributed architecture where content is cooperatively distributed among users in a fair manner; 2) A Mobile Quality-oriented Cooperative Multimedia Delivery Solution (MENCO) brings mobility awareness into the architecture which expands the scope of the solution by providing devices the ability to select the best hosts to communicate with based on their direction and speed of travel.

In this chapter, the architecture of the overall solution is described. Detailed descriptions of each of the contributions are presented including the components and algorithms.

4.2 Distributed Energy-Aware Cooperative delivery (ENCO)

In this proposed solution, the user devices are responsible for sharing content among themselves as requested. Each of the user devices maintains a record of the neighbours available as well as its own performance. When a client wishes to request content, they must first consider their own performance and adapt the requested content to suit. This adaptation allows the device to achieve a balance between quality and energy, but can also be adjusted to suit the users needs. Next the client must select a neighbour to request content from using the stored record. This is based on prior knowledge of content availability, device performance, load and distance. Finally, when the content has been received, the involved devices will share their updated personal record with each other. This method of shar-

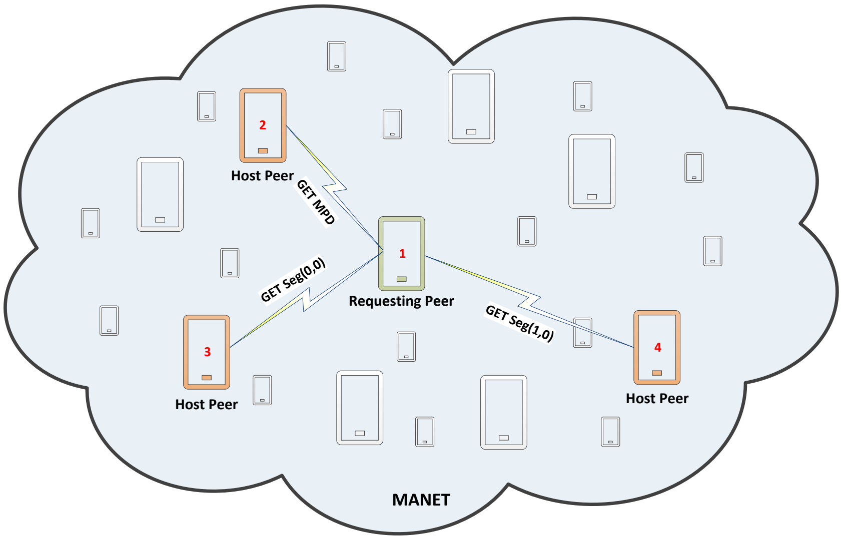

Figure 4.1: Overview of ENCO Solution

ing control information is low impact and allows regular neighbours to stay up to date with each other. A detailed description of the architecture components and algorithms follows.

4.2.1 ENCO Architecture

An overview of the introduced architecture is illustrated in Figure 4.2. Each peer consists of four separate components. The *Storage module* is responsible for storing segments received from peer hosts for future distribution. All of the decisions are made in the *Decision Module* including quality adaptation and host selection. The *Monitor* regularly measures the energy consumption, quality, location and load of the peer device. Finally, the *Communications Manager* is responsible for identifying and interacting with neighbouring peers. In addition to these modules, a modified MPD file is also introduced to enable distributed peer-aware DASH delivery. This file is called the area MPD (aMPD).

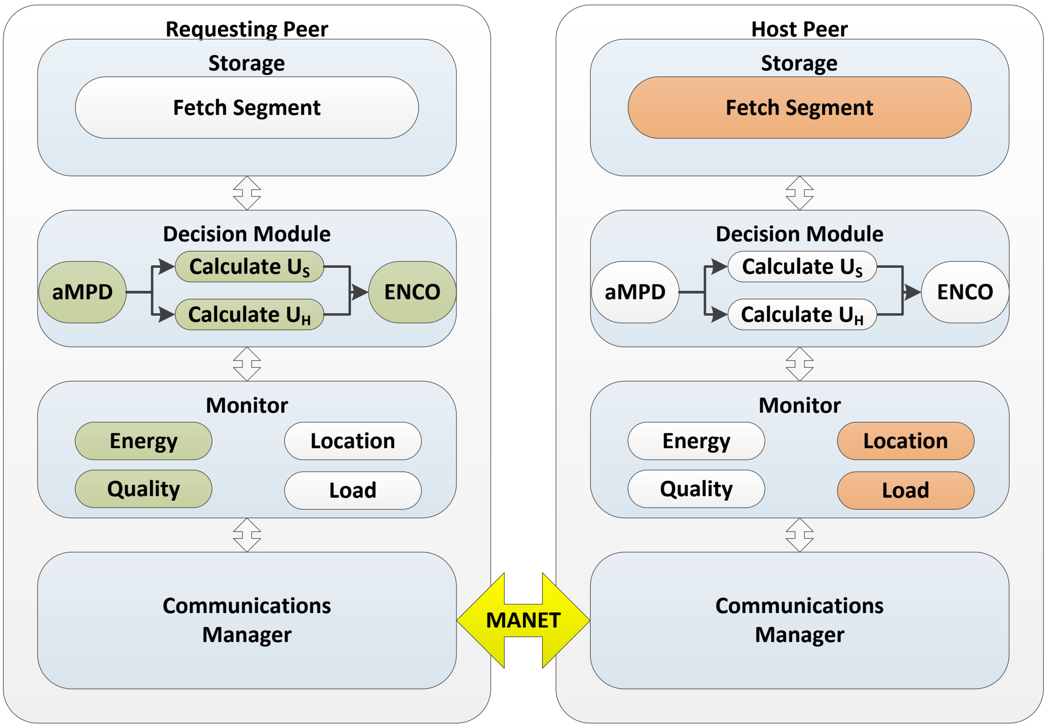

Figure 4.2: Illustration of ENCO Architecture

4.2.1.1 Area Media Presentation Description File

The aMPD file is an enhanced MPD file which not only details the locations of the segments for a particular piece of content, but also contains the previously reported geographical location and load of the hosting peers. This addition gives the file its name as the locations detailed in each aMPD file will depend on the location of the requesting peer effectively forming an area around the user. As a result, each user will have a unique aMPD file. The role of the aMPD can be visualised in Figure 4.1. In this example, the central peer is a user who is interested in accessing the multimedia content. The surrounding devices are the neighbouring peers visible to the user using short range communication (e.g. WiFi Direct [103]). To access the content, the user declares its interest to neighbouring devices. In response, the peers share their aMPD files with the requesting peer who can then generate their own aMPD from the information. Upon creating the aMPD file, the user device proceeds to request the first segments of the multimedia content at the lowest quality in order for the user to get and play the content with the minimum start-up delay. Next, the ENergy aware COoperative multimedia delivery (ENCO) algorithm is employed to determine the most appropriate quality level, as well as the most suitable host to request the following segments from.

4.2.1.2 Storage

When a peer receives a segment from a neighbour, it will be stored locally. This thesis considers the segments have enough storage capacity and the management mechanism is not the focus of this research. However, a solution to this storage could be to store the most popular segments, or the latest segments only. Regardless, storage on devices is a minor issue as the latest technologies regularly reaching 128GB and above.

After storing a segment, the local aMPD file is updated to reflect the additional segment location. Other peer devices are made aware through the aMPD of the new segment being available at an additional location and the current condition of that location. Updates can be sent to neighbouring peers at regular intervals. However, in order to reduce the impact on traffic overhead, the peer will distribute aMPD updates only at the end of a segment transmission. This method should minimise the overhead by only extending an existing transmission instead of generating additional, unexpected bursts of traffic. Although, this solution may not be sufficient for keeping neighbour devices up to date as the local device may not regularly transmit segments.

When a segment is requested from a peer, the Communications Manager will request the segment from the storage via a fetch segment method. This method will be device dependent and will simply pass a copy of the segment to the Communications Manager for transmission. The storage module is a crucial component in a distributed cooperative solution as certain unique issues, such as management and update methods, need to be addressed.

4.2.1.3 Decision Module

This module is responsible for deciding what quality to request for future segments and what host to request from. These decisions are each based on two measurements provided by the Monitor Module. The quality adaptation considers a utility consisting of throughput and energy consumption from previous segment playback. Secondly, the host selection considers a separate utility including the location and experienced load of each of the hosts to decide the most suited host to request the segment from.

These utilities are compared in the ENCO algorithm situated inside this module. The decision module is effectively a distributed controller for this architecture as it is responsible for making all request decisions for the participating peer devices. By further developing this module, it is possible to further enhance the architecture with future contributions such as best network interface selection and mobility prediction.

4.2.1.4 Monitor Module

This module is responsible for measuring the various attributes of interest as the mobile device sends, receives and plays multimedia content. Currently, the metrics being measured are the throughput, energy consumption, device location and amount of data transmitted.

The throughput experienced by the device is measured at a regular interval in order to determine the capacity of the network and the most suitable quality to request for future segments. In order to translate the throughput into a quality metric, an equation defined in [98] is used to map the throughput value to a normalised 0-1 measurement.

Alongside the throughput, the energy consumption of multimedia retrieval and playback is predicted by the equation developed in [78], [98]. This function calculates the expected energy consumption based on the expected throughput of the requested quality level. By comparing against the expected energy cost of the lowest and highest rates, the utility is normalised.

Each device is aware of their current position. As part of the aMPD update mechanism, this position is regularly shared with neighbouring peers to be used in their decision making. The monitor considers the shared location of each of the peers in the aMPD, and the current location of the local peer, in order to calculate a location metric. The value gained from this metric is a type of reputation metric that determines how suitable a device is to be a host based on the distance between the two devices. Further away devices will have a near 0 value as they are less likely to be able to deliver the desired content at a high quality, while much closer devices will achieve closer to 1.

Finally, the load of each device is monitored. This is achieved by measuring exactly how much content is being sent. The load here is directly related to the amount of requests being received and answered. Therefore, the more segments a host sends, the higher their load. However, this metric is averaged over time, so the longer the host sends no or few segments, the lower their respective load goes.

These four metrics are used to calculate the two separate utilities. A future work would be to consider the combination of all of the metrics into a single utility. However, as some of these metrics are provided by neighbouring peers and are used for two different decisions, this will be complicated. Each of these metrics is measured periodically in order to have a steady understanding of the performance of the peer devices and the network overall. All of the discussed equations are introduced in details later in this section.

4.2.1.5 Communications Manager

This module is responsible for all communications to and from a device. When the decision module decides on the next segment and host, the request is generated and sent through the Communications Manager module. When the request is answered, the response is handled by this module. Any incoming segments are stored in the storage module. If a request is received from another peer, the storage module will be requested to provide a copy of the segment to start transmitting in response. Finally, upon finishing the transmission of a segment, an updated aMPD is also sent to the peer to keep it up to date on the local peers current condition.

The Communications Manager is agnostic and does not depend on a particular network interface to function. Therefore, this module will be ideal for future works involving network selection across multiple interfaces. Furthermore, few changes should be necessary in order to implement this solution to other interfaces in its current scenario.

4.2.2 ENCO Algorithms

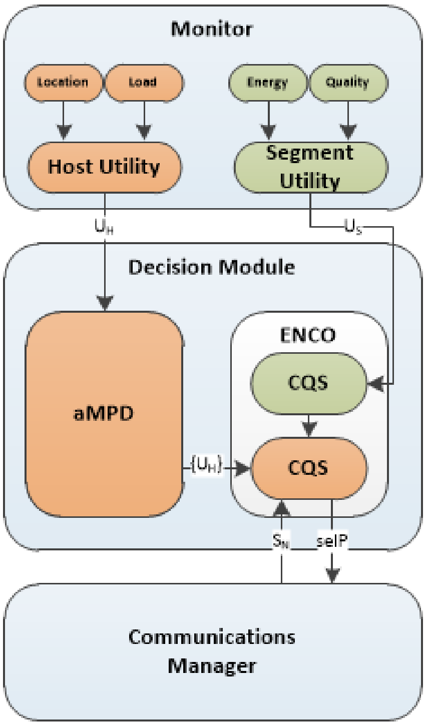

Figure 4.3: Illustration of ENCO Module

The proposed solution aims to improve the energy consumption and quality for the user while fairly distributing the workload across the hosting peers. As can be seen in Figure 4.2, each peer contains an ENCO module allowing each peer to be self adaptive. To avoid introducing significant overheads into the network, each host will only send an update of its current metrics to the requesting peer at the end of its current segment transmission. ENCO consists of two separate stages, as shown in Figure 4.3, quality adaptation and host selection. These are presented as two individual algorithms, *Content Quality Selection (CQS)* and *Host Peer Selection (HPS)*.

4.2.2.1 Content Quality Selection (CQS)

The CQS algorithm is used to determine the next segment quality to request. It weighs the trade-off between the predicted perceived quality and energy consumption of the previous segment to make the decision. It includes two utilities which assess the perceived quality and energy consumption [98] and employs a Multiplicative Exponentially Weighted (MEW) utility method to weigh them for the final decision algorithm.

- **Quality Utility**

  In order to measure the quality of the received content, the quality utility function defined in equation (4.1), [98] is employed. This utility maps the measured throughput to user satisfaction, represented by a [0,1] interval, using a zone-based sigmoid function.

$$u_Q = \begin{cases} 0 & for\ Th < Th_{min} \\ 1 - e^{\frac{-\alpha Th^2}{\beta + Th}} & for\ Th_{min} < Th \leq Th_{max} \\ 1 & otherwise \end{cases} \tag{4.1}$$

  In equation (4.1), α and β are two parameters which define the shape of the function, Th is the measured throughput, Th_{min} is the minimum threshold for the throughput below which the quality is unacceptable and Th_{max} is the maximum threshold for throughput, above which there will be no noticeable improvements to the user.

- **Energy Utility**

  The predicted energy consumption of multimedia playback is mapped to a utility value in the [0,1] interval using the energy utility function from equation (4.2) [78], [98].

$$u_E = \begin{cases} 1 & for\ E \leq E_{min} \\ \frac{E_{max}-E}{E_{max}-E_{min}} & for\ E_{min} < E \leq E_{max} \\ 0 & otherwise \end{cases} \tag{4.2}$$

In equation (4.2), E_{min} is the energy consumed for the lowest multimedia quality throughput (Joule), while E_{max} is the energy associated with the highest quality throughput (Joule). E is the predicted energy consumption for the device (Joule) which receives data based on the model described by Mahmud et. al. [104]. E_{min} and E_{max} are calculated using Th_{min} and Th_{max} respectively using equation (4.3).

$$E = t_{seg}\left[r_t + Th_{req}r_d\right] + c \tag{4.3}$$

In equation (4.3), Th_{req} is the required throughput for the current quality level (Mbps), t_{seg} is the duration of a segment (seconds), r_t and r_d are energy consumption per second (W) and energy consumption per data received (Joule/KB), respectively. These rates can be found either from device specifications or from practical measurements. A literature work with this energy model [78], measured typical values for r_d and r_t for a modern mobile device while receiving multimedia content at different quality levels.

- **Segment Utility**

  To weigh the multiple metrics, a Multi Attribute Decision Making (MADM) method and a multiplicative exponentially weighted (MEW) utility is chosen. Authors of [78] show that employing a MEW method outperforms other state-of-the-art MADM methods when performing a trade-off between multiple attributes such as energy and quality.

$$U_S = u_Q^{w_Q} . u_E^{w_E} \tag{4.4}$$

In equation (4.4), u_Q and u_E are quality and energy utilities, respectively. w_Q and w_E are the weights assigned to the quality and energy utilities, respectively with $w_Q + w_E = 1$. U_S is the segment utility metric which is later used to adapt the requested quality. The weights can be adjusted directly by the user in order to prioritise one metric over the other, or set to suit a users profile.

- **CQS Algorithm**

The first discussed algorithm, CQS, is responsible for deciding what quality to request next based on the aforementioned segment utility. A utility threshold (U_{Thresh}) is determined using the previously requested quality level and corresponding predicted energy using equation (4.4). A *Count* value is incremented each time the utility exceeds the threshold and is set to 0 otherwise. Another value called Hold Off (*HO*) is an exponential variable which helps reduce quality variations when they become too rapid and avoid the ping-pong effect. If the utility is under the threshold, *HO* is updated by doubling its previous value. Otherwise, *HO* is halved if *Count* exceeds it by a predefined factor which can be set based on experimentation. In this study, it was observed that without a factor, when *Count* exceeds the *HO* value it was immediately halved. This often resulted in the algorithm quickly having to decrease the quality because the *Count* was not stable. By setting a "buffering" factor, the *Count* had to stabilise before *HO* could be reduced. Through testing, a factor of 1.5 was determined as an acceptable value. An optimal factor may be determined but will be reliant on the preference of the user.

Algorithm 4.1 Content Quality Selection (CQS) Algorithm

Input: Segment Utility, U_S; Utility Threshold, U_{Thresh}

Output: Desired Quality, QL

```
Update U_Thresh
if U_S ≤ U_Thresh then
    if QL > QL_min then
        Decrease QL one level
    end if
    Count = 0
    Update HO
else
    Count++
    if Count ≥ HO then
        Update HO
        Increase QL one level
    end if
end if
return QL
```

In order to select the most suitable quality level for the next segment, CQS compares the segment utility (U_S) against the utility threshold. If the utility falls below the threshold, the quality level is decremented to a minimum acceptable level of QL_{min}. Otherwise, the current *Count* is incremented and compared against the *HO*. The quality level will only be incremented when *Count* exceeds the *HO* value. The resulting algorithm is shown in Algorithm 4.1.

4.2.2.2 Host Peer Selection (HPS)

Due to the decentralised nature of this solution, a separate decision is necessary not only to determine the most suitable host peer to request a segment from, but also to fairly share requests across all available peers. The second algorithm of this solution is the HPS algorithm which considers the load experienced by each potential host peer as well as the distance from the user peer.

- **Load Utility**

  To avoid repeated requests being made to the same peers, each device monitors their load, which is considered as the average amount of data transmitted from the device, expressed in Mbps. As such, if a host is sending successive packets over a short period of time, their load will increase. However, the longer a host does not send packets, the lower its load will become. In mapping the load to a [0,1] interval, a high load is associated with a utility value of 0 and a low load with 1. The rest of the range is populated using the normalised load. The load utility, u_L, can be seen in equation (4.5).

$$u_L = \begin{cases} 1 & for\ Tx_{Av} \leq Tx_{min} \\ \frac{Tx_{max} - Tx_{Av}}{Tx_{max} - Tx_{min}} & for\ Tx_{min} < Tx_{Av} \leq Tx_{max} \\ 0 & otherwise \end{cases} \tag{4.5}$$

In equation (4.5), Tx_{Av} is the average transmitted load (Mbps), Tx_{min} is the minimum load (Mbps) and Tx_{max} is the highest load (Mbps). Th_{min} and Th_{max} are used to calculate Tx_{min} and Tx_{max} respectively. These are the required rates to achieve the lowest quality and the highest quality level, respectively.

- **Distance Utility**

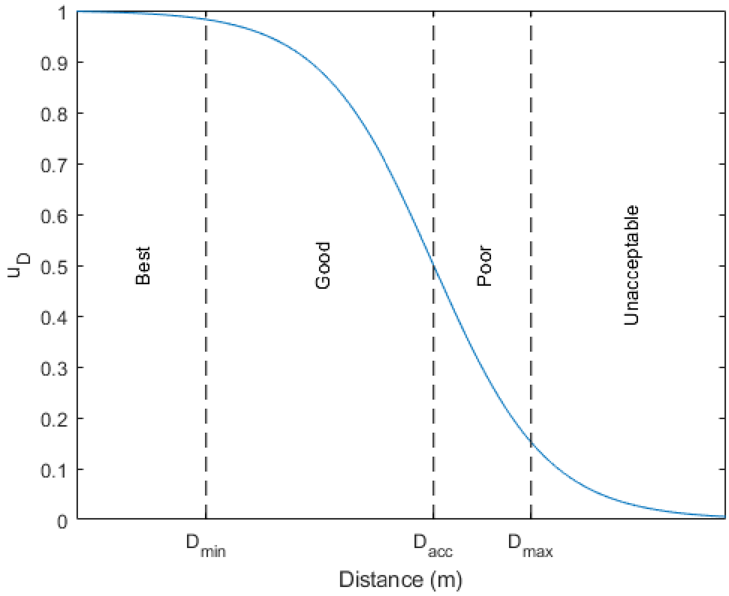

Figure 4.4: Distance Utility Sigmoid Function

The distance utility was defined considering the relationship between WiFi rate adaptation and distance. A modified sigmoid function is used to produce a utility in the [0,1] interval, as shown in Figure 4.4. Three parameters were required to compute the utility function: the minimum distance below which no change is noticed (D_{min}), the acceptable distance up to which the rate is high and above which the rate begins to deteriorate (D_{acc}), and the maximum distance above which the distance is unacceptable (D_{max}). The equation for this function is given in equation (4.6).

$$u_D = \begin{cases} 1 & for\ D \leq D_{min} \\ 1 - \frac{1}{1+e^{\frac{\gamma - D}{\delta}}} & for\ D_{min} < D \leq D_{max} \\ 0 & otherwise \end{cases} \tag{4.6}$$

In equation (4.6), D is the estimated distance between devices (m), D_{min} is the distance where the rate starts to diminish (m), and D_{max} is the distance at which the rate is too poor to maintain suitable throughput (m). Two parameters are used to shape the utility function, γ and δ. To calculate these, two relationships were used to find equations to solve for the required values.

First, when $D = D_{min}$, then $u_D = u_{max}$. From this, the first equation can be found so that γ is expressed in terms of δ as in equation (4.7).

$$\gamma = \delta ln \left(\frac{u_{max}}{1 - u_{max}} \right) + D_{min} \tag{4.7}$$

Next, the acceptable distance (D_{acc}) is the point on the curve that the utility's shape changes from concave to convex. Mathematically, the second derivative of the function is equal to zero at D_{acc}. The second equation presented in equation (4.8) is calculated by substituting equation (4.7) into the second derivative = 0 equation. This gives us the value for δ when solved.

$$\frac{e^{\frac{\mathrm{D_{min}} - D_{acc} + \delta \log\left(-\frac{1}{\mathrm{u_{max}} - 1} - 1\right)}{\delta}}}{\delta^2 \left(e^{\frac{\mathrm{D_{min}} - D_{acc} + \delta \log\left(-\frac{1}{\mathrm{u_{max}} - 1} - 1\right)}{\delta}} + 1\right)^2} - \frac{2\, e^{\frac{2\left(\mathrm{D_{min}} - D_{acc} + \delta \log\left(-\frac{1}{\mathrm{u_{max}} - 1} - 1\right)\right)}{\delta}}}{\delta^2 \left(e^{\frac{\mathrm{D_{min}} - D_{acc} + \delta \log\left(-\frac{1}{\mathrm{u_{max}} - 1} - 1\right)}{\delta}} + 1\right)^3} = 0 \tag{4.8}$$

- **Host Utility** Another MEW utility is used to balance two metrics of interest to get the host utility, U_H as shown in equation (4.10).

$$U_H = u_L^{w_L}.u_D^{w_D} \tag{4.9}$$

The load and distance utilities, u_L and u_D, respectively, are assigned weights, w_L and w_D, where $w_L + w_D = 1$. These weights are adjusted to give more priority to load or distance depending on the preferences of the host peer.

- **HPS Algorithm**

The HPS algorithm, presented in Algorithm 4.2, uses the previously discussed host utility to determine the most suitable host peer to request the next segment from. It traverses the aMPD file to find the peer with the highest host utility that contains the segment at the desired quality level. If the desired quality is not available among the neighbouring peers, the algorithm will drop the quality and check again until either QL reaches QL_{min} or the segment is found. If no peer is found, the Communications Manager will send a request to a server if possible. Otherwise, a request will be sent to the chosen peer for the desired segment.

Algorithm 4.2 Host Peer Selection (HPS) Algorithm

Input: Desired Segment, S; Desired Quality, QL_{des}; aMPD File;

Output: Selected Host, $selP$;

$selP = NULL$
$QL = QL_{des}$
while $QL >= QL_{min}$ **do**
    for all Peers $crtP$ in aMPD **do**
        if $crtP$ stores S at QL **then**
            if $crtP.U_H > selP.U_H$ **then**
                $selP = crtP$
            end if
        end if
    end for
    if $selP! = NULL$ **then**
        break
    else
        Decrease QL one level
    end if
end while
if $selP == NULL$ **then**
    $RequestFromServer(S, QL_{des})$
else
    return $selP$
end if

4.3 Mobile Quality-oriented Cooperative Multimedia Delivery (MENCO)

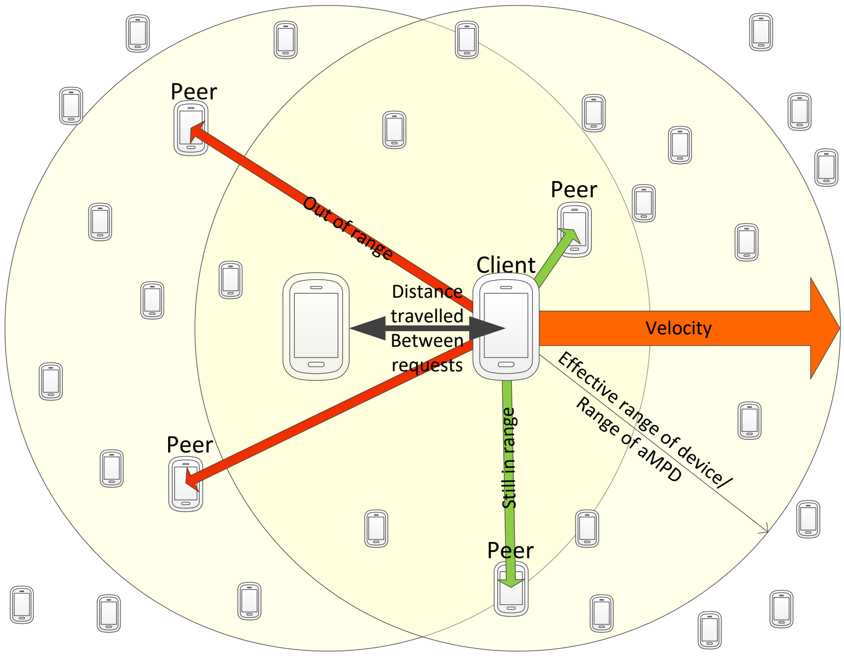

Figure 4.5: Example of architecture scenario

In the previous section, ENCO is introduced as a distributed cooperative energy aware solution. However, the performance of ENCO depends on stationary devices and does not consider the impact of dynamic mobility which is always present among mobile devices. A Mobile Quality-oriented Cooperative Multimedia Delivery Solution (MENCO) brings an innovative mobility awareness to the distributed cooperative environment. The proposed solution as illustrated in Fig. 4.5, observes the velocity of the user device and uses this information as part of the neighbour selection process. By considering this mobility before requesting content during a multimedia streaming session, MENCO aims to reduce delayed and failed requests being made to unsuitable neighbours. Without such a mechanism, neighbours can be selected which may be out of range before the transmission can finish causing significant delays and loss in quality. This section introduces the general peer architecture and the file structure introduced for maintaining relationships between client devices.

4.3.1 MENCO Architecture

Additional components must be considered in order to monitor the velocity of the device. Moreover MENCO introduces additional mechanisms which are reflected in the updated architecture. The MENCO architecture and components will be discussed here in detail.

4.3.1.1 Peer Overview

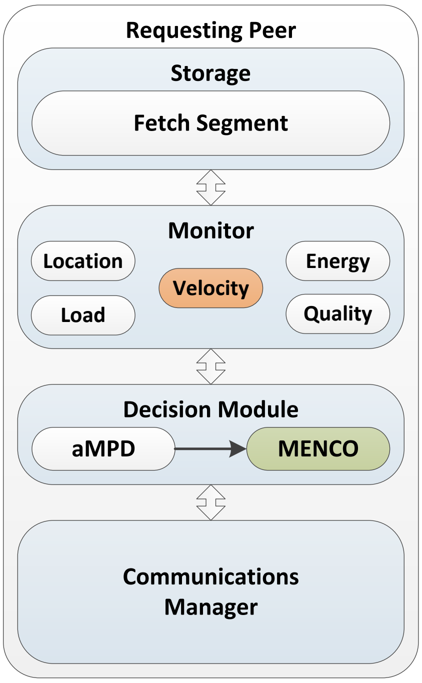

Figure 4.6: Representation of Peer Structure

Fig. 4.6 provides an illustration of the peer architecture. Each peer has an identical structure with the *Communications Manager* responsible for handling requests from other clients. Four main components make up the structure of a peer: *Storage*, *Monitor*, *Decision Module* and *Communications Manager*. Segments that are contained on a device are stored within the *Storage* module and are fetched when necessary via a fetch method unique to the device. The *Monitor* unit observes the performance of the device by monitoring a number of individual metrics including throughput, energy consumption, location, load and velocity. These metrics are used by the algorithms to determine the best quality and host for future segment requests.

All decisions are made within the *Decision Module* which contains the various algorithms and the aMPD file. The *Decision* module is also responsible for updating the aMPD. Finally, a *Communications Manager* enables the peer device to identify and communicate with neighbouring peers regardless of the network interface used. Interface selection is not considered within this work and is considered to be IEEE 802.11, a wireless LAN standard widely used in practice. Further details of the original components can be found in the previous section.

4.3.2 MENCO Algorithm

The algorithm proposed in this section, MENCO, is designed for host peer selection in a mobile cooperative distributed wireless environment. The MENCO algorithm runs once every segment duration to identify the desired segment quality and host of the next request. The quality adaptation mechanism CQS is reused from ENCO and can be seen in Fig. 4.7 within the MENCO module. This introduces the Mobile Host Peer Selection (MHPS) module to the architecture in Fig. 4.7 which is responsible for observing the velocity of the client and determining the most suitable host from the aMPD. A mobility metric is introduced which considers the angular distance between a host and the client's path of travel, and also the speed of the client. The distance utility is adjusted to accommodate for the uncertainty introduced by client mobility during a communication session. Finally,

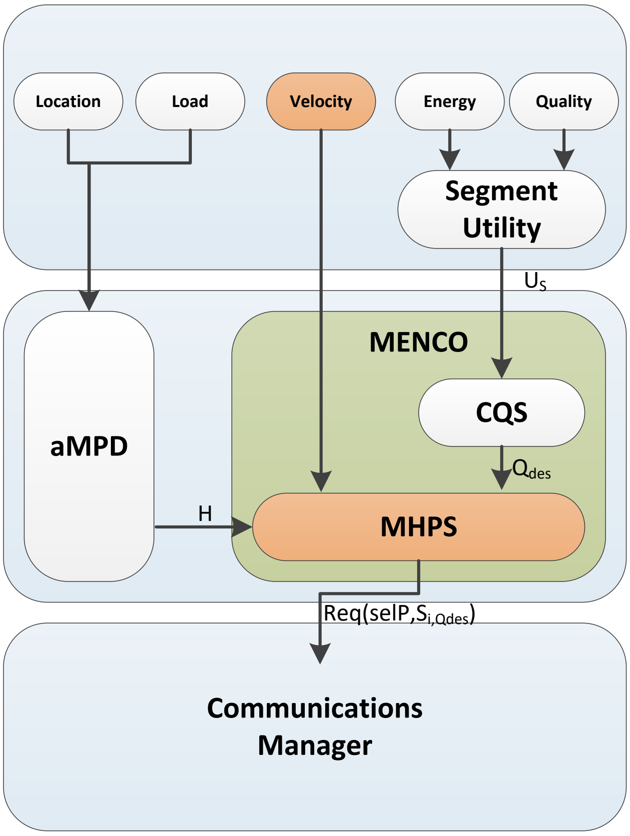

Figure 4.7: MENCO Peer Architecture

the host selection algorithm is extended to allow the client to wait or perform fast requests when necessary.

4.3.2.1 Host Utility

A host rating utility U_H was introduced in section 4.2.2.2 which gives a rating to a host based on the distance the device is located from the client, and the average load observed on the host. This metric was chosen as it was shown to reflect the load of the peer devices well in ENCO. Eq. (4.10) shows the original utility function where u_L is the load utility described in Section 4.2.2.2, u_D is the distance utility which has been further modified for this work, w_L and w_D are the load and distance weights respectively which sum to 1. A brief description of each of these utilities follows.

- **Distance Utility**

$$U_H = u_L^{w_L}.u_D^{w_D} \tag{4.10}$$

The distance utility function is shaped based on the rate-distance adaptation profile of WiFi. The general equation is represented in eq. (4.11).

$$u_D = \begin{cases} 1 & for\ D \leq D_{min} \\ 1 - \frac{1}{1+e^{\frac{\gamma - D}{\delta}}} & for\ D_{min} < D \leq D_{max} \\ 0 & otherwise \end{cases} \tag{4.11}$$

Here, D is the distance from the client to host, D_{min} is the distance up to which rate adaptation does not occur, D_{max} is the maximum distance above which the rates are no longer viable. γ and δ are two constants used to shape the curve based on the modelled profile. Original values considered only a static environment allowing the curve to give ratings to further distances. Based on the observed performance of TCP over varying distances, distant nodes are not as reliable in mobile scenarios as they are in a static scenario. If a device goes out of range, there is a high potential for large timeouts, drastically impacting the quality for the user. To compensate for this, the sigmoid was adjusted to reduce the reputation given to distant nodes. The shaping parameters, γ and δ, were recomputed to consider a smaller range. By considering the maximum distance to be $D_{max} = 35m$ and minimum distance $D_{min} = 5m$, updated values of $\gamma = 25$ and $\delta = 4.35$ were found.

- **Load Utility**

  The load utility, u_L, metric reflects the amount of data a host has shared with neighbouring devices, indicating their workload. It is used to encourage fairness in the host selection mechanism and is calculated using eq. (4.12).

$$u_L = \begin{cases} 1, & for\ Tx_{Av} \leq Tx_{min} \\ \frac{Tx_{max} - TxAv}{Tx_{max} - Tx_{min}}, & for\ Tx_{min} < Tx_{Av} \leq Tx_{max} \\ 0, & otherwise \end{cases} \tag{4.12}$$

  Here, Tx_{Av} is the average transmitted load (Mbps), Tx_{min} is the minimum load (Mbps) and Tx_{max} is the highest load (Mbps). Th_{min} and Th_{max} are used to calculate Tx_{min} and Tx_{max} respectively. These are the required rates to achieve the lowest quality and the highest quality level, respectively.

4.3.2.2 Mobility Utility

The mobility utility uses a sigmoid curve to map the angular distance and speed of the user to a utility function. The angular distance can be described as the angle between the path of travel vector, and the straight line vector from the peer to the host in question. An angular distance of 0 is ideal while an angular distance of $\geq \pm 90$ deg or $\geq \pm\frac{\pi}{2}$ is treated as the worst case scenario. The value of ϵ is directly related to the midpoint of the linear curve which is considered to be at ± 45 deg or $\pm\frac{\pi}{4}$. X is a scaling constant for the speed and is adjusted based on the scenario. Finally, ζ is calculated by considering the point of inflection. At this point, the second derivative $u''_M = 0$ and substituting the values for $X, \epsilon, S = 0, \theta = \frac{\pi}{4}$ gives ζ. The final utility is shown in eq. (4.13).

$$u_M(\theta, S) = 1 - \frac{1}{1 + e^{\frac{\epsilon X - |\theta| S}{\zeta S}}} \tag{4.13}$$

where u_M is the mobility utility, θ is the angle between the path of travel and the host, S is the speed of the client, ϵ, ζ, X are constants for controlling the shape of the curve. In order to normalise the output of the function, the utility is divided by the maximum u_M which is considered to be at $\theta = 0$ for any speed S.

4.3.2.3 Host Mobility Utility

The host utility is combined with the mobility utility to create the *Host Mobility utility* U_{HM}. The new utility acts as a multiplicative weight to the existing host utility. The mobility utility will have no influence when the user is stationary. The host mobility utility is shown in eq. (4.14).

$$U_{HM} = U_H . u_M \tag{4.14}$$

Here, U_H is the Host utility from [105] and u_M is the novel mobility utility. U_{HM} is the host mobility utility used for selecting the best host to request from.

4.3.3 MENCO Algorithm

MENCO employs a novel algorithm, *Mobile Host Peer Selection (MHPS)*, that

Algorithm 4.3 Mobile Host Peer Selection (MHPS) Algorithm

Input: Host List $aMPD$; Quality level to request, QL;

Output: Best Host to request from, $bestHost$;

```
for all host in aMPD do
    Compute host.U_H using eq. (4.10)
    Compute host.u_M using eq. (4.13)
    Compute host.U_HM using eq. (4.14)
end for
Select bestHost=argmax(host.u_HM) from aMPD
if bestHost.U_HM < k_L then
    Hold QL
    Perform requestWait
else if bestHost.U_HM > k_H then
    Hold QL and bestHost
    Perform fastRequest
else
    Request from bestHost
end if
```

searches for the best host peer to request from based on their past performance as recorded in the local aMPD. Furthermore, MHPS introduces two new mechanisms to improve the overall performance, *requestWait* and *fastRequest*.

- **requestWait** Triggered when U_{HM} is below the lower threshold, k_L. This wait mechanism holds the clients current selected quality but waits until the next request period before selecting a host. When the hosts available to the peer have poor ratings according to the U_{HM}, the *requestWait* mechanism stops the peer from requesting content which will likely fail to be answered.

- **fastRequest** Triggered when U_{HM} is above an upper threshold, k_H. This mechanism is introduced to offset the reduced performance introduced by the Wait. It allows the client to hold the current selected quality and host in order to rapidly request each segment as the previous is received. This condition ends when the measured U_{HM} of the host drops below the threshold indicating that the selected host is going out of range or struggling to maintain high performance. Algorithm 4.3 provides a pseudo-code explanation of the host selection algorithm.

4.4 Chapter Summary

This chapter presents the overall system architecture proposed in this thesis as well as the details of the proposed solutions ENCO and MENCO. ENCO is introduced as a distributed multimedia solution which uses adaptive multimedia to not alone balance the quality and energy cost for the user but also to fairly share the cooperative workload across available neighbours. MENCO extends the architecture by introducing mobility awareness into the cooperative paradigm. It was observed the direction and speed of the user have an impact on the performance of the communication between devices. A smart neighbour selection mechanism is therefore introduced which takes these characteristics into consideration during host discovery and selection. The next chapter describes the testing performed and the results gathered to demonstrate the benefits of ENCO & MENCO.

Chapter 5

Performance Evaluation

This chapter presents the simulation-based testing scenarios and results and analysis of the main contributions introduced in Chapter 4: 1) A distributed Energy-Aware Cooperative delivery solution (ENCO); 2) A Mobile Quality-oriented Cooperative Multimedia Delivery Solution (MENCO). The evaluation of these solutions is carried out under various scenarios against different state-of-the-art solutions from the literature.

5.1 Performance Analysis of ENCO

In this section the performance of the proposed solution (ENCO) is analysed through simulation testing using NS-3. ENCO introduces a cooperative distributed environment in which each client device is responsible for adapting their own content to improve performance while also requesting and sharing content with neighbours. In this scenario, network infrastructures such as LTE and WiFi are considered unavailable due to overwhelming traffic or remote environments.

5.1.1 Simulation Scenarios and Settings

Simulation based testing is used to assess the performance of ENCO. The considered scenario consists of a user requesting multimedia content on a busy university campus under different circumstances. Two separate studies are performed as part of the evaluation. An overview of this scenario is represented by Figure 4.1 where the central device represents the user and the cloud contains the visible neighbouring peers in the users vicinity. The Network Simulator 3 (NS-3), version 3.25, is used to implement the solution and compare against two other works: BaseAmy [82] and DAV [89]. Each of these solutions proposes an adaptive multimedia approach to improve the experience for the user. BaseAmy focuses on improving energy efficiency while DAV highlights the potential of getting data from local peers. However, neither solution considers the full scenario as this study does. They have been selected as comparison studies due to the similar goals of the authors as well as the attention they have gained from the research community. BaseAmy uses the remaining battery and measured stream quality to adapt the requested quality to achieve the longest battery life possible. On the other hand, DAV predicts the available bandwidth using the previously measured throughput and chooses a suitable host peer based on a reputation generator located on a gateway device which considers the device specifications and the connection type (wireless/wired). However, in this study devices are considered to be of equal performance and all wireless, therefore the host peer is randomly chosen by DAV from the available peers.

$$PSNR = 20log_{10}\left(\frac{Exp_{Bitrate}}{\sqrt{(Exp_{Bitrate} - Meas_{Thr})^2}}\right) \tag{5.1}$$

Three configurations of ENCO are considered in the testing based on the adaptation weights w_Q and w_E. The Balanced ENCO configuration (B-ENCO) will adapt the requested quality in order to maintain a balanced trade-off between quality and energy consumption. In this case the two weights used are equal: $w_Q = w_E = 0.5$. The Throughput-focus ENCO (T-ENCO) involves adaptation which focuses on getting the highest throughput without considering the corresponding energy costs and employs $w_Q = 1, w_E = 0$. Finally the Energy-focused ENCO (E-ENCO) will aim at achieving the lowest energy consumption with no consideration for the received throughput and uses $w_Q = 0, w_E = 1$. These configurations are not optimal, just three suggestions for different types of users.

The performance of each solution is assessed using four different metrics: throughput (Mbps), energy consumption (J), PLR (%) and PSNR (dB). Also, in this simulation a packet is considered lost if it is not received by the start of the next segment. The PSNR is calculated using the quality estimation formula introduced by Lee at al. [106] as shown in Eq. 5.1. When using the PSNR formula, the adaptive nature of ENCO is considered by using the expected bitrate for each segment as the peak bitrate. Therefore, $Exp_{Bitrate}$ is the requested rate for the segment (Mbps) and $Meas_{Thr}$ is the measured throughput (Mbps). The settings used in the simulation can be seen in Table 5.1.

Two studies are performed to fully analyse the three ENCO configurations (B-ENCO, T-ENCO and E-ENCO) against BaseAmy and DAV in terms of performance. These are described below.

5.1.1.1 Study 1 - Effect of Number of Host Peers on Performance

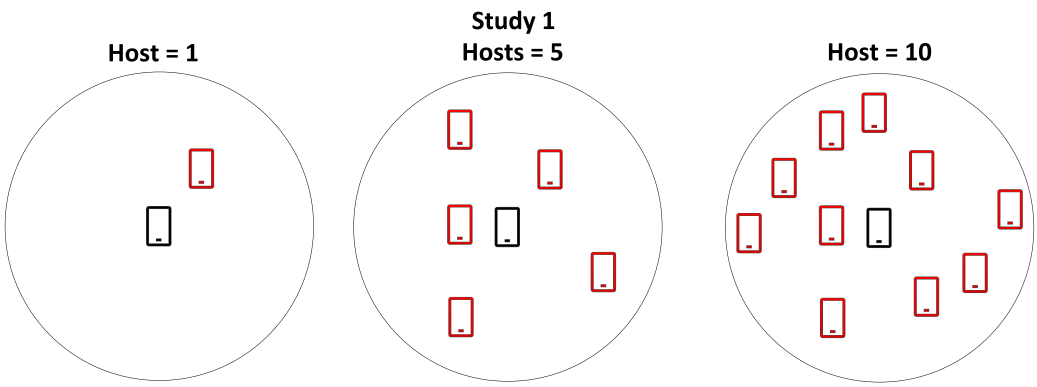

Figure 5.1: Study 1 - Simulation topology

Table 5.1: Simulation Settings

Parameter	Value
Peer Distribution	RandomDiscPositionAllocator, r=70m
Mobility Model	ConstantPositionMobilityModel
Wi-Fi Standard	WIFI_PHY_STANDARD_80211n_2_4GHZ
MAC Config	NqosWifiMacHelper Type = "ns3::AdhocWifiMac"
PHY Config	YansWifiPhyHelper, YansWifiChannelHelper Type = Default
Traffic Model	CBR (1920, 960, 480, 240, 120) kbps; Uniform Background Traffic
Simulation Length	150s
Segment Length	1s
Number of runs per test	10
Number of Host Peers	1; 5; 10
Number of Background Peers	20; 40; 60
r_t	0.6690961 W
r_d	0.0002377 J/kB

Table 5.2: Encoding Parameters

Video Quality Level	Resolution	Frame Rate (fps)	Bitrate (kbps)
QL1	800x448	30	1920
QL2	512x288	25	960
QL3	320x176	20	480
QL4	320x176	15	240
QL5	320x176	10	120

The first study performed focuses on the influence of the number of host peers on the performance of the solutions. For this study, a specific situation of the overall scenario is considered. The considered scenario involves a full lecture where the lecturer wishes to stream a video for the class. This video is part of the course and has been viewed by a handful of students prior to the class. It is assumed that the campus network is not currently available. Therefore, the lecturer has to stream the video with the help of the students in the lecture hall. Three tests are run for each of the five solutions considered with the number of host students changed for each run. The simulation topology for this can be seen in Figure 5.1. An ideal network environment is considered in this study with no additional traffic introduced.

5.1.1.2 Study 2 - Effect of Background Traffic on Performance

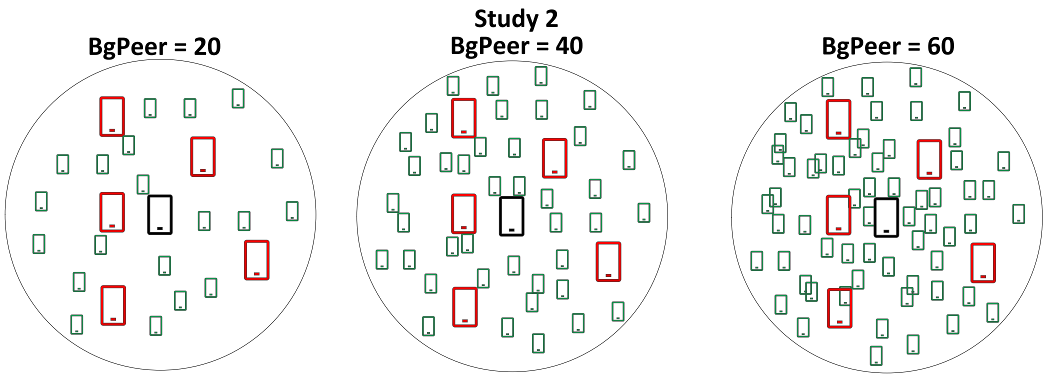

Figure 5.2: Study 2 - Simulation topology

A second study focuses on the effect of background traffic on performance. In this study, another scenario is considered within the overall scenario of a university campus. This time, a new video has recently been released on a social network (e.g. YouTube). A student wishes to watch the video but does not wish to spend money on LTE and is not allowed to use the campus network for such material. Instead, the student uses the ENCO architecture to get the segments from other students in the area who have already watched the video. However, as this is a public area of the campus, there is a lot of traffic with much congestion and interference from other people. Therefore, three tests are run for each solution where the number of host students is set constantly to 5. Background traffic is introduced by adding additional background peers (passers-by) capable of generating traffic. This traffic is goes from the minimum to maximum rates available for requested content. The number of background peers was set to 20, 40 and 60 across the three tests. The simulation topology for this study can be seen in Figure 5.2.

5.1.2 Results and Analysis

5.1.2.1 Study 1 - Effect of Number of Host Peers on Performance

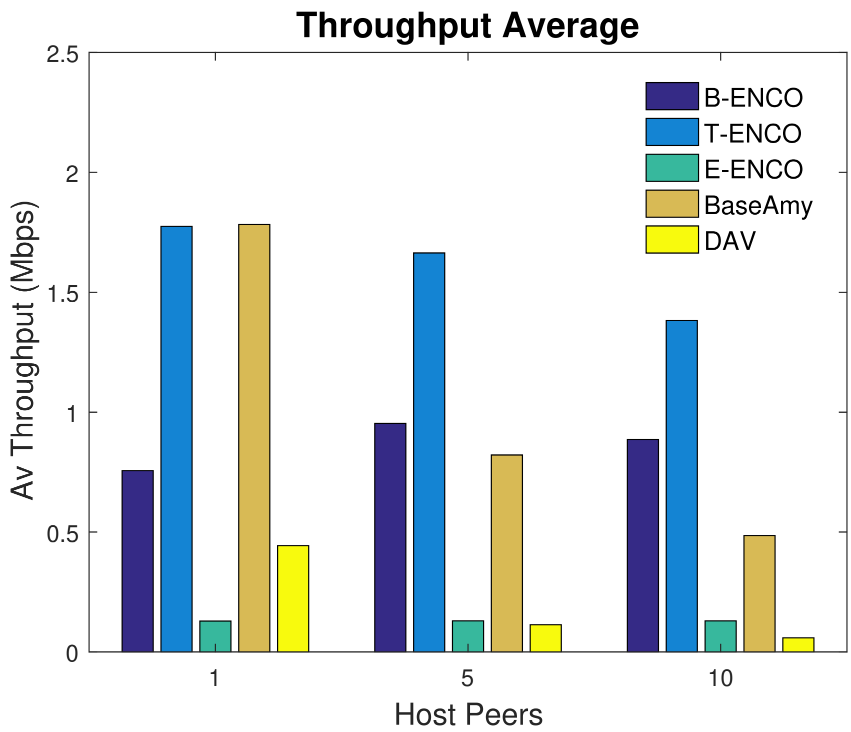

Figure 5.3: Average Throughput

The results for study 1 are shown in Figures 5.3 - 5.6. It can be seen that, as the

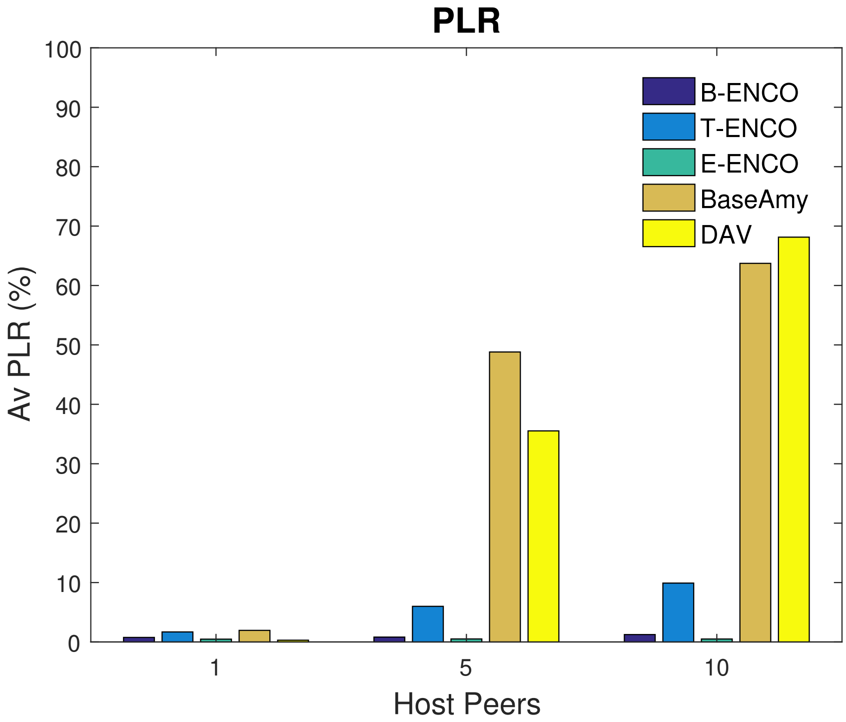

Figure 5.4: Average PLR

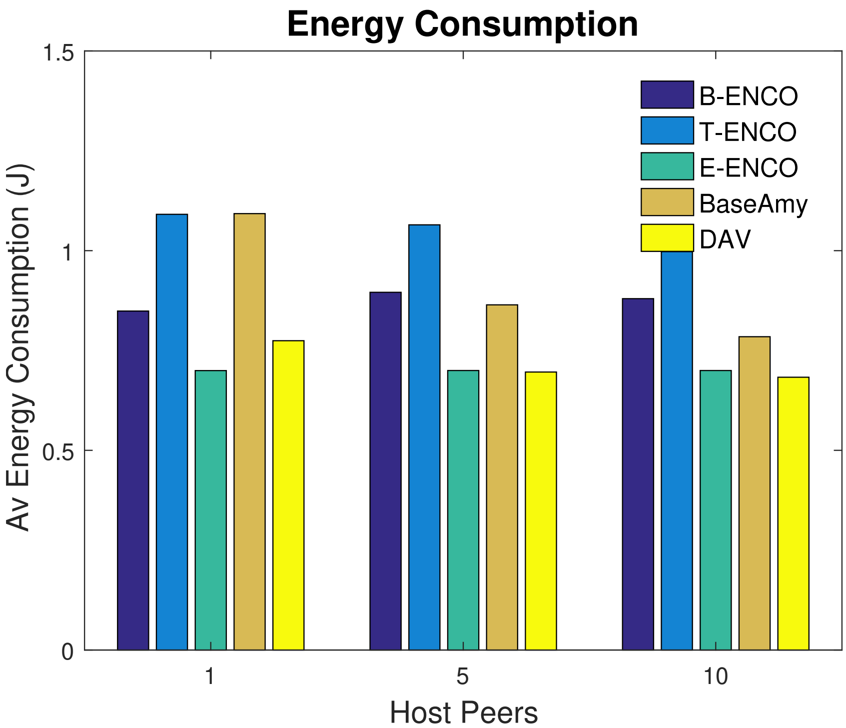

Figure 5.5: Average Energy

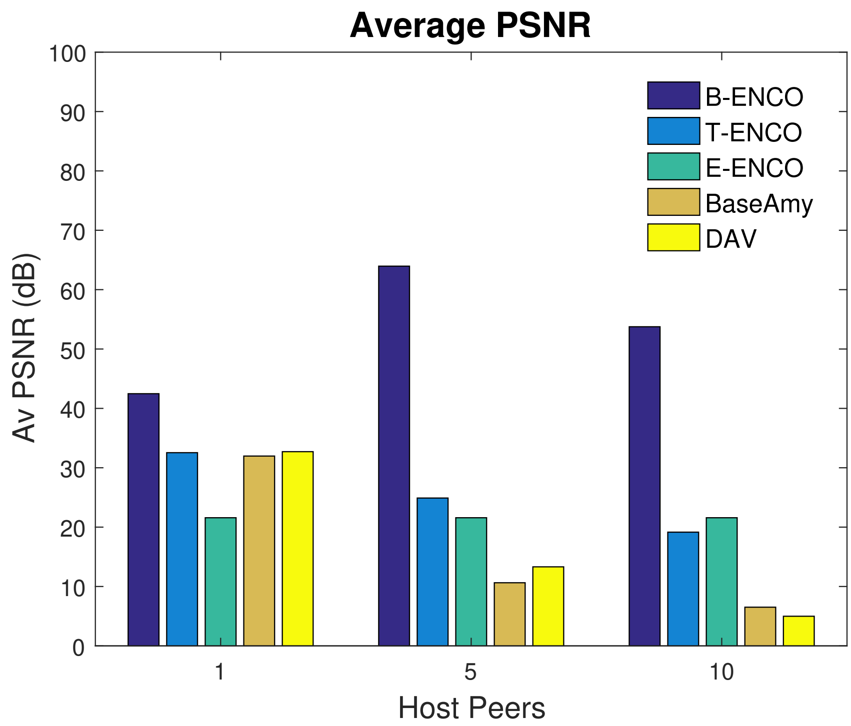

Figure 5.6: Average PSNR

number of available hosting peers increases, the performance of B-ENCO is stable with the lowest PSNR value being approximately 40 dB for 1 host and exceeding 60 dB with additional hosts. On the other hand, the performance of both BaseAmy and DAV rapidly deteriorates with both reaching over 60% PLR and peaking at an approximate PSNR of 30 dB. The lack of neighbour awareness in these solutions, in a highly decentralised environment, leads to random host choice with no consideration for host performance or location.

Looking at the energy performance of the solutions, DAV manages to consistently achieve lower consumption than B-ENCO and equivalent to that of E-ENCO. BaseAmy manages to outperform B-ENCO by up to 12% as the number of peers increases. However, this small reduction in energy cost comes with a large performance cost. Considering the 5 peers scenario, B-ENCO has higher energy cost than both DAV and BaseAmy. However, for the same scenario, B-ENCO outperforms BaseAmy and DAV with an average PSNR of 63.96 dB compared to 10.63 dB and 13.31 dB respectively.

5.1.2.2 Study 2 - Effect of Background Traffic on Performance

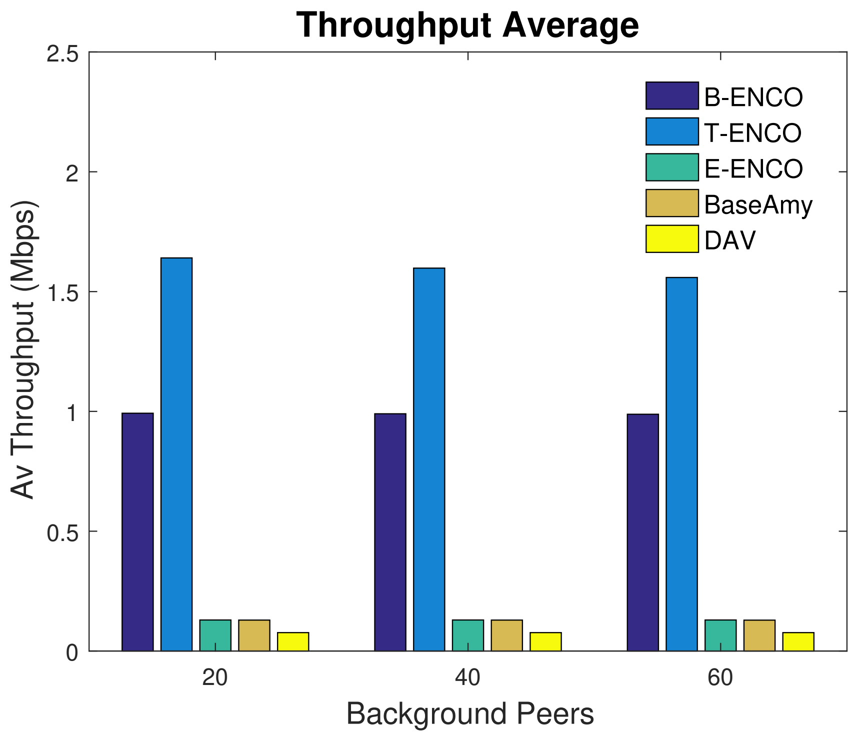

Figure 5.7: Average Throughput

Study 2 looks at the impact of background traffic on the performance of each of the solutions, with the results illustrated in Figures 5.7 - 5.10. It is immediately clear that the additional traffic has little effect on any of the ENCO configurations with B-ENCO, T-ENCO and E-ENCO performing nearly as well as in study 1. However, BaseAmy and DAV both suffer greatly with BaseAmy reaching nearly 80% PLR and DAV exceeding 40%. Again, the lack of neighbour awareness is a major downfall for BaseAmy in this largely decentralised scenario and DAV suffers without access to a central gateway for proper host decisions.

It is noted that BaseAmy and DAV have lower energy consumptions than B-

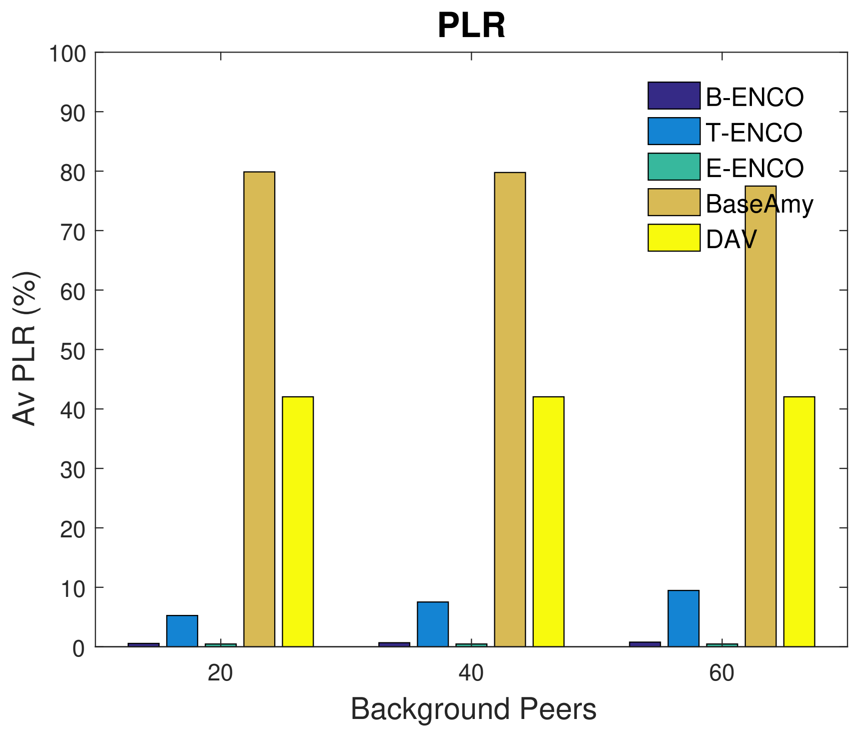

Figure 5.8: Average PLR

ENCO. However, E-ENCO is nearly equal with DAV in this regard (i.e DAV uses 1.82% less energy than E-ENCO). Meanwhile, E-ENCO's PSNR is substantially higher than either of the alternative solutions. Considering the 40 background peer scenario, E-ENCO has an average PSNR of 21.58 dB, while BaseAmy and DAV have 1.71 dB and 8.88 dB, respectively. The reason the comparatives achieve low energy consumption is mainly because of the small throughput they are achieving.

5.1.3 ENCO Complexity Analysis

Considering the CQS algorithm 4.1, it is clear that it does not rely on any input length. This algorithm takes a utility and a threshold value and computes the best quality level. Therefore, the complexity associated with this algorithm is constant, or using O notation, O(1).

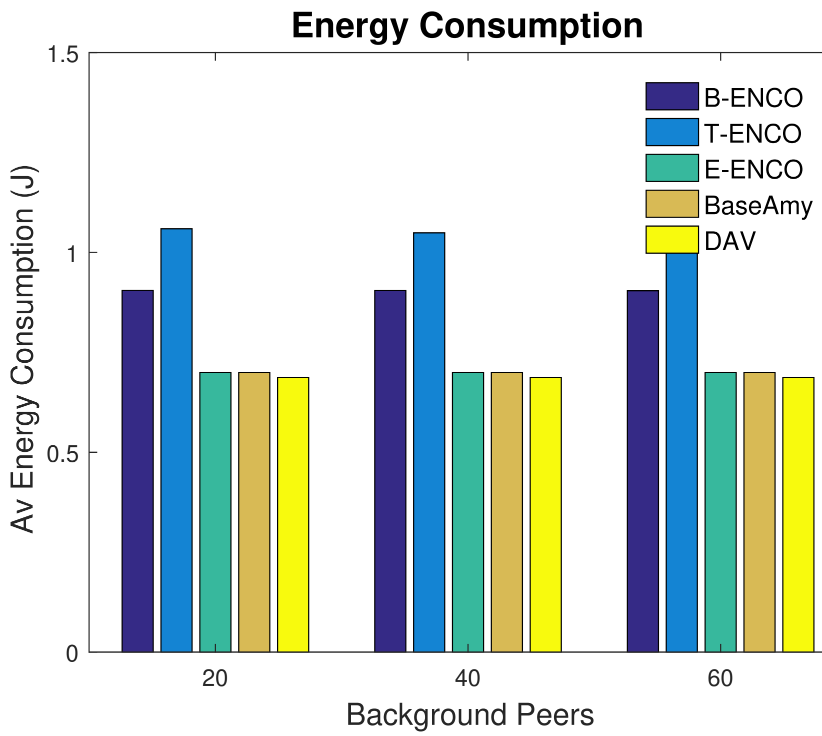

Figure 5.9: Average Energy

The HPS algorithm 4.2 is responsible for traversing the aMPD file and identifying the most suitable host to request the content from at the desired quality level. The worst case scenario here would be if the desired QL is the highest level, and no peer can be found in the aMPD. Considering HPS has a for loop nested in a while loop, the complexity will rely on how many iterations will be invoked. The inner loop goes through the whole aMPD file, lets consider this file to have N entries. The outer loop goes through each QL, let's consider there to be M quality levels. Typically M will be significantly smaller than N. From this, we can estimate a worst case complexity of O(MN).

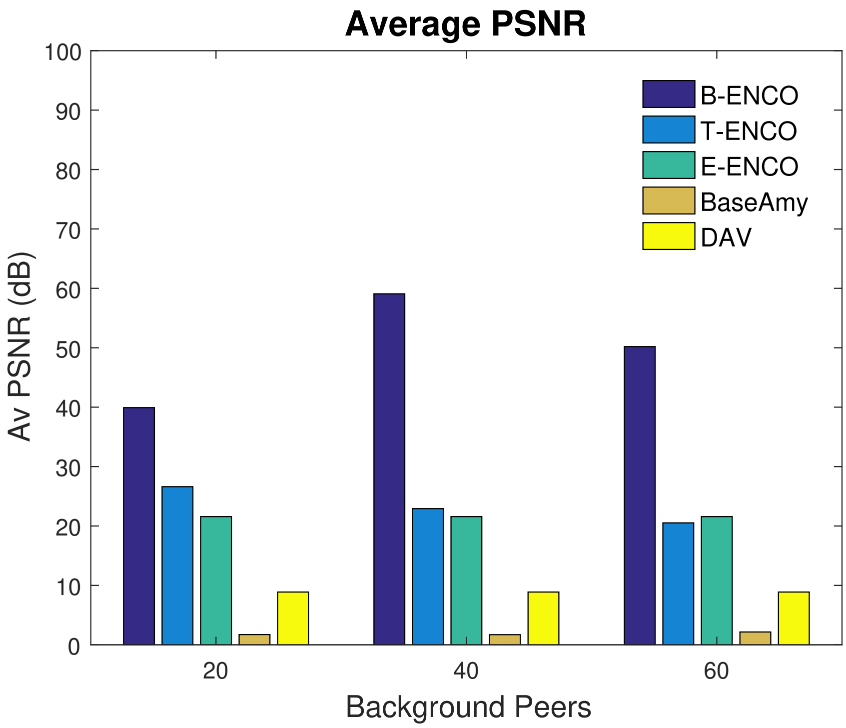

Figure 5.10: Average PSNR

5.2 Performance Analysis of MENCO

This section introduces the performance testing for the MENCO solution. Different scenarios are presented which were designed to test a variety of conditions for MENCO while comparing against other models. For the purpose of this work, the users relies on D2D communications as other networks are unavailable due to remoteness and/or congestion.

5.2.1 Simulation Scenarios and Settings

A basic initial scenario is created in NS-3 to simulate a client travelling through a busy environment while viewing a video on their mobile device. The density of the scenario is considered to be 1000 persons per km^2 to reflect a small urban environment with users viewing similar content to the user. Three speeds are con-

Table 5.3: Simulation Settings

Parameter	Value
Peer Distribution	RandomBoxPositionAllocator
Peer Density	1000 Persons per sq. km.
Host Mobility Model	ConstantPositionMobilityModel
Client Mobility Model	ConstantVelocityMobilityModel
Wi-Fi Mode	IEEE 802.11n 2.4GHz Ad-hoc
Simulation Length	100s
Segment Duration	1s
Number of runs per test	5
Client Speed (km/hr)	0; 25; 50
Video Rates (Mbps)	0.12; 0.24; 0.48; 0.96; 1.92
X	15
k_L	0.1
k_H	0.9

sidered as part of the study, $0km/hr$ or stationary, $25km/hr$ and $50km/hr$. This initial scenario is considered as a basic measure of the performance of MENCO, hence the choice of low speeds. Each test is run five times with different simulation seeds for a duration of 100s each.

Two studies are then performed on the performance of MENCO under different conditions.

- *Study 1: Impact of aMPD range* - For this study, the range of awareness maintained by the local aMPD, as described in Section 4, is varied between 20, 50 and 80 meters. The idea behind this is a smaller range will result in fewer neighbours being searched during the algorithms. However, this also means there are fewer locations to request content from. On the other hand, a large range will cost most resources to maintain and may result in requesting from far away hosts which fail to respond.

- *Study 2: Impact of change in speed* - When the client is travelling at different speeds, the performance of the multimedia stream will be affected depending on the location of hosts in relation to the client as well as other factors. This study observes the performance of the proposed MENCO solution under three different speeds, $0km/hr$ or stationary, $40km/hr$ and $80km/hr$.

To assess the performance of the proposed solution, four main metrics are considered: throughput (Mbps), PSNR (dB) as estimated using the formula introduced in [106] eq. (5.1), delay (ms) and energy (J). The settings used in the simulation can be seen in Table 5.3.

5.2.2 Results and Analysis

Simple Scenario

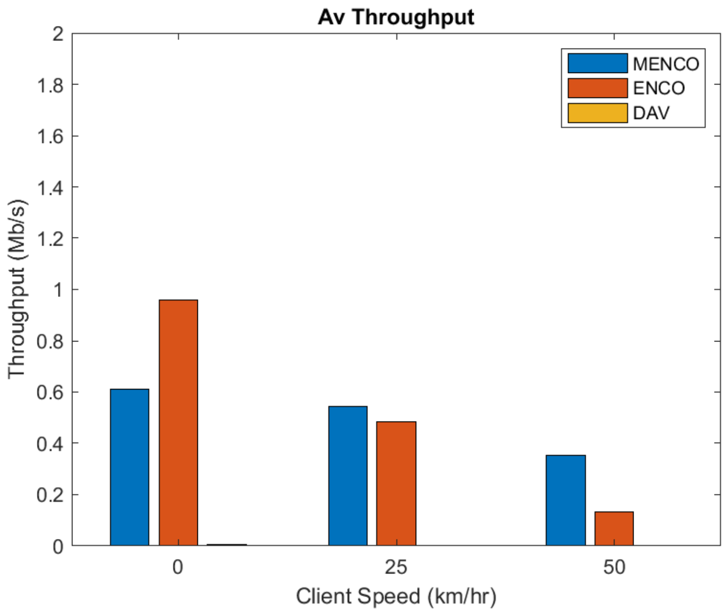

Figure 5.11: Average Throughput Results

In Fig. (5.11), the average throughput for each scenario is shown. Considering the stationary scenario, it is clear that ENCO is outperforming the proposed

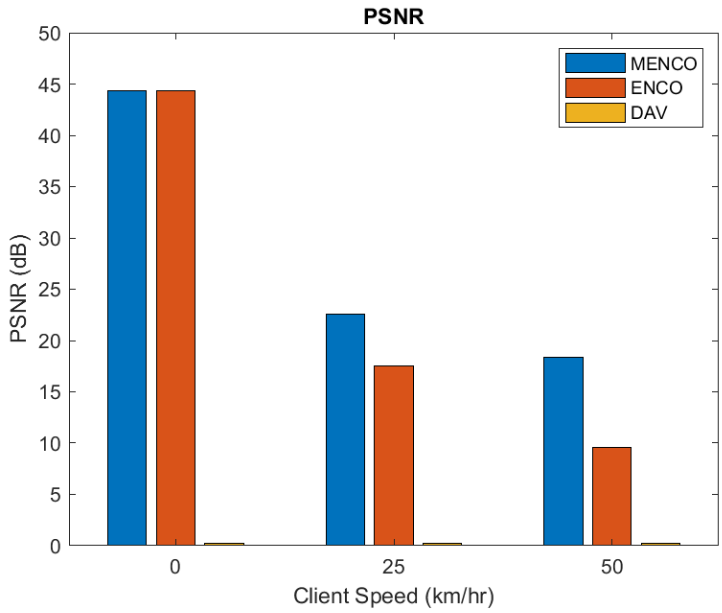

Figure 5.12: Average PSNR Results

solution. This is due to the modified distance utility and the Wait mechanism. ENCO will always send a request once a host is found, while MENCO will only send a request when the host is strong enough. This can lead to situations were ENCO can more freely request data than MENCO. However, once mobility is introduced, ENCO suffers from this. Further devices now become quickly unreachable leading to large timeout delays in the TCP stream. Meanwhile, MENCO only commits to a host if its performance and location are suitable, as described in Algorithm 4.2. MENCO marginally outperforms ENCO in the $25km/hr$ scenario but almost triples the throughput of ENCO in the $50km/hr$ scenario. Most importantly, MENCO holds a reasonably consistent performance across the scenarios. DAV struggles in these tests due to its throughput averaging mechanism and lack of proper awareness of the distributed environment.

Looking at the PSNR values achieved in Figure. 5.12, MENCO and ENCO perform equally when stationary. This occurs despite ENCO having higher average throughput. It can be concluded from this that the MENCO fast request mechanism is working to rapidly retrieve segments at stable rates in order to maintain quality. This brings the observed throughput down due to the maintained quality level. ENCO is only requesting segments as needed, more accurately reflecting the available throughput. Both MENCO and ENCO drop approximately 50% in PSNR when mobility is introduced. However, MENCO maintains a stable value around 20dB between $25km/hr$ and $50km/hr$, unlike ENCO which drops a further 50%.

Study 1: Impact of aMPD range

The change in range was considered for three different speed scenarios as it was expected that the performance impact of the range would be influenced by the speed. The three speed scenarios considered for this study were, $0km/hr$, $40km/hr$ and $80km/hr$.

This first study was conducted to analyse the impact of changing the range awareness of the aMPD file used in the MENCO solution. Three speed scenarios are considered and the results are illustrated in the following figures.

Looking at the throughout results shown in Figure 5.13, a difference in performance between the ranges can be observed. At the shortest value, the throughput struggles even at stationary. As the value increases it can be seen that the performance improves with the mid and highest values performing equally as well.

Figure 5.14 illustrates the Average PSNR results gathered during the testing for this study. It is clear that in the stationary scenario, each range performs equally well which would be expected. However, once mobility is introduced, each range loses performance as the speed increases. Out of the three ranges, the middle value, $50m$, performs the best especially around the $40km/hr$ rate.

The delay results shown in Figure 5.15 are interesting. Delay would be expec-

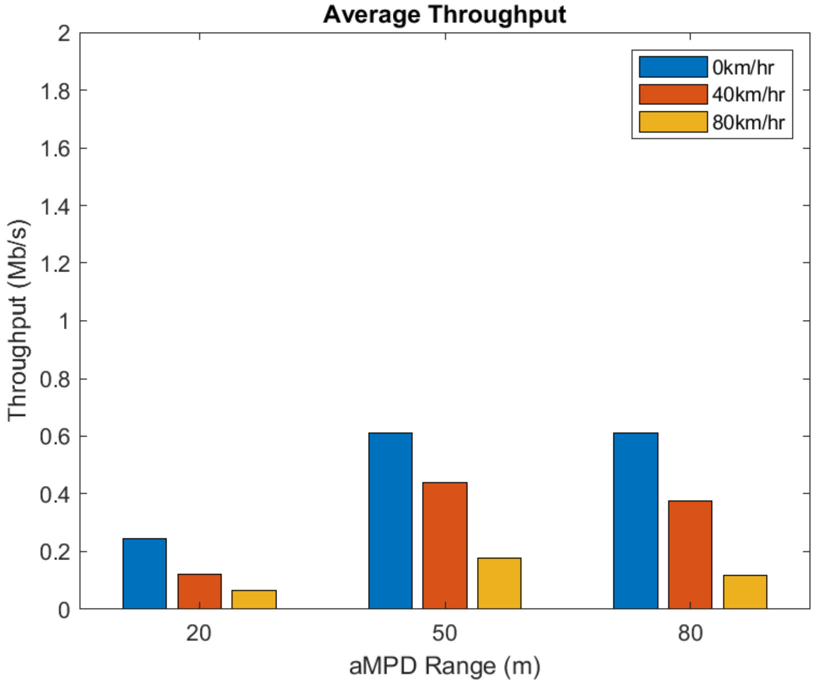

Figure 5.13: Study 1 Average Throughput

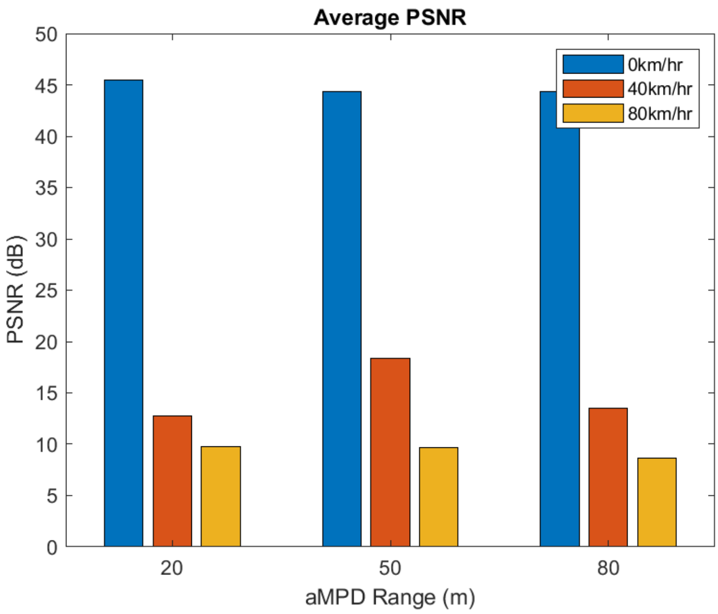

Figure 5.14: Study 1 Average PSNR

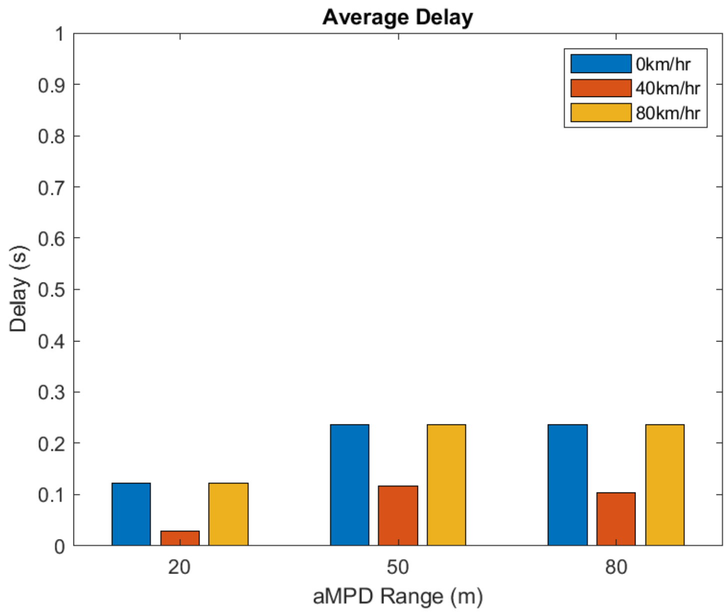

Figure 5.15: Study 1 Average Delay

ted to follow the throughput behaviour. The results however are strange because delay is relatively high in the 50m stationary scenario despite the throughput being reasonable. Meanwhile a low throughput for the 20m 80km/hr scenario achieved a lower delay, although still higher than expected. The cause for this behaviour seems to be the mobility of the device. At low speeds, the device experiences high delays despite achieving the strongest throughputs. While at the highest speeds, similar delays are experienced despite having low throughputs. Resulting from this behaviour, the middle speed of 40km/hr provides the best trade-off with the 50m range slightly outperforming 80m.

Energy costs behave as expected in Figure 5.16 with the highest energy costs occurring alongside the highest throughput levels. It is also noted that as the range and speed are increased, the energy is impacted with higher values having a higher energy cost. While the difference is not large between the 50m and 80m configurations, there is a considerable jump between 20m and 50m.

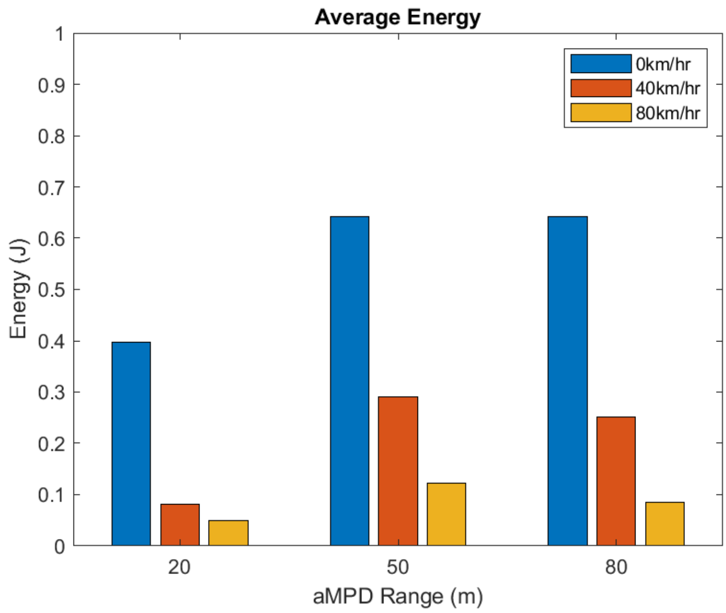

Figure 5.16: Study 1 Average Energy

Summary

Study 1 observed the impact of the aMPD awareness range on the performance of the device. It is shown that the performance is impacted considerably as the range is changed from 20m to 50m to 80m. Across the four metrics, the 50m configuration provides the most stable performance with the best throughput for each speed and the lowest delay achieved at $40km/hr$. This will be used in the next study as the constant aMPD range.

Study 2: Impact of change in speed

The impact of speed on the performance of the user device is considered in this study. Four models are compared in the results: 'MENCO50' is MENCO with a range of 50m as decided in Study 1, a balanced ENCO configuration, DAV [89] a host selection mechanism for MPEG-DASH, and a base scenario which chooses whichever host it can.

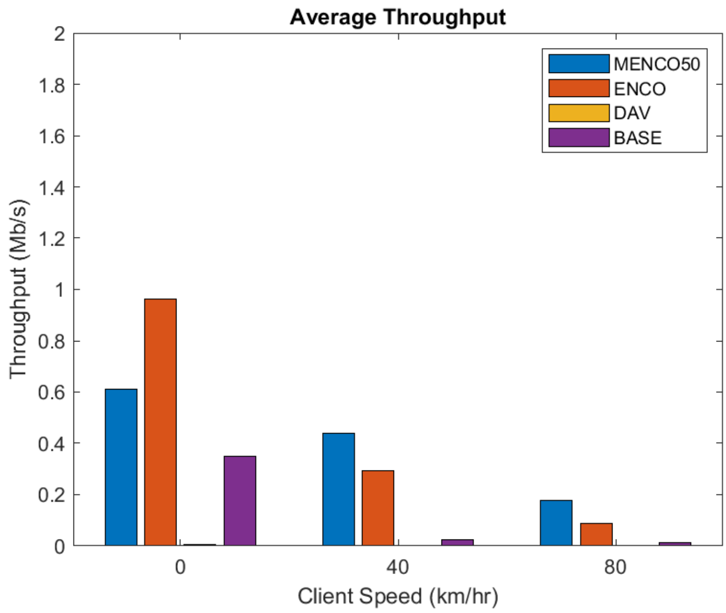

Figure 5.17: Study 2 Average Throughput

The throughput results represented in Figure 5.17 show that MENCO is outperformed in the stationary scenario. However, as speed is introduced MENCO rapidly takes over and comfortably outperforms the other solutions as speed increases up until higher speeds of $80km/hr$ where each solution struggles. The BASE solution performs reasonably well in the stationary scenario but once mobility is introduced it struggles to keep up.

In Figure 5.18 MENCO and ENCO perform equally well across the three scenarios with MENCO having an upper hand overall. Again BASE can be seen performing well in the stationary scenario with a respectable PSNR of approximately 30dB. As the speed increases, BASE manages to maintain its PSNR at around 12dB by requesting the lowest quality available and therefore suffering only minor loss of data.

Results for delay can be found in Figure 5.19. The results are mostly as expected with higher delays observed at times of higher PSNR and throughput. Large delay

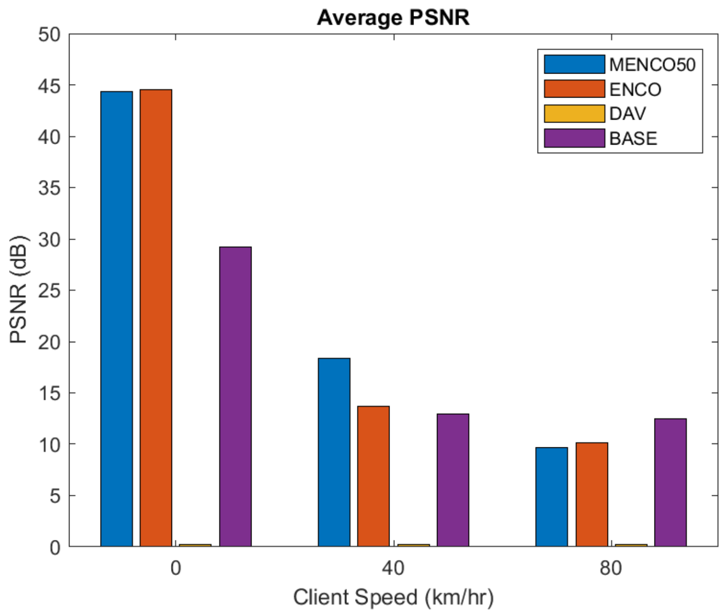

Figure 5.18: Study 2 Average PSNR

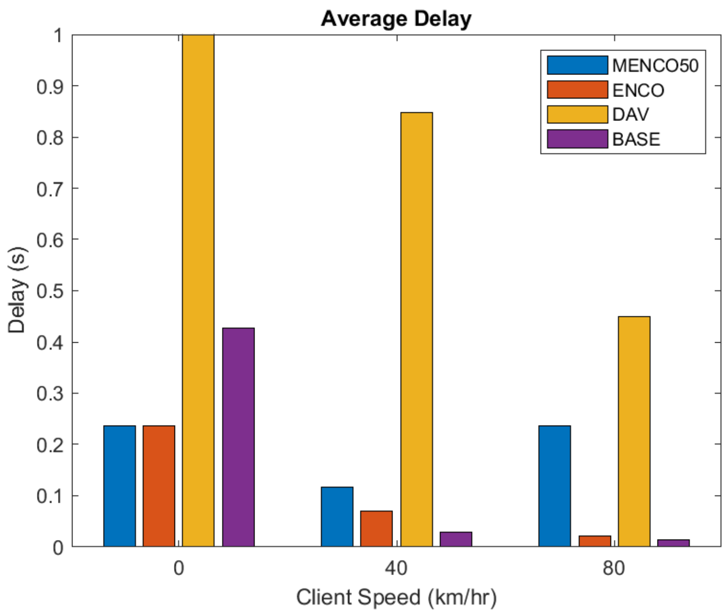

Figure 5.19: Study 2 Average Delay

spikes also occur frequently for DAV due to the poor throughput that the solution achieves. It gets stuck waiting indefinitely. One strange behaviour that can be observed is a spike in delay for MENCO50 at 80km/hr. A behaviour that was observed during testing involved MENCO getting stuck due to the TCP back off mechanism. More specifically, if this occurred while travelling away from a host, if the host goes out of range then the TCP back off causes a long hang time for the user device. This is most likely to be the cause for this delay spike.

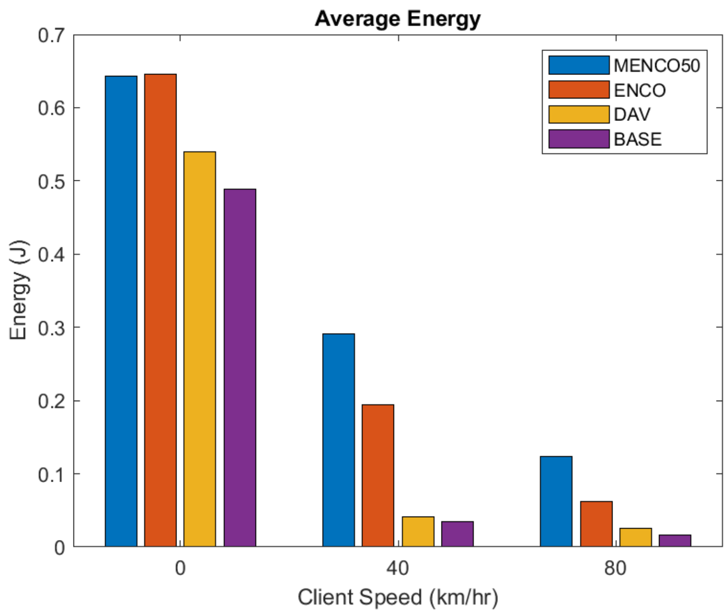

Figure 5.20: Study 2 Average Energy

The energy results are represented in Figure 5.20. It can be seen that in the stationary scenario each of the solutions has a similar energy cost. As the speed increases, DAV and BASE go to near 0 due to low performance at speed, meaning nothing is happening to consume energy. MENCO performs well and reduces energy cost as the speed increases, however ENCO shows it can maintain a balance between the quality it is achieving and the energy cost.

Summary

This Study looked at the impact of speed on the performance of MENCO when compared to other works from the literature. MENCO outperforms DAV and the base scenario comfortably across all scenarios and performed best in high mobility scenarios when compared to ENCO.

5.2.3 MENCO Complexity Analysis

MENCO again makes use of the CQS algorithm to determine the QL to be requested, but as previously described, this has a constant complexity of O(1). So to determine the complexity of MENCO, the MHPS algorithm 4.3 must be analysed. It is immediately clear that this algorithm depends on the size of the aMPD. In the worst case scenario, it is assumed that the client has no prior knowledge of the aMPD contents and therefore must compute each value. Considering a single loop through the file, where the length of the file is assumed to be N hosts, it is clear that the complexity of this solution is O(N).

5.3 Chapter Summary

This chapter introduces the testing carried out on the contributions proposed in this thesis. ENCO is a energy aware cooperative solution which strives to balance quality and energy in a mobile distributed environment. Utility functions are designed to enable an adaptive solution and to allow the client to select a host to request from by considering neighbouring devices. Throughput, energy, load and distance are all considered as part of the solution. Simulation based testing showed that ENCO outperforms comparative studies in terms of PSNR while maintaining a balance between quality and energy.

MENCO is a mobility aware cooperative streaming solution which introduces a mobile host selection mechanism capable of considering mobility during the selection process. The novelty of MENCO comes from the novel mobility awareness that allows it to not only consider the load and distance of a host, but also the relative speed and direction of travel between the client and potential host. This allows MENCO the select hosts based on those that can actually satisfy the clients needs in the time they will be in range of each other. Testing using the NS3 simulator prove the MENCO works and selects the best host as it is travelling while maintaining an acceptable PSNR value.

Chapter 6

Conclusion

This chapter summarises the work introduced in this thesis. The three main contributions are summarised along with their related results. Future directions of work are then highlighted.

6.1 Problem Overview

Multimedia is rapidly becoming the largest contributor to global internet traffic. Also, more of this traffic is being delivered to mobile devices than ever before. With the growth of the performance of these devices, this trend will only get worse as users expect more. Current networking solutions are not capable of supporting the demand of this rapidly growing traffic.

Another point of interest is the slow advancement of battery technologies in mobile devices. Despite the performance rapidly growing, the battery capacity is increasing very slowly. This leads to a bottleneck between performance and energy. Most mobile users struggle to get through a full day of moderate use without the need to charge their devices. When multimedia streaming is introduced, this becomes even worse.

With 5G soon to be a common mobile technology for users, devices will be more connected than ever before. Not only will 5G bring improved speeds and capacities, it will also connect more devices than ever before. One of the major categories of use for 5G will be connecting devices to each other to help realise the Internet of Everything (IoE). This improved connectivity will only help to encourage and promote cooperative approaches, which will be crucial if users want to be able to sustain the ever growing amount of data required to stream their desired content.

From the above observations, the problems analysed within this thesis can be defined by three main points. Firstly, there is a need for smarter networking solutions that can keep up with the rapid traffic growth being experienced. Emerging and future technologies such as 4k, 8k and Virtual reality are only going to introduce more issues. A cooperative approach is considered as a viable solution which would fit well with current future network proposals. Secondly, more consideration needs to be given to energy efficiency, especially on mobile devices. The communications industry is currently one of the largest contributors to the global carbon footprint. Furthermore, mobile users are struggling to get a full days use out of their mobile devices. Finally, it is essential for a balance between energy and performance to be achieved for mobile users.

6.2 Contributions

Existing solutions and standards have been analysed as a part of this thesis and based on the findings two main contributions are proposed from this thesis:

1. This first contribution is a novel energy aware cooperative multimedia delivery (ENCO) solution which monitors throughout and energy during a stream. Using this information ENCO adapts the multimedia stream to achieve a balance between quality and energy. Furthermore, ENCO monitors load and distance of neighbouring devices to allow the client to select suitable

devices to request content from without overloading them or experiencing loss due to distance. ENCO manages to not only achieve a balance between energy and quality in a distributed cooperative environment, but also promotes fairness between host devices by considering historical load as part of the selection process.

2. A mobile aware Quality-oriented Cooperative Multimedia Delivery (MENCO) solution is the second contribution introduced in this thesis. MENCO brings a novel host selection mechanism to a distributed heterogeneous cooperative environment. By considering the velocity between devices, MENCO is able to determine if a host will stay in range for a suitable amount of time before making a request. This reduces the risk of requests being made to hosts which are lost therefore causing considerable delays. Results show that MENCO performs well across a variety of speeds with throughput improvements of nearly 50% achieved.

6.3 Future work

Both of the solutions proposed make up the fundamentals of a purely distributed cooperative architecture which is aware of the individual users. The next step would be to build a real test bed to enable practical results to be gathered as well as to perform a subjective study. Further on from this, a real world implementation would provide even more useful data including user behaviours.

The work in this thesis only considers multimedia traffic, but users are regularly sharing data other than just video. A future study would be to analyse the behaviour of different traffic types across the proposed architecture. Using this data, smart methods such as Multi-path TCP could be leveraged to enable multiple hosts to stream content simultaneously, increasing the capacity of the network.

Of course, the cooperative paradigm is only capable of performing as long as the underlying devices agree to cooperate. A big direction of future work would be to introduce an incentive mechanism to encourage clients to participate and to not take advantage of others. People will naturally have their own interests in mind, so such an incentive would need to provide something in return for their cooperation.

ENCO and MENCO have been shown to perform well, but in a heterogeneous mobile environment there are a limitless number of influencing factors that can be considered to further improve performance. The obvious problem here is the rapid growth in complexity of such solutions. To keep a distributed solution, a two layered solution could be taken where the devices can make a basic decision themselves but a more complex solution could be situated at the cloud level. This would allow a more sophisticated method such as machine learning to be leveraged under certain conditions while maintaining a distributed solution.

Bibliography

[1] Cisco, 'Forecast and Methodology, 2016-2021, White Paper', Tech. Rep., 2016, p. 22. arXiv: `1454457600809267`.

[2] ——, 'The Zettabyte Era: Trends and Analysis', Tech. Rep. June, 2017, pp. 1–29.

[3] Ericsson, 'Ericsson Mobility Report', Tech. Rep. June, 2017.

[4] OEM Off Highway, 'Moving Beyond Lithium-Ion', Feb. 2016.

[5] J. Janek and W. G. Zeier, 'A solid future for battery development', *Nature Energy*, vol. 1, no. 9, p. 16 141, 2016.

[6] Z. Specification, 'ZigBee Alliance', *ZigBee Document 053474r06, Version*, vol. 1, 2006.

[7] B. Specification, 'Version 1.0', *Technical specification, URL https://trustedcomputinggroup.org*, p. 26, 2003.

[8] M. WiMAX-Part, 'I: A technical overview and performance evaluation', in *WiMAX forum*, vol. 1, 2006.

[9] I. Committee *et al.*, 'Wireless LAN medium access control (MAC) and physical layer (PHY) specifications', *IEEE Std*, vol. 802, p. 50, 1997.

[10] R. M. Metcalfe and D. R. Boggs, 'Ethernet: distributed packet switching for local computer networks', *Communications of the ACM*, vol. 19, no. 7, pp. 395–404, 1976.

[11] IEEE 802.11 Working Group and others, 'Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band', Tech. Rep., 2003.

[12] 'Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band', *IEEE Std*, 2012.

[13] ——, 'Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendmeny 2: Sub 1 GHz License Exempt Operation', *IEEE Std*, 2016.

[14] ——, 'Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications High-speed Physical Layer in the 5 GHz Band', *IEEE Std*, 1999.

[15] ——, 'Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz Band', *IEEE Std*, 2003.

[16] ——, 'Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 5: Enhancements for Higher Throughput', *IEEE Std*, 2009.

[17] 'Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz', *IEEE Std*, 2013.

[18] B. Bellalta, 'IEEE 802.11ax: High-efficiency WLANS', *IEEE Wireless Communications*, vol. 23, no. 1, pp. 38–46, 2016. arXiv: `1501.01496`.

[19] Y. Ghasempour, C. R. Da Silva, C. Cordeiro and E. W. Knightly, 'IEEE 802.11ay: Next-Generation 60 GHz Communication for 100 Gb/s Wi-Fi', *IEEE Communications Magazine*, vol. 55, no. 12, pp. 186–192, 2017.

[20] A. A. Goldberg, 'PCM NTSC television characteristics', *SMPTE Journal*, vol. 85, no. 3, pp. 141–145, 1976.

[21] H. V. Sims, *Principles of PAL colour television and related systems*. Newnes-Butterworths, 1969.

[22] G. B. Townsend, 'New developments in the SECAM colour television system', *The Television Society Journal*, vol. 10, no. 6, pp. 172–190, 1963.

[23] E. N. Etsi, '302 307:" Digital Video Broadcasting (DVB)', *Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications*, vol. 1, p. 2, 2006.

[24] A. ATSC, '52/10]: United states advanced television systems committee digital audio compression (ac-3) standard, doc', A/52/10, December, Tech. Rep., 1995.

[25] H. Asami and M. Sasaki, 'Outline of isdb systems', *Proceedings of the IEEE*, vol. 94, no. 1, pp. 248–250, 2006.

[26] J. Song, Z. Yang, L. Yang, K. Gong, C. Pan, J. Wang and Y. Wu, 'Technical review on Chinese digital terrestrial television broadcasting standard and measurements on some working modes', *IEEE Transactions on Broadcasting*, vol. 53, no. 1, pp. 1–7, 2007.

[27] U. Reimers, 'DVB-T: the COFDM-based system for terrestrial television', 1996.

[28] G. Jo, *Digital Multimedia Broadcasting*, 2007.

[29] J. Lehenkari and R. Miettinen, 'Standardisation in the construction of a large technological system - The case of the Nordic mobile telephone system', *Telecommunications Policy*, vol. 26, no. 3-4, pp. 109–127, 2002.

[30] V. H. Mac Donald, 'Advanced mobile phone service: The cellular concept', *The bell system technical Journal*, vol. 58, no. 1, pp. 15–41, 1979.

[31] Jay E. Padgett, Christoph G. Gunther and T. Hattori, 'Overview of Wireless Personal Communications', *IEEE Communications Magazine*, no. January, 1995.

[32] M. Mouly and M.-B. Pautet, *The GSM System for Mobile Communications*. Telecom Publishing, 1992.

[33] H. D. Rate, 'cdmaOne optimized for high speed, high capacity data', *Wireless Infrastructure, Qualcomm*, 1998.

[34] R. J. Bates, *GPRS: general packet radio service*. McGraw-Hill Professional, 2001.

[35] A. Furuskär, J. Näslund and H. Olofsson, 'Edge - enhanced data rates for GSM and TDMA/136 evolution', *Ericsson Review (English Edition)*, vol. 76, no. 1, pp. 28–37, 1999.

[36] T. Ojanpera and R. Prasad, 'An overview of air interface multiple access for IMT-2000/UMTS', *IEEE Communications Magazine*, vol. 36, no. 9, pp. 82–86, 1998.

[37] D. N. Knisely, Q. Li and N. S. Ramesh, 'Cdma2000: A third-generation radio transmission technology', *Bell Labs Technical Journal*, vol. 3, no. 3, pp. 63–78, 1998.

[38] E. Dahlman, S. Parkvall, J. Skold and P. Beming, *3G evolution: HSPA and LTE for mobile broadband*. Academic press, 2010.

[39] H. Holma, A. Toskala and P. Tapia, *HSPA+ Evolution to release 12: performance and optimization*. John Wiley & Sons, 2014.

[40] D. Astely, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom and S. Parkvall, 'LTE: the evolution of mobile broadband', *IEEE Communications Magazine*, vol. 47, no. 4, pp. 44–51, Apr. 2009.

[41] E. Dahlman, S. Parkvall and J. Skold, *4G: LTE/LTE-advanced for mobile broadband*. Academic press, 2013.

[42] ITU-T, 'SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE, INTERNET PROTOCOL ASPECTS AND NEXT-GENERATION NETWORKS, INTERNET OF THINGS AND SMART CITIES Next Generation Networks – Enhancements to NGN', *International telecommunication union*, 2016.

[43] D. Roddy, *Satellite communications*. McGraw-Hill Prof Med/Tech, 2006.

[44] F. Tobagi and J. Pang, 'StarWorks-a video applications server', in *Digest of Papers. Compcon Spring*, IEEE Comput. Soc. Press, 1993, pp. 4–11.

[45] K. Brandenburg and G. Stoll, 'Iso/mpeg-1 audio: A generic standard for coding of high-quality digital audio', *Journal of the Audio Engineering Society*, vol. 42, no. 10, pp. 780–792, 1994.

[46] H. Schulzrinne, 'Real time streaming protocol (RTSP)', 1998.

[47] V. Jacobson, R. Frederick, S. Casner and H. Schulzrinne, 'RTP: A transport protocol for real-time applications', 2003.

[48] D. Wing, 'Symmetric RTP/RTP Control Protocol (RTCP)', 2007.

[49] H. Schwarz, D. Marpe and T. Wiegand, 'Overview of the scalable video coding extension of the H.264/AVC standard', *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[50] O. Forum, 'Release 2 Specification Volume 2a - HTTP Adaptive Streaming', vol. 2, pp. 1–25, 2011.

[51] T. Wiegand, 'Overview of the H. 264/AVC video coding standard', *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, 2003.

[52] T. Stockhammer, 'Dynamic Adaptive Streaming over HTTP: Standards and Design Principles', *Proceedings of the Second Annual ACM Conference on Multimedia Systems*, vol. 2014, no. i, pp. 133–144, 2011.

[53] M. Pantos, *Apple HLS (HTTP Live Streaming). IETF draft*, 2013.

[54] A. Zambelli, 'IIS Smooth Streaming Technical Overview', *Microsoft Corporation*, no. March, 2009.

[55] *Adobe HTTP Dynamic Streaming*, 2011.

[56] International Telecommunication Union, 'ITU-T Recommendation P.910: Subjective video quality assessment methods for multimedia applications', 2008.

[57] ——, 'ITU-T Recommendation P.911: Subjective audiovisual quality assessment methods for multimedia applications', 1998.

[58] ——, 'ITU-T Recommendation P.913: Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment', 2016.

[59] P. Le Callet, S. Möller, A. Perkis *et al.*, 'Qualinet white paper on definitions of quality of experience', *European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)*, vol. 3, 2012.

[60] R. L. Oliver, 'A conceptual model of service quality and service satisfaction: compatible goals, different concepts', *Advances in services marketing and management*, vol. 2, no. 4, pp. 65–85, 1993.

[61] Q. Huynh-Thu and M. Ghanbari, 'Scope of validity of PSNR in image/video quality assessment', *Electronics letters*, vol. 44, no. 13, pp. 800–801, 2008.

[62] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, 'Image quality assessment: from error visibility to structural similarity', *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[63] D. M. Chandler and S. S. Hemami, 'VSNR: A wavelet-based visual signal-to-noise ratio for natural images', *IEEE transactions on image processing*, vol. 16, no. 9, pp. 2284–2298, 2007.

[64] M. H. Pinson and S. Wolf, 'A new standardized method for objectively measuring video quality', *IEEE Transactions on broadcasting*, vol. 50, no. 3, pp. 312–322, 2004.

[65] H. R. Sheikh and A. C. Bovik, 'Image information and visual quality', *IEEE Transactions on image processing*, vol. 15, no. 2, pp. 430–444, 2006.

[66] A. Beben, P. Wiśniewski, J. M. Batalla and P. Krawiec, 'ABMA+ : lightweight and efficient algorithm for HTTP adaptive streaming', in *Proceedings of the 7th International Conference on Multimedia Systems - MMSys '16*, New York, New York, USA: ACM Press, 2016, pp. 1–11.

[67] K. Spiteri, R. Urgaonkar and R. K. Sitaraman, 'BOLA: Near-optimal bitrate adaptation for online videos', in *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, IEEE, Apr. 2016, pp. 1–9. arXiv: `1601.06748`.

[68] D. J. Vergados, A. Michalas, A. Sgora, D. D. Vergados and P. Chatzimisios, 'FDASH: A Fuzzy-Based MPEG/DASH Adaptation Algorithm', *IEEE Systems Journal*, vol. 10, no. 2, pp. 859–868, Jun. 2016.

[69] X. Yin, A. Jindal, V. Sekar and B. Sinopoli, 'A Control-Theoretic Approach for Dynamic Adaptive Video Streaming over HTTP', in *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication - SIGCOMM '15*, ACM Press, 2015, pp. 325–338.

[70] J. Jiang, V. Sekar and H. Zhang, 'Improving Fairness, Efficiency, and Stability in HTTP-based Adaptive Video Streaming with FESTIVE', in *Proceedings of the 8th international conference on Emerging networking experiments and technologies*, 2012, pp. 97–108.

[71] Z. Li, X. Zhu, J. Gahm, R. Pan, H. Hu, A. C. Begen and D. Oran, 'Probe and Adapt: Rate Adaptation for HTTP Video Streaming At Scale', *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 4, pp. 719–733, Apr. 2014. arXiv: `1305.0510v2`.

[72] L. De Cicco, V. Caldaralo, V. Palmisano and S. Mascolo, 'ELASTIC: A Client-Side Controller for Dynamic Adaptive Streaming over HTTP (DASH)', in *2013 20th International Packet Video Workshop*, IEEE, Dec. 2013, pp. 1–8.

[73] P. Juluri, V. Tamarapalli and D. Medhi, 'SARA: Segment aware rate adaptation algorithm for dynamic adaptive streaming over HTTP', in *2015 IEEE International Conference on Communication Workshop (ICCW)*, IEEE, Jun. 2015, pp. 1765–1770.

[74] R. K. P. Mok, X. Luo, E. W. W. Chan and R. K. C. Chang, 'QDASH: A QoE-aware DASH system', in *Proceedings of the 3rd Multimedia Systems Conference*, New York, New York, USA: ACM Press, 2012, p. 11.

[75] R. Fang, R. Al-Bayaty and D. Wu, 'BNB Method for No-Reference Image Quality Assessment', *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 7, pp. 1381–1391, 2017.

[76] D. Zegarra Rodriguez, R. Lopes Rosa, E. Costa Alfaia, J. Issy Abrahao and G. Bressan, 'Video Quality Metric for Streaming Service Using DASH Standard', *IEEE Transactions on Broadcasting*, vol. 62, no. 3, pp. 628–639, Sep. 2016.

[77] Y. Xue, B. Erkin and Y. Wang, 'A Novel No-Reference Video Quality Metric for Evaluating Temporal Jerkiness due to Frame Freezing', *IEEE Transactions on Multimedia*, vol. 17, no. 1, pp. 134–139, Jan. 2015.

[78] R. Trestian, O. Ormond and G.-M. Muntean, 'Performance evaluation of MADM-based methods for network selection in a multimedia wireless environment', *Wireless Networks*, vol. 21, no. 5, pp. 1745–1763, Jul. 2015.

[79] T. Zhao, Q. Liu and C. W. Chen, 'QoE in Video Transmission: A User Experience-Driven Strategy', *IEEE Communications Surveys & Tutorials*, no. c, pp. 1–1, 2016.

[80] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hobfeld and P. Tran-Gia, 'A Survey on Quality of Experience of HTTP Adaptive Streaming', *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 469–492, 2015.

[81] D. Mcmullin, R. Trestian and G.-m. Muntean, 'Power Save-based Adaptive Multimedia Delivery Mechanism', *The 9th Information Technology and Telecomunication Conference (IT&T)*, 2009.

[82] M. Kennedy, H. Venkataraman and G.-M. Muntean, 'Battery and stream-aware adaptive multimedia delivery for wireless devices', *Proceedings - Conference on Local Computer Networks, LCN*, pp. 843–846, 2010.

[83] S. A. Jalal, N. Gibbins, D. Millard, B. Al-Hashimi and N. R. Aljohani, 'Energy-Aware Adaptation of Educational Multimedia in Mobile Learning', *Proceedings of International Conference on Advances in Mobile Computing & Multimedia - MoMM '13*, pp. 434–439, 2013.

[84] L. Zou, R. Trestian and G.-M. Muntean, 'eDOAS: Energy-aware device-oriented adaptive multimedia scheme for Wi-Fi offload', in *IEEE Wireless Communications and Networking Conference (WCNC)*, vol. 3, IEEE, Apr. 2014, pp. 2916–2921.

[85] ——, 'E 2 DOAS : User Experience Meets Energy Saving for Multi-Device Adaptive Video Delivery', in *Computer Communications Workshops (INFOCOM WKSHPS), 2015 IEEE Conference on*, 2015, pp. 493–498.

[86] L. Zou, R. Trestian and G.-m. Muntean, 'E 3 DOAS : Balancing QoE and Energy-Saving for Multi-Device Adaptation in Future Mobile Wireless Video Delivery', pp. 1–15, 2017.

[87] R. Trestian, Q.-T. Vien, H. X. Nguyen and O. Gemikonakli, 'On the Impact of Video Content Type on the Mobile Video Quality Assessment and Energy Consumption', *International Symposium on Broadband Multimedia Systems and Broadcasting*, pp. 1–6, Jun. 2015.

[88] M. A. Hoque, M. Siekkinen and J. K. Nurminen, 'Energy Efficient Multimedia Streaming to Mobile Devices - A Survey', *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 579–597, 2014.

[89] L. Rovcanin and G.-M. Muntean, 'A DASH-based performance-oriented Adaptive Video distribution solution', *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, BMSB*, pp. 1–7, Jun. 2013.

[90] S. Lederer and. M{\"u}ller, Christopher Timmerer, 'Towards peer-assisted dynamic adaptive streaming over HTTP', in *International Packet Video Workshop (PV), 2012 19th International*, IEEE, May 2012, pp. 161–166.

[91] G. Tian, Yang Xu, Y. Liu and K. Ross, 'Mechanism design for dynamic P2P streaming', in *IEEE P2P 2013 Proceedings*, IEEE, Sep. 2013, pp. 1–10.

[92] L. Natali and M. L. Merani, 'A Novel Rate Control Strategy for Adaptive Video Streaming in P2P Overlays', in *2015 IEEE Global Communications Conference (GLOBECOM)*, IEEE, Dec. 2014, pp. 1–7.

[93] ——, 'Successfully Mapping DASH over a P2P Live Streaming Architecture', *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 6, pp. 1326–1339, 2017.

[94] X. Zhang, R. Yu, Y. Zhang, Y. Gao, M. Im, L. Cuthbert and W. Wang, 'Energy-efficient multimedia transmissions through base station cooperation over heterogeneous cellular networks exploiting user behavior', *IEEE Wireless Communications*, vol. 21, no. 4, pp. 54–61, 2014.

[95] H. Abou-Zeid and H. S. Hassenein, 'Toward Green Media Delivery: Location-Aware Opportunities and Approaches', *IEEE Wireless Communications*, no. August, pp. 38–46, 2014.

[96] J. Wu, C. Yuen, B. Cheng, M. Wang and J. Chen, 'Energy-minimized multipath video transport to mobile devices in heterogeneous wireless networks', *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 5, pp. 1160–1178, 2016.

[97] M. A. Hoque, M. Siekkinen, J. K. Nurminen, S. Tarkoma and M. Aalto, 'Saving Energy in Mobile Devices for On-Demand Multimedia Streaming – A Cross-Layer Approach', *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 10, no. 3, pp. 1–23, Apr. 2014.

[98] R. Trestian, O. Ormond and G.-M. Muntean, 'Enhanced Power-Friendly Access Network Selection Strategy for Multimedia Delivery Over Heterogeneous Wireless Networks', *IEEE Transactions on Broadcasting*, vol. 60, no. 1, pp. 85–101, Mar. 2014.

[99] C. Singhal, S. De, R. Trestian, G.-M. Muntean and S. Member, 'Joint Optimization of User-Experience and Energy-Efficiency in Wireless Multimedia Broadcast', *IEEE Transactions on Mobile Computing*, vol. 13, no. 7, pp. 1522–1535, Jul. 2014.

[100] M. Anedda, M. Murroni and G.-M. Muntean, 'E-ARMANS: Energy-aware device-oriented video delivery in heterogeneous wireless networks', in *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, IEEE, Jun. 2017, pp. 1–6.

[101] L. Al-Kanj, Z. Dawy and E. Yaacoub, 'Energy-Aware Cooperative Content Distribution over Wireless Networks: Design Alternatives and Implementation Aspects', *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1736–1760, 2013.

[102] J. Hecht, 'The bandwidth bottleneck that is throttling the Internet', *Nature*, vol. 536, no. 7615, pp. 139–142, Aug. 2016.

[103] A. Asadi and V. Mancuso, 'WiFi Direct and LTE D2D in action', in *2013 IFIP Wireless Days (WD)*, IEEE, Nov. 2013, pp. 1–8.

[104] K. Mahmud, M. Inoue, H. Murakami, M. Hasegawa and H. Morikawa, 'Measurement and usage of power consumption parameters of wireless interfaces in energy-aware multi-service mobile terminals', *Personal, Indoor and Mobile Radio Communications, 2004. PIMRC 2004. 15th IEEE International Symposium on*, vol. 2, pp. 1090–1094, 2004.

[105] J. Monks and G.-M. Muntean, 'A Distributed Energy-Aware Cooperative Multimedia Delivery Solution', in *2017 IEEE 42nd Conference on Local Computer Networks Workshops (LCN Workshops)*, IEEE, Oct. 2017, pp. 66–74.

[106] S. B. Lee, G.-M. Muntean and A. F. Smeaton, 'Performance-Aware Replication of Distributed Pre-Recorded IPTV Content', *IEEE Transactions on Broadcasting*, vol. 55, no. 2, pp. 516–526, 2009.