Università degli Studi di Cagliari

DOTTORATO DI RICERCA
Ingegneria Elettronica e Informatica
Ciclo XXIX

QoS OPTIMIZATION FOR MULTIMEDIA DELIVERY CONTENT OVER HETEROGENEOUS WIRELESS NETWORKS

Settore/i scientifico disciplinari di afferenza
ING-INF-03 (Telecomunicazioni)

Presentata da:	MATTEO ANEDDA
Coordinatore Dottorato	Prof. FABIO ROLI
Tutor	Prof. MAURIZIO MURRONI

Esame finale anno accademico 2015 – 2016
Tesi discussa nella sessione d'esame marzo – aprile 2017

TABLE OF CONTENT

Table of Content i

Introduction iv

Related Papers vii

List of Tables viii

List of Figures ix

List of Abbreviations xi

Part I - Heuristic Optimization 1

1 Digital Video Broadcasting 2

1.1 DVB-T2 2

1.1.1 DVB-T2 History 2

1.1.2 The DVB-T2 specification 3

1.1.3 Tests 4

1.1.4 The standard 4

1.1.5 System differences with DVB-T 5

1.2 DVB-SH 5

2 Heuristic Algorithms: Definition and Classification 7

2.1 Simulated Annealing 7

2.2 Advanced simulated annealing 8

2.3 Genetic Algorithm 9

3 Coverage Optimization in DVB-T/T2 11

3.1 Introduction 11

3.2 Planning SFN networks 12

3.2.1 Problem description 12

3.2.2 COFDM symbol and static delay overview 13

3.3 SFN network simulation model 13

3.4 Heuristic approach 16

3.4.1 SFN network heuristic optimization 16

3.5 DVB-T2 Scenarios 18

3.5.1 Basque Country study case 18

3.5.2 Sardinia study case 19

3.6 ITU-R Path Loss comparison 20

3.6.1 The GA based approach 21
3.6.2 Region of Interest: Basque Country, Sardinia Island, Flanders Region 22
3.7 DVB-T2 simulation results 26
3.8 Conclusions 29
4 Coverage Optimization of Satellite Broadcasting 31
4.1 Introduction 31
4.2 System Model 33
4.3 Genetic Algorithms Optimization 36
4.4 Mapping analysis 37
4.4.1 CCSDS Mapping for the 4+12+20+28 constellation 37
4.4.2 Proposed Mapping for the 4+12+20+28 constellation 38
4.4.3 Proposed Mapping for the 4+12+16+32 constellation 39
4.5 Optimization Results 41
4.5.1 Results without constellation optimization 41
4.5.2 Results with constellation optimization 41
4.5.3 Validation in realistic scenario 43
4.6 Conclusions 45
Part II - Heterogeneous Wireless Networks Optimization 47
5 Adaptive Real-time Multi-user Access Network Selection Algorithm for Load-balancing over Heterogeneous Wireless Networks 48
5.1 Introduction 48
5.2 Related Works 51
5.2.1 Load balancing state of the art 51
5.2.2 Prioritization and radio access network selection 51
5.2.3 Markov Decision Process in Access Network Selection 52
5.3 ARMANS and P-ARMANS System Models 54
5.3.1 Framework of the Proposed ARMANS 54
5.3.2 Framework of the Proposed P-ARMANS 54
5.4 Network Load Model 56
5.4.1 Problem Formulation 57
5.4.2 ARMANS and P-ARMANS Proposed Load Balancing Scheme 58
5.5 ARMANS Simulation Plan 61
5.6 P-ARMANS MDP-Based Priority Model 62
5.6.1 Markov Decision Process 62
5.6.2 MDP States 64

5.6.3 MDP Actions 65
5.6.4 User Priority Model 66
5.6.5 MDP States Transition Probability 66
5.6.6 MDP Algorithm and Reward Function 68
5.7 ARMANS Simulation-based Testing 70
5.8 P-ARMANS Simulation-based Testing 71
5.8.1 Multimedia Content 71
5.8.2 Simulation Scenario 72
5.9 Simulation-Based Testing Results 75
5.9.1 Packet Delay Results 76
5.9.2 Average PSNR Delay Results 77
5.9.3 Packet Loss Rate Results 78
5.9.4 Aggregate Throughput Results 80
5.9.5 P-ARMANS MDP and ARMANS results 78
5.10 Conclusion 82
References 84

INTRODUCTION

The transition from analogue TV to digital TV has led a fast-growing of available multimedia contents. In recent decades there has been a great development of digital broadcasting systems, in addition to those relating to the existing DVB consortium other examples just as representative as DAB (Digital Audio Broadcasting) and DRM (Digital Radio Mondiale). Nowadays, we can mention three important aspects for success in services development that businesses need, and for which has not been provided a satisfying solution.

The **first** factor is that single frequency networks (SFN) are a essential tool to enable efficient use of resources spectrum available for spectrum bands allocated for broadcasting services spread. Radio spectrum is a limited resource and is a growing problem especially in the UHF band as discussed in Geneve treaty in 2006. Moreover, a part of UHF band used by TV broadcasting has been assigned to mobile phone services, thus limiting in that sense the available spectrum for digital television services in the future.

The **second** of these factors is based on digital services complexity. The most of the analog broadcast networks have been successfully designed manually, but the complexity introduced by new services digital broadcasting, it has become evident in recent years.

Specifically with implementation of single frequency networks, it becomes the task of planning networks manually using the traditional model, in a risky task, which involves subsequent amendments and adjustments that in turn give rise to an elevation of costs and reduced quality of service.

To date, the service planning has been done with planning tools that have been, in many cases, legacy analog systems and in other cases tools are adapted from analog to include some the specific characteristics of digital networks. These adaptations have failed to achieve the best planning tool that allows important parameters include the new networks. This is especially significant because the mechanisms sync single frequency networks (SFN).

The **third** factor is related to digital receivers and COFDM modulation. Although the use of COFDM multicarrier modulation has been widely accepted in digital broadcasting systems there are important shortcomings regarding the characterization of receptors and in the knowledge on its actual performance.

Evaluation tests of different systems show that these shortcomings are especially evident when these receptors operate within a single frequency network, and this naturally affects the planning process of these network configurations.

Given these factors, it is possible consider that the business for networks digital broadcasting requires a tool to exploit the features of digital broadcasting systems, in particular the advantages provided in the use of single frequency networks (SFN). It is a clear need providing solutions to broadcasters to enable them to address the task of design and implementation of new networks in a reliable and the most optimal way possible. Thus

arises the need to provide a coverage optimization algorithm, applicable in the planning of digital broadcasting systems with modulation multicarrier COFDM, in his entry for a single frequency network.

Part I of this thesis proposes to plan and model the functionality of broadcasting systems: in standardization phase, in testing phase and finally in commercial deployment phase. The aim is to further analyze the influences that have different parameters of single frequency networks (SFN) in the receivers.

This will have acquired professional receivers of different digital services in both fixed and mobile services. The aim is then to use such theoric models together with project execution and compare the result of these models is more faithful to reality than the "traditional" model used up today, comparing simulations with measurement campaigns carried out.

The purpose of the first part of this project therefore is as follows: the study and implementation of an algorithm for optimizing the planning of digital broadcasting networks based on COFDM modulation for use in single frequency networks (SFN). This algorithm manipulates the information on a broadcast network and estimates degree of coverage and interference that offers a particular network, and does it the best way possible and at minimum cost. Subsequently the study and implementation of this algorithm optimization, we have studied in terms of functionality and performance, to evaluate the usefulness of them and if conclusions drawn from their application can be applied reliably.

On the other hand, the mobile data traffic reached two major milestones and continues to increase and diversify. First milestone refers to the offload traffic which exceeded the cellular traffic for the first time in 2015. Globally, the total public number of Wi-Fi hotspots is expected to grow seven fold from 64.2 million in 2015 to 432.5 million by 2020. Originally, due to the high availability of Wi-Fi hotspots globally, and continued pressure on users imposed by data caps, the mobile users have opted for Wi-Fi wherever possible. Lately, technology advances including dual mode devices have enabled smarter switching between Wi-Fi and cellular networks. The second milestone reached was that the 4G traffic surpassed the 3G traffic. Although 4G connections represented only 14 percent of mobile connections in 2015, they have accounted for 47 percent of mobile data traffic, while 3G connections represented 34 percent of mobile connections, but carried 43 percent of the traffic. This trend continues today and is expected to include 5G solutions whenever they will become available.

A major contributor to this traffic growth is the enhancement process of mobile devices which resulted in user access to smartphones and tablets with high processing power, diverse screen sizes and very ever-greater high display resolutions. Moreover, video-on-demand services such as Vimeo, YouTube, Netflix and Amazon Instant Video offer diverse high data rate content which puts pressure in terms of resources on both network for delivery and device for processing display. In particular, mobile device innovations are being introduced at unprecedented speeds and have improved exponentially with time. The last years have seen a 20 times increase in cellular speed, a 30 times increase in Wi-Fi bandwidth and a 34 times increase in video resolution (compared to 2010).

Part II of this thesis proposes two novel algorithms: the **adaptive real-time multi user access network selection (ARMANS)** and **Prioritized ARMANS (P-ARMANS).** ARMANS is a load balancing algorithm that takes into account not only the real-time global traffic load on each network, but also considering the different classes of traffic, voice, video, best-effort, and background. The proposed algorithm improves both QoS and load balancing in comparison with the case when a classic network selection with no traffic type load balancing is employed.
The novel P-ARMANS load balancing algorithm Markov decision process (MDP)-based multi-agent framework was introduced to perform the bandwidth allocation on users with different typology and priority and also considering and classifying the type of traffic involved in a multimedia data user demand. The proposed approach performs the available bandwidth on users with different screen resolution and priority. We demonstrate that P-ARMANS framework ensures high QoS distributing different type of traffic load among typical and business users compared to a classic no-priority model.

Both ARMANS and P-ARMANS operate over heterogeneous wireless networks (HWNs), 4G-LTE and 802.11n.

The rest of the thesis is structured as follows.

The Part I defines the heuristic optimization both for Digital Video Broadcasting (DVB) Terrestrial (DVB-T/T2) and Satellite Broadcasting (DVB-S/S2/SH), whereas the Part II proposes two algorithms for real-time access network selection in heterogeneous networks.

In Chapter 1, the DVB-T2 and DVB-SH are briefly detailed. This chapter defines the considered network to be optimized in terms of Quality of Service (QoS) and coverage.

In Chapter 2, the heuristic algorithms employed in this work are briefly defined and classified. Simulated annealing (SA) and Genetic Algorithms (GA) were used in order to improve and optimized different study cases on DVB.

In Chapter 3, two propagation prediction tools are proposed in three different scenarios: Basque Country (northern Spain), Sardinia Island (Italy), and Flanders (Belgium). The performance of the algorithm will be evaluated by means of simulation results

In Chapter 4, the GA optimizes the 64-APSK constellation in order to reduce the distortion while preserving bandwidth.

The Part II treats the heterogeneous wireless network optimization with particular attention to the overloaded networks.

In Chapter 5 the ARMANS and P-ARMANS algorithms were proposed to ensure high QoS for different type of traffic load, type of user in Wi-Fi and LTE scenarios.

RELATED PAPERS

M. Anedda, P. Angueira, J. Morgade, and M. Murroni, "Heuristic performance evaluation for DVB-T/T2 SFN network," *Telecommun Syst*, vol. 63, no. 1, pp. 55–63, Mar. 2015.

M. Anedda, A. Anedda, and M. Murroni, "Coverage optimization for DVB-T2 SFNs using ITU-R P.1546 and ITU-R P.1812," in *2015 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Ghent, 2015, pp. 1-5. doi: 10.1109/BMSB.2015.7177188*, 2015, pp. 1–5.

M. Anedda, A. Meloni, and M. Murroni, "64-APSK Constellation and Mapping Optimization for Satellite Broadcasting Using Genetic Algorithms," *IEEE Transactions on Broadcasting*, vol. 62 (1), 7271031, no. 99, pp. 1–9, 2016.

M. Anedda, G.-M. Muntean, and M. Murroni, "Adaptive Real-time Multi-user Access Network Selection Algorithm for Load-balancing over Heterogeneous Wireless Networks," in *2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), Nara, 2016, pp. 1-4. doi: 10.1109/BMSB.2016.7521907*, 2016.

In press:

M. Anedda, G.-M. Muntean, and M. Murroni, "P-ARMANS: Priority Users Access Network Selection Algorithm for Load-balancing using the Markov Decision Process", *IEEE Transactions on Broadcasting*

M. Anedda, A. Anedda, and M. Murroni, "Coverage Optimization for DVB-T2 SFN: a Comparative Analysis of PSO, A-SA, and A-GA", *International Journal of Digital Multimedia Broadcasting, Hindawi*

LIST OF TABLES

Table I - Differences between DVB-T and DVB-T2 5
Table II - Threshold values for DVB-T2 18
Table III - Genetic Algorithm Parameters 22
Table IV - Threshold Values for DVB-T2 23
Table V - Static delay optimization in Basque Country 27
Table VI - Comparison of the SA and ASA optimization schemes under different conditions of reception and transmission in Basque Country 27
Table VII - Comparison of the SA and ASA optimization schemes in Sardinia 27
Table VIII - Static delays optimization for cog in Sardinia 27
Table IX - Comparison of the SA and ASA optimization schemes under different conditions of reception and transmission in Sardinia 28
Table X - Static Delay optimization in Sardinia Region 29
Table XI - Static Delay optimization in Basque Country Region 29
Table XII - Static Delay optimization in Flanders Region 29
Table XIII - Comparison of the Average Distortion among Neighbour Symbols on Different Rings 38
Table XIV - comparison of the average distortion among neighbor symbols on the same ring 38
Table XV - Comparison of the PEAQ Using Non-Linear and Ideal HPA Filter 42
Table XVI - Type of Traffic Characteristics 62
Table XVII - Encoding video data rate 62
Table XVIII - Simulation Parameters 69
Table XIX - Simulation Results of Scenario 1 and Scenario 2 70
Table XX - Encoding Video Data Rate 72
Table XXI - Simulation Parameters 73
Table XXII - Type of Users and Traffic Characteristics 74
Table XXIII - P-ARMANS MDP Actions Sequence and Data Rate 75
Table XXIV - Simulation Results Evaluation 80

List of Figures

Figure 1 - DVB-T2 logo. 3
Figure 2 - DVB-SH supported spectrum 6
Figure 3 - Example of network simulation model: a yellow dot generic receiver contains information represented by three couples of power received and relative delay by three TXs 14
Figure 4 – DVB-T MODEL – splitting of the signal power into useful and interfering components. 15
Figure 5 - Block diagram of the approach presented. 17
Figure 6 - Details of the Basque Country SFN with T=18 transmitters in which the analysis has been carried out. 19
Figure 7 - Sardinia study case, 78 transmitters SNF network 19
Figure 8 - Optimized SFN Sardinia network (Scen2): gain sectors and relative orientation. The white, blue and red sectors represent the gains in the ranges [-0.99, 0], [-2.99, -1] and [-5, -3] dB, respectively. 20
Figure 9 – Block diagram of the SFN planning tool using ITU-R P.1546 and P.1812 21
Figure 10 – basque country orography of territory 23
Figure 11 – Flanders orography of territory 24
Figure 12 – Sardinia orography of territory 25
Figure 13 – Flanders SFN 16 transmitters 26
Figure 14 - System model 34
Figure 15 - 16-APSK constellation (left) and 32-APSK constellation 34
Figure 16 - Distortion of the CCSDS mapping for the 4+12+20+28 (a), the 4+12+20+28 (b) and the 4+12+16+32 composed 64-APSK constellation (c). Symbols and relative mutual distortion are expressed in decimal representation. 39
Figure 17 - MSE results for the CCSDS mapping and the 4+12+20+28 (a), 4+12+16+32 (b) and overall comparison of the presented constellations without GA optimization, (c). 40
Figure 18 - Optimized constellation for the 4+12+20+28 CCSDS mapping (a), 4+12+20+28 proposed mapping (b), and 4+12+16+32 proposed mapping (c). Symbols are expressed in decimal representation 43
Figure 19 - Performance comparison between optimized and non-optimized CCSDS mapping (a), performance comparison between optimized and non-optimized 4+12+20+28 proposed mappings (b), MSE results obtained optimizing the CCSDS and the 4+12+16+32 proposed mappings (c), and overall comparison of the presented constellation with GA optimization (d) 45
Figure 20 - Original “Lena” 512x512, 8bpp (a), “Lena” transmitted for SNR = 5dB, with the proposed modified 64-APSK 4+12+20+28 mapping (b), and conventional 64-APSK with CCSDS mapping (c). 45
Figure 21 - Load balancing ARMANS server in heterogeneous network 49
Figure 22 - Load balancing P-ARMANS server in heterogeneous network 50
Figure 23 - ARMANS scheduling framework 55
Figure 24 - P-ARMANS scheduling framework 56
Figure 25. The policy π of the MDP algorithm 64
Figure 26 - Aggregate throughput for different load balancing schemes, with increasing amount of offered load (a), average delay for different load balancing schemes, with increasing the amount of offered load (b), and average packet loss rate for ARMANS and classic load balancing (c). 70
Figure 27 - PSNR achieved for a 1080p resolution using UDP traffic 71

Figure 28 - MDP Process **75**
Figure 29 - Scenario 1 (a), 2 (b) and 3 (c) **77**
Figure 30 - Packet delay for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP **78**
Figure 31 - Average PSNR for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP **79**
Figure 32 - Packet loss rate for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP **79**
Figure 33 - Aggregate throughput for 4K, 2K and FHDi TUs and BUs evaluated with P-ARMANS MDP **79**
Figure 34 - Packet loss rate for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP and ARMANS **80**

LIST OF ABBREVIATIONS

AM	**Amplitude Modulation**
AP	**Access Point**
APSK	**Amplitude Phase Shift Keying**
ARMANS	**Adaptive Real-time Multi-user Access Network Selection**
ASA	**Advanced Simulated Annealing**
AVC	**Advanced Video Coding**
AWGN	**Additive White Gaussian Noise**
BER	**Bit Error Rate**
BU	**Business User**
CCSDS	**Consultative Committee for Space Data Systems**
CD	**Compact Disk**
COFDM	**Coded Orthogonal Frequency-Division Modulation**
CQI	**Channel Quality Indicator**
DAB	**Digital Audio Broadcasting**
DTG	**Digital TV Group**
DTT	**Digital Terrestrial Television**
DVB	**Digital Video Broadcasting**
DVB-H	**Digital Video Broadcasting - Handled**
DVB-S2	**Digital Video Broadcasting – Satellite version 2**
DVB-SH	**Satellite services to Handheld**
DVB-T	**Digital Video Broadcasting – Terrestrial**
DVB-T2	**Digital Video Broadcasting – Terrestrial version 2**
EIRP	**Equivalent Isotropic Radiated Power**
eNodeB	**evolved Node B**
ETSI	**European Telecommunications Standards Institute**
FEC	**Forward Error Correction**
FFT	**Fast Fourier Transform**
FHDi	**Full High Definition**
GA	**Genetic Algorithm**
GI	**Guard Interval**
HD	**High-Definition**
HDTV	**High-Definition Television**
HPA	**High Power Amplifier**
HQ	**High Quality**
HWNs	**Heterogeneous Wireless Networks**
iPAS	**intelligent Prioritized Adaptive Scheme**
IPDC	**DVB IP Datacast**
ISI	**Inter-Symbol Interference**
ITU-R	**International Telecommunication Union – Radiocommunication Sector**
LTE	**Long Term Evolution**
MADM	**Multiple Attribute Decision-making**
MC	**Markov chain**
MDP	**Markov Decision Process**
MISO	**Multiple Input Single Output**
ML	**Maximum Likelihood**
MNO	**Mobile Network Operator**
MP	**Markov process**

MPEG	**Motion Picture Expert Group**
MPI	**Multipath Interference**
MP-TCP	**Multipath TCP**
MSB	**Most Significant Bit**
MSE	**Mean Square Error**
MUPA	**Modulation with Unequal Power Allocation**
OFDM	**Orthogonal Frequency-division Multiplexing**
OFDMA	**Orthogonal Frequency-division Multiple Access**
OP	**Optimization Problem**
P-ARMANS	**Prioritized-ARMANS**
PEAQ	**Perceptual Evaluation of Audio Quality**
PM	**Phase Modulation**
PSI	**Program and System Information**
PSNR	**Peak Signal to Noise ratio**
PSNR	**Peak Signal to Noise Ratio**
PSO	**Particle Swarm Optimization**
QAM	**Quadrature Amplitude Modulation**
QCI	**QoS Class Identifier**
QoE	**Quality of Experience**
QoS	**Quality of Service**
RAT	**Radio Access Technologies**
S2P	**Serial-to-parallel**
SA	**Simulated Annealing**
SD	**Standard Definition**
SFN	**Single Frequency Network**
S-GRA	**Strengthen-Gray Relative Analysis**
SMDP	**Semi-Markov Decision Process**
SQ	**Standard Quality**
SRRC	**Squared Root Raised Cosine**
TCP	**Transmission Control Protocol**
TM-T2	**Technical Module on Next Generation DVB-T**
TTI	**Transmission Time Interval**
TU	**Typical User**
UDP	**User Datagram protocol**
UEP	**Unequal Error Protection**
UHD	**Ultra High Definition**
UHF	**Ultra high Frequency**
UMTS	**Universal Mobile Telecommunication System**
UPR	**User Perceived Rate**
VHF	**Vey High Frequency**
Wi-Fi	**Wireless Fidelity**
WLAN	**Wireless Local Area Network**

PART I - HEURISTIC OPTIMIZATION

1 DIGITAL VIDEO BROADCASTING

The DVB Project is an Alliance of about 250-300 companies, originally of European origin but now worldwide. Its objective is to agree specifications for digital media delivery systems, including broadcasting. It is an open, private sector initiative with an annual membership fee, governed by a Memorandum of Understanding (MoU). The DVB Project has used and continues to draw extensively on standards from the ISO/IEC JTC MPEG. The transport for all systems is the MPEG2 transport stream. The DVB report-giving candidates for baseband systems follows the systems developed in the JTC MPEG. For convenience, the final documentation is arranged with sets of hyphenated initials that identify the area. For example, DVB-S2 is the specification for the second generation version of the DVB digital satellite system. Other areas include DVB-S (the first generation version of the digital satellite system), DVB-C (the digital cable system), DVB-T/T2 (the digital terrestrial broadcasting system, first and second generation), DVB-H (bringing digital terrestrial broadcasting to battery-powered handheld receivers), DVB-DATA (the cyclical data delivery system), DVB-SI (the service information system), and DVB-MHP (middleware for interactive television).

1.1 DVB-T2.

DVB-T2 is an abbreviation of **Digital Video Broadcasting – Terrestrial second generation**; it is the extension of the television standard DVB-T, issued by the consortium DVB, devised for the broadcast transmission of digital terrestrial television. This system transmits compressed digital audio, video, and other data in "physical layer pipes" (PLPs), using OFDM modulation with concatenated channel coding and interleaving. It is currently broadcasting in parts of the UK under the brand name Free view HD.

1.1.1 DVB-T2 HISTORY

In March 2006 DVB decided to study options for an upgraded DVB-T standard. In June 2006, a formal study group named TM-T2 (Technical Module on Next Generation DVB-T) was established by the DVB Group to develop an advanced modulation scheme that could be adopted by a second generation digital terrestrial television standard, to be named DVB-T2.

According to the commercial requirements and call for technologies issued in April 2007, the first phase of DVB-T2 would be devoted to provide optimum reception for stationary (fixed) and portable receivers (i.e., units which can be nomadic, but not fully mobile) using existing aerials, whereas a second and third phase would study methods to deliver higher payloads (with new aerials) and the mobile reception issue. The novel system should provide a minimum 30% increase in payload, under similar channel conditions already used for DVB-T.

The BBC, ITV, Channel 4 and Five agreed with the regulator Federal Office of Communications in Switzerland (OFCOM) to convert one UK multiplex (B, or PSB3) to DVB-T2 to increase capacity for HDTV via DTT. They expected the first TV region to use the new standard would be Granada in November 2009 (with existing switched over regions being changed at the same time). It was expected that over time there would be enough DVB-T2 receivers sold to switch all DTT transmissions to DVB-T2, and H.264.

OFCOM published its final decision on April 3, 2008 for HDTV using DVB-T2 and H.264: BBC HD would have one HD slot after DSO at Granada. ITV and C4 had, as expected, applied to OFCOM for the 2 additional HD slots available from 2009 to 2012.

OFCOM indicated that it found an unused channel covering 3.7 million households in London, which could be used to broadcast the DVB-T2 HD multiplex from 2010, i.e., before DSO in London. OFCOM indicated that they would look for more unused UHF channels in other parts of the UK, that can be used for the DVB-T2 HD multiplex from 2010 until DSO.

Figure 1 - DVB-T2 logo.

1.1.2 THE DVB-T2 SPECIFICATION

DVB-T2 test modulator developed by BBC Research & Innovation. The DVB-T2 draft standard was ratified by the DVB Steering Board on June 26, 2008, and published on the DVB homepage as DVB-T2 standard BlueBook. It was handed over to the European Telecommunications Standards Institute (ETSI) by DVB.ORG on June 20, 2008. The ETSI process resulted in the DVB-T2 standard being adopted on September 9, 2009.The ETSI process had several phases, but the only changes were text clarifications. Since the DVB-T2 physical layer specification was complete, and there would be no further technical

enhancements, receiver VLSI chip design started with confidence in stability of specification. A draft PSI/SI (program and system information) specification document was also agreed with the DVB-TM-GBS group.

1.1.3 TESTS

Prototype receivers were shown in September IBC 2008 and more recent version at the IBC 2009 in Amsterdam. A number of other manufacturers demonstrated DVB-T2 at IBC 2009 including Albis Technologies, Arqiva, DekTec, Enensys, Harris, Pace, Rohde & Schwarz, Tandberg, and TeamCast. Other manufacturers planning DVB-T2 equipment launches include CellMetric, Cisco, Digital TV Labs, Humax, NXP Semiconductors, Panasonic, ProTelevision Technologies, Screen Service, SIDSA, Sony, ST Microelectronics and T-VIPS. The first test from a real TV transmitter was performed by the BBC Research & Innovation in the last weeks of June 2008 using channel 53 from the Guildford transmitter, southwest of London: BBC had developed and built the modulator/demodulator prototype in parallel with the DVB-T2 standard being drafted. NORDIG published a DVB-T2 receiver specification and performance requirement on the 1 July 2009. In March 2009 the Digital TV Group (DTG), the industry association for digital TV in the UK, published the technical specification for high definition services on digital terrestrial television (Freeview) using the new DVB-T2 standard. The DTG's test house: DTG Testing are testing Freeview HD products against this specification.

1.1.4 THE STANDARD

The following characteristics have been devised for the T2 standard:

- COFDM modulation with QPSK, 16-QAM, 64-QAM, or 256-QAM constellations.
- OFDM modes are 1k, 2k, 4k, 8k, 16k, and 32k. The symbol length for 32k mode is about 4 ms.
- Guard intervals are 1/128, 1/32, 1/16, 19/256, 1/8, 19/128, and 1/4. (For 32k mode, the maximum is 1/8.)
- FEC is concatenated LDPC and BCH codes (as in DVB-S2), with rates 1/2, 3/5, 2/3, 3/4, 4/5, and 5/6.
- There are fewer pilots, in 8 different pilot-patterns, and equalization can be based also on the RAI CD3 system.
- In the 32k mode, a larger part of the standard 8 MHz channel can be used, adding about 2% extra capacity.
- DVB-T2 is specified for 1.7, 5, 6, 7, 8, and 10 MHz channel bandwidth.

- MISO (Multiple-Input Single-Output) may be used (Alamouti scheme), but MIMO will not be used. Diversity receivers can be used (as they are with DVB-T).
- Bundling of more channels into a SuperMUX (called TFS) is not in the standard, but may be added later.

1.1.5 SYSTEM DIFFERENCES WITH DVB-T

The following Table I reports a comparison of available modes in DVB-T and DVB-T2.

	DVB-T	**DVB-T2**
FEC	Convolutional Coding + Reed Solomon 1/2, 2/3, 3/4, 5/6, 7/8	LDPC + BCH 1/2, **3/5**, 2/3, 3/4, **4/5**, 5/6
Modes	QPSK, 16QAM, 64QAM	QPSK, 16QAM, 64QAM, **256QAM**
Guard Interval	1/4, 1/8, 1/16, 1/32	1/4, **19/256**, 1/8, **19/128**, 1/16, 1/32, **1/128**
FFT size	2k, 8k	**1k**, 2k, **4k**, 8k, **16k**, **32k**
Scattered Pilots	8% of total	**1%**, **2%**, **4%**, 8% of total
Continual Pilots	2.6% of total	**0.35%** of total

Table I - Differences between DVB-T and DVB-T2

For instance, a UK MFN DVB-T profile (64-QAM, 2k mode, coding rate 2/3, guard interval 1/32) and a DVB-T2 equivalent (256-QAM, 32k, coding rate 3/5, guard interval 1/128) allows for an increase in bit rate from 24.13 Mbit/s to 35.4 Mbit/s (+46.5%).

Another example, for an Italian SFN DVB-T profile (64-QAM, 8k, coding rate 2/3, guard interval 1/4) and a DVB-T2 equivalent (256-QAM, 32k, coding rate 3/5, guard interval 1/16), achieves an increase in bit rate from 19.91 Mbit/s to 33.3 Mbit/s (+67%).

1.2 DVB-SH

The **DVB-SH** (**Satellite services to Handheld**) standard was published in February 2007. Trials are ongoing in several European countries.

DVB-SH, Digital Video Broadcasting - Satellite services to Handhelds, is a physical layer standard for delivering IP based media content and data to handheld terminals such as mobile phones or PDAs, based on a hybrid satellite/terrestrial downlink and for example a GPRS uplink. The DVB Project published the DVB-SH standard in February 2007.

The DVB-SH system was designed for frequencies below 3 GHz, supporting UHF band, L-Band or S-band. It complements and improves the existing DVB-H physical layer standard. Like its sister specification (DVB-H), it is based on DVB IP Datacast (IPDC)

delivery, electronic service guides and service purchase and protection standards.

DVB-SH specifies two operational modes:

- SH-A: specifies the use of COFDM modulation on both satellite and terrestrial links with the possibility of running both links in SFN mode.
- SH-B: uses Time-Division Multiplexing (TDM) on the satellite link and COFDM on the terrestrial link.
- FEC using Turbo coding
- Improved time interleaving
- Support for antenna diversity in terminals

The improvements to signal requirements translates to better in-building penetration, better in-car coverage and extension of outdoor coverage. DVB-SH chipsets are being developed now by DiBcom and NXP Semiconductors, and are expected to be available in beginning of 2008. Initial specifications show that the chipsets supports both UHF and S-Band and are compatible with DVB-H.

DiBcom has announced a DVB-SH chip with availability in 2008 Q3. Dibcom DVB-SH 2008 Q3. The chip "has dual RF tuners supporting VHF, UHF, L-Band and S-Band frequencies".

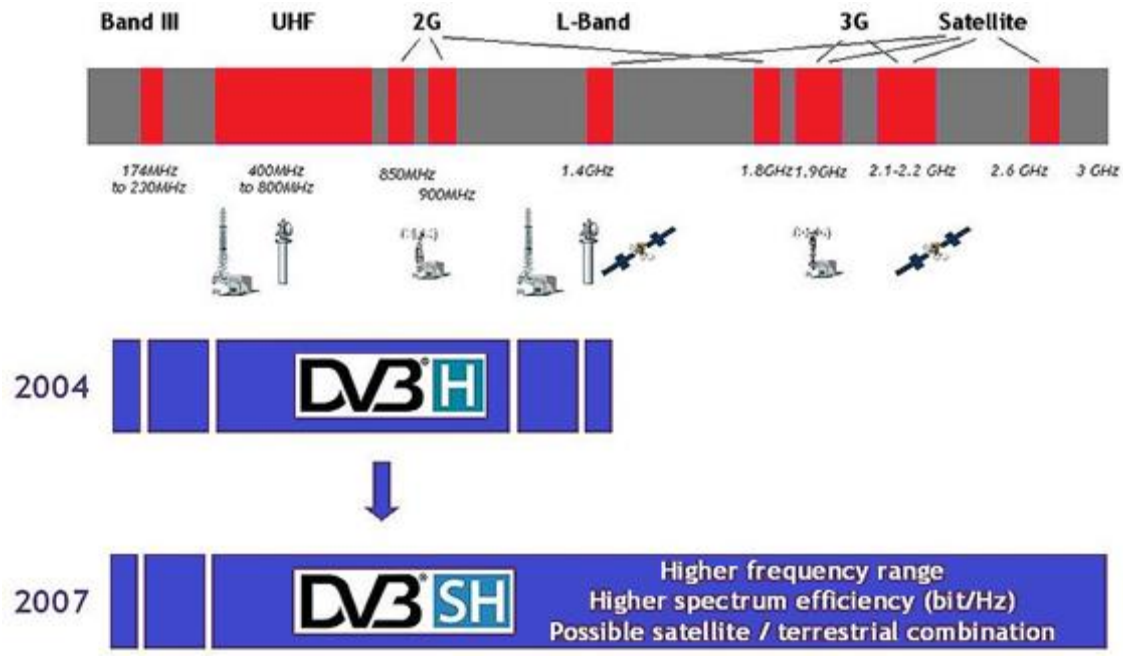


Figure 2 - DVB-SH supported spectrum

http://www.dvb.org/news_events/news/dvb_approves_dvbsh_specif/index.xml

Alcatel Unlimited mobile TV

2 HEURISTIC ALGORITHMS: DEFINITION AND CLASSIFICATION

The optimization methods can be classified into traditional methods and modern heuristic methods. Within the probabilistic heuristic methods, there are two differentiated families: those that use a single starting point and those that use a population. With regard to simulated annealing (SA), this method simulates at the computational level the physical process to be followed to obtain solids with minimum energy configurations. All these algorithms have in common the fact of using a population or set of potential solutions and subjecting them to an iterative process, using different schemes, operators and strategies depending on the type of algorithm. The most extensive family of this type of algorithms is the grouping of so-called evolutionary algorithms, establishing the four large areas that make up what is known in the literature as evolutionary computation. Of these four methodologies, the genetic algorithms deserve a special mention being one of the methods that more attention has received in the last years. Genetic algorithms (GA) have been used successfully in multiple fields of science, notably by way of example in electromagnetism, its application to the design of printed antennas and wire antennas, the synthesis of arrays or the design of microwave filters or the planning of distribution networks. Based on the principles of natural selection and evolution, the classical scheme of genetic algorithms evolves the population based on the influence exerted by the selection, crossing and mutation operators.

2.1 SIMULATED ANNEALING

The idea of SA comes from a paper published by Metropolis et al. in 1953 [1]. The algorithm implemented in this paper simulates the cooling of material in a heat bath. This is a process known as annealing, in metallurgy a technique involving heating and controlled cooling of a material to increase the size of its crystals and reduce their defects. The heat causes the atoms to become unstuck from their initial positions (a local minimum of the internal energy) and wander randomly through states of higher energy; the slow cooling gives them more chances of finding configurations with lower internal energy than the initial one.

In SA, the fitness function establishes the only relationship between the physical problem

and the SA algorithm, so the performance of the approach is directly related to the fitness function used. In fact, the algorithm works by representing the parameters of the fitness function as continuous numbers, and as dimensions of a hypercube (N dimensional space). Parameters of SA algorithm are delineated as follows [2]: T_k is a control variable called temperature gradually decreased along the process, k is the index of steps during which the temperature is fixed, f_k denotes the current function value, f_{k+1} is the new function value and f_{val} the final function value. The temperature is decreased at each step k according to $T_{k+1} = \alpha T_k$. The choice of α is crucial. Several simulations have been needed to verify the optimal value.

The detected value is α = 0.89. Compared to other values, α = 0.89 proves the best convergence performance in term of capability to avoid local minimum. Each step of the SA algorithm replaces the current solution by a random "nearby" solution, chosen with a probability (1)

$$P(k,k+1,T_k) = \begin{cases} 1 & if\ E(T_{k+1}) \le E(T_k),\ \ f_{k+1} \le f_k \\ \exp\left(-\dfrac{\Delta E}{k_B T_k}\right) & if\ E(T_{k+1}) > E(T_k),\ \ f_{k+1} > f_k \end{cases} \tag{1}$$

that depends both on the difference between the corresponding function values and also on the global parameter T_k; for a negative difference of energy (2)

$$\Delta E = E(T_k) - E(T_{k+1}) \tag{2}$$

the new solution is accepted with a probability defined by Boltzmann k_B distribution whereas for a positive difference of energy, the new solution is always accepted. The dependency is such that the current solution changes almost randomly when T_k is large, but increasingly "downhill" as T_k goes to zero. The allowance for "uphill" moves saves the method from becoming stuck at local optima which are the bane of greedier methods.

2.2 ADVANCED SIMULATED ANNEALING

Another configuration of SA, the advanced simulated annealing (ASA) [3] is also considered. In ASA the algorithm parameters that control temperature schedule and random step selection are automatically adjusted according to algorithm progress. This makes the algorithm more efficient and less sensitive to user-defined parameters than canonical SA. The temperature and the step size are adjusted so that all of the search space is sampled to a coarse resolution in the early stages, whilst the state is directed to favorable

areas in the late stages. For these reasons is not possible to consider the parameter α as a constant for each step as considered in general SA.

2.3 GENETIC ALGORITHM

The GA [4] is a method for solving both constrained and unconstrained optimization problems based on a natural selection process that mimics biological evolution. The algorithm repeatedly modifies a population of individual solutions. At each step, the GA randomly selects individuals from the current population and uses them as parents to produce the children for the next generation. Over successive generations, the population evolves toward an optimal solution. Similar to other stochastic algorithms like simulated annealing (SA) or particle swarm optimization (PSO), GA uses different operators to drive the global solution search. Selection operator is the stage in which individual genomes are chosen from a population for later breeding; Crossover operator is used to vary the programming of a chromosome or chromosomes from one generation to the next; Mutation operator is used to maintain genetic diversity from one generation of a population of GA chromosomes to the next.

The evolution usually starts from a Population of randomly generated individuals, and is an iterative process called Generation. In each generation, the fitness of every individual in the population is evaluated; the fitness is usually the value of the objective function in the optimization problem being solved.

Population and Generation influence GA convergence capacity towards best solution and computational complexity within optimization process. The computational load of the GA is measured substantially as the number of fitness assessments. The total number of iterations Neval played by the GA is calculated as follows:

$$N_{eval} = PopSize \cdot N_{gen} \tag{3}$$

where *PopSize* is the population size and N_{gen} represents the number of generations. It is evident that choosing a high number of generations N_{gen}, involves a high number of evaluations N_{eval} that beyond a certain limit, produce negligible improvements in terms of the value of the fitness function. GA uses different operators to drive the global solution search:

- *Selection* operator is the stage in which individual genomes are chosen from a population for later breeding.
- *Mutation* operator is used to maintain genetic diversity from one generation of a population of A-GA chromosomes to the next.

- *Crossover* operator is used to vary the programming of a chromosome or chromosomes from one generation to the next.

3 COVERAGE OPTIMIZATION IN DVB-T/T2

In DVB-T/T2 broadcasting, single frequency network (SFN) multipath interference (MPI) at the receiver antenna is a critical issue. The capability to elaborate those signals depends on both synchronization strategy used and coded orthogonal frequency-division multiplexing (COFDM) modulation properties. An efficient COFDM system requires the guard interval as short as possible and such as to protect the system from the MPI. To cope with this problem a static delay adjustment obtained by fitting the launching time of each broadcast station is introduced so that the time difference of the incoming signals within the reception area falls into the COFDM guard interval. This work investigates the use of advanced simulated annealing to optimize transmission parameters such as the static delay, gain and orientation of sector antennas of DVB-T/T2 transmitters in SFNs. In particular, we identify the static delay adjustment map to maximize the coverage area according to quality of service requirements. The proposed approach has been validated for DVB-T/T2 network planning in the Basque Country (Spain) and Sardinia island (Italy). Different density of transmitters, generalized geography of territory are considered for the best setting of the parameters involved. Reliability studies are included to demonstrate the efficacy of the method.

3.1 INTRODUCTION

The digital broadcasting systems based on coded orthogonal frequency-division multiplexing (COFDM) modulation, such as DAB [5], DVB-H [6], DVB terrestrial DVB-T [7] and more recently DVB-T2 [8], commonly exploit the single-frequency network (SFN) technique [9]. An SFN is a broad-cast network where several transmitters simultaneously send the same signal over the same frequency channel. In SFNs, setting the synchronization strategy receiver side and reducing the inter-symbol interference (ISI) between signals from different transmitters are important issues in network planning. Both aspects are involved to provide a good quality of service (QoS) in the intended coverage area. As described in [10], heuristic and local search [11] algorithms are good candidate solutions to solve problems involving digital terrestrial broadcasting. In [12], particle swarm optimization (PSO) [13], [14] and simulated annealing (SA) [2] heuristic algorithms were considered and compared: a propagation prediction tool was developed to estimate and optimize the transmission parameters based on 3D digital terrain maps and path loss based on ITU-R P.1546 [15] in case of DVB-T and DVB-H. The treatment of ISI and static delays [16] to

improve synchronization of the signals that arrive within the area of interest of the receiver was considered to optimize the coverage. Results showed that PSO outperforms the SA in terms of accuracy, but requires large convergence time especially in case of large networks.

In this work we increase the density of transmitters which turned out in a greater computational complexity due to higher self-interference data matrices dimension involved in the static delay optimization. To cope with the computational burden SA heuristic optimization has been considered and the relevant performance in a double SFN real scenario: Basque Country (Spain) and Sardinia island (Italy) was studied. We defined the generalization of the SA method by adapting it to a different orography of the territory in the two above scenarios and evaluated the effectiveness of advanced SA (ASA) in a coverage optimization problem. 4K and 8K mode with a guard interval (GI) of 1/4 of the duration of the COFDM symbol mode were selected which correspond to available and frequently used modes for the DVB-T/T2 networks [17]. In addition, the effect of sector and omnidirectional antennas both at transmitter and receiver sides was also studied. Results show that the proposed ASA allows an efficient set-up of the search algorithm by faster locating the global minimum of the objective function, thus reducing the global convergence time with respect to SA. Moreover, higher accuracy in the coverage is achieved for fixed QoS.

Coverage optimization involving a third region with a different orography of territory has been studied and optimized using genetic algorithms (GA).

3.2 PLANNING SFN NETWORKS

3.2.1 PROBLEM DESCRIPTION

An SFN is a broadcast network where several transmitters simultaneously send the same signal over the same frequency channel. An SFN requires all transmitters to be synchronized in frequency and time so that receivers in the coverage area may receive signals coming from different transmitters within the guard interval.

Emission time of each COFDM symbol can be adjusted (i.e., delayed) to a certain percentage of the guard interval. This is called “SFN static delay adjustment” and the relative delay configuration at each location of the coverage area is changed.

Moreover, when deploying an SFN, a set of system parameters needs to be well configured: equivalent isotropic radiated power (EIRP), radiating system height and antenna characteristics (i.e., radiation diagram, azimuth and tilt).

The theoretical model of DVB-T/T2 COFDM receivers, under SFN conditions, supposes that the total amount of received signals within the guard interval at the receiver contributes

to the total usable signal strength. Similarly, the echoes coming from transmitters whose relative delay exceeds the guard interval contribute to the overall interferences.

Coverage optimization is obtained by adjusting the relative delays between different transmitters in order to find the optimum relative delay combination that maximizes the useful signal strength and minimize interferences at the receivers, in as many sites as possible.

3.2.2 COFDM SYMBOL AND STATIC DELAY OVERVIEW

In COFDM modulation the lower the symbol rate the lower the inter-symbol interference caused by multipath propagation (MPI). The GI reduces the sensitivity to time synchronization problems: the greater the GI, the lower the self-interference. On the other hand, longer GI implies less efficient transmission and so the effective data rate is reduced. An efficient system requires the GI as short as possible, but able to protect the system from the MPI. In SFNs a static delay adjustment is introduced so that the time difference of the incoming signals falls into the COFDM GI within the reception area. The static delay adjustment is performed by fitting the launching time of each broadcast station and has a direct impact on the relative delay configuration at each location and therefore on the global coverage. In the case of a dense SFN network, the delay adjustment of a high number of transmitters is not a trivial issue and due to the number and complexity of the variables involved, heuristics are a good candidate solution as already shown in [12], [13], [16], [18].

3.3 SFN NETWORK SIMULATION MODEL

This work considers two real existing SFNs and aims at providing optimized future expansion of the network [19] without changing the physical location of the transmitters [20]. Once the frequency and transmission power of each transmitter are defined, a propagation prediction loss model provides a database that contains a contribution of delay and power received for each location within area of interest.

Figure 3 shows a simplified example. The network simulation model includes the terrain height database, determined generally by a grid 100 × 100m accuracy. A configuration with three transmitters and three pairs of power and propagation delay (PT_x,DT_x) is represented for a generic receiving point. In digital broadcasting systems based on COFDM modulation and statistical models, the QoS is estimated as a function of carrier to interference (C/I) ratio. In our study, the QoS is determined in terms of the threshold value C/I required for the correct operation of the DVB-T/T2 system.

Figure 3 - Example of network simulation model: a yellow dot generic receiver contains information represented by three couples of power received and relative delay by three TXs

Let us consider an SFN consisting of N transmitters distributed over a certain region for which the 3D digital terrain maps are known. According to the COFDM properties the signals with relative propagation delay shorter than the GI contributes constructively to the useful part of combined signals, C, while the signals that arrive outside the GI are treated as interference, I. Therefore, at the receiver

$$\frac{C}{I} = \frac{\sum_{i \in \Omega} c_i \cdot w_i \cdot (\delta_i - \delta_0)}{\sum_{i \in \Omega} c_i \cdot [1 - w_i \cdot (\delta_i - \delta_0)]} \tag{4}$$

where, c_i [dBm] denotes the power received from the i-th transmitter, w_i [0,1] is the fast Fourier transform (FFT) receiver mask weighting value of i-th echo, δ_i [μs] the i–th echo delay relative to the synchronization reference δ_0 [μs] selected among δ_i which maximize (1). The set of transmitters is represented by $\Omega = \{1, \ldots, N\}$.

$$w_i(t) = \begin{cases} 0 & if \quad t \le \Delta - T_p \\ \left(\frac{T_u + t}{T_u}\right)^2 & if \quad \Delta - T_p < t \le 0 \\ 1 & if \quad 0 \le t \le \Delta \\ \left[\frac{(T_u + \Delta) - t}{T_u}\right]^2 & if \quad \Delta < t \le T_p \\ 0 & if \quad T_p < t \end{cases} \tag{5}$$

Several studies [21], [22] provide a model of FFT weighting function, which considers the effect of possible signal paths that arrive faster or slower than the current symbol synchronization time reference, i.e., pre-echoes and post-echoes respectively. These echoes are partially contributing and partially interfering. In this work we consider the weighting function [21] where T_u denotes the useful symbol length, Δ is the GI length and T_p the interval during which signals contribute constructively, as shown in Figure 4.

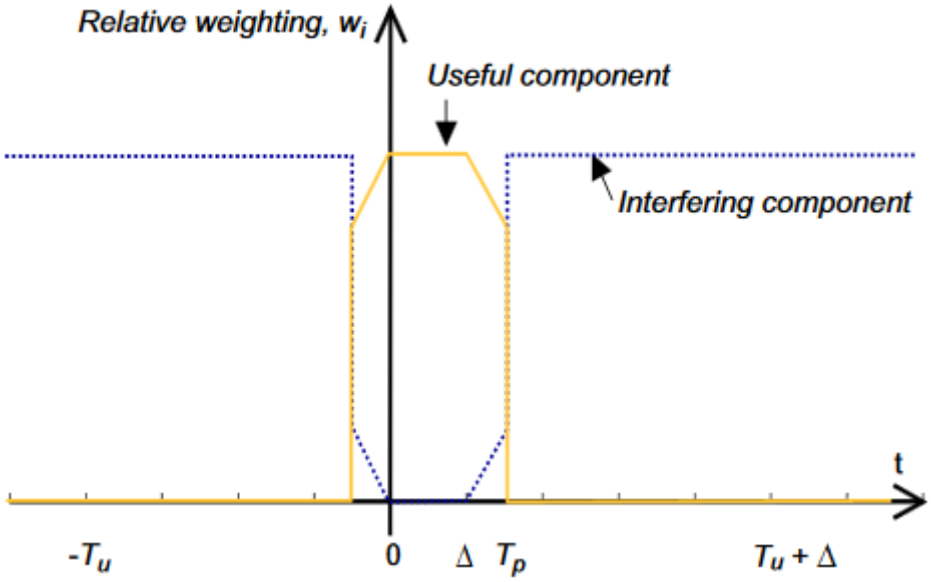


Figure 4 – DVB-T MODEL – splitting of the signal power into useful and interfering components.

Another crucial issue in COFDM-based broadcasting network planning is the FFT synchronization at receiver side. In this work we rely on Centre of Gravity (CoG) strategy [19].

CoG locates the FFT window on:

$$t_c = \frac{\sum_i p_i t_i}{\sum_i p_i} \tag{6}$$

where t_i is the arrival time of the i-th signal and p_i represents its corresponding power level. The CoG approach responds well to pre-echoes and delayed signals of similar amplitude.

It does not fix the FFT window to a particular signal, but takes into account the average behavior of the impulse response of the transmission channel.

3.4 HEURISTIC APPROACH

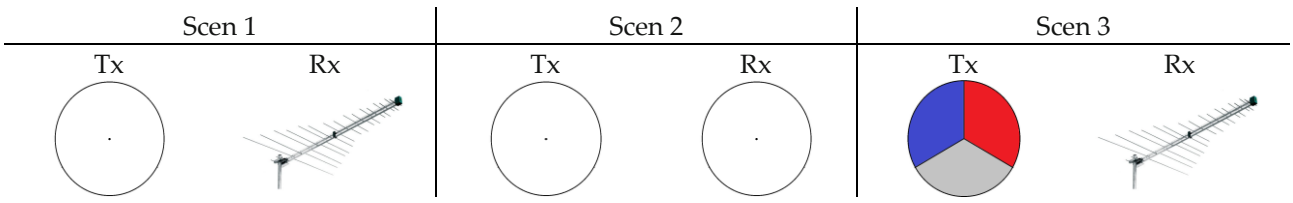
3.4.1 SFN NETWORK HEURISTIC OPTIMIZATION

As to our optimization, let us define the fitness function F as follows:

$$F(C,C/I)=\frac{L}{\sum_{l=1}^{L} COV_l},\ COV_l=\begin{cases}1, & if\ \ C>C_{\min}\ and\ \dfrac{C}{I}>\left(\dfrac{C}{I}\right)_{\min}\\ 0, & otherwise\end{cases} \tag{7}$$

where l is a generic receiver location within area of interest, L is the total receiver locations and COV_l is the coverage in the l-th location. The optimization process identifies the parameter δ_0 among the δ_i which minimizes (7) by means of (4) to achieve the highest percentage of coverage for fixed QoS requirements. QoS requirements are fulfilled when the total power C of the useful signals and the C/I ratio at the receiver antenna concurrently are above the threshold values C_{min} and $(C/I)_{\min}$ respectively. Different scenarios have been considered for this optimization:

- Scen1 optimizes transmitters delay considering at the receiver side a typical directive antenna pattern with a beam width of 60° and 20° in azimuth and elevation, respectively, a gain of 13 dBi and−12 dB sidelobes level. Omnidirectional antennas are considered at the transmitters;
- Scen2 optimizes transmitters delay but now considering omnidirectional antennas at both transmitters and receivers sides. Scen2 carrying out the optimization of the static delay for the N transmitters;
- Scen3 is similar to Scen1 but now 120° sectors are considered for the transmitters, and besides transmitters delay, the gain to be applied to any of the sectors as well as their relative orientation are also optimized.



In Scen2, omnidirectional antennas are considered for each location and therefore a N dimensional vector $X_1 = (x_{L,1}, x_{L,2}, ..., x_{L,N})$ is optimized, where $x_{L,i}$ is a L dimensional vector consisting of the static delays for the i-th transmitter and in which the dynamic range for any static delay lies in the interval $[0, \Delta]$. In Scen3 analysis, 120° sector antennas are considered at each location, the algorithm therefore considers a N+4 dimensional vector $X_2 = (X_1, x_{L,N+1}, x_{L,N+2}, x_{L,N+3}, x_{L,N+4})$ where $(x_{L,N+1}, x_{L,N+2}, x_{L,N+3}) \in [-5\ dB, 0]$ are the attenuations to be applied to each sector antenna and $x_{L,N+4} \in [0, 120°]$ their relative shift orientation with respect to the North cardinal point.

The complete approach is shown in Figure 5. Let us consider an SFN consisting of N transmitters distributed over a certain region for which the 3D digital terrain maps are known. The optimization method proposed shows different blocks that interact in order to obtain the optimum configuration for the transmitters of the SFN network. For the network infrastructure, the prediction of the propagation is used to estimate the contribution in terms of power, Ci and delays δ_i for all the transmitters and receivers. In order to evaluate the optimization algorithm the ITUR P.1546 propagation prediction loss has been used in this work. Path loss block also allows the variation of sectorization and orientation of the antennas and their optimum gain, (i.e., $x_{L,N+1}$, $x_{L,N+2}$, $x_{L,N+3}$, $x_{L,N+4}$). The use of heuristic algorithms and a propagation prediction tool allows estimating and increasing the power received at each location according to the ITU–R recommendation P.1546 [15]. This information consists of couples ($C_{L,N}$, $\delta_{L,N}$) and represents the input to the SA and ASA algorithms.

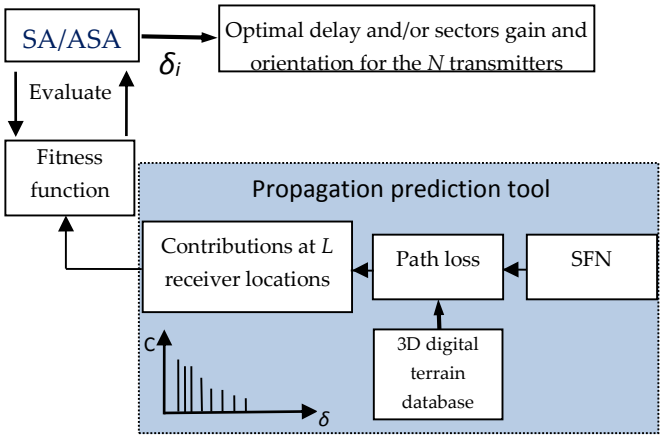


Figure 5 - Block diagram of the approach presented.

3.5 DVB-T2 SCENARIOS

In this section we consider two SFNs for DVB-T2, 4K and 8K modes involving mobile reception and optimizing the coverage by efficiency of the ASA

3.5.1 BASQUE COUNTRY STUDY CASE

The SFN topology analyzed for SA and ASA, concerns the geographic region of the Basque Country, in northern Spain. The region exhibits a very complex orography. The area is around 150 km long and 115 km wide. The region is shown in detail in Figure 6 with an indication of the location of the N=18 transmitters considered for the SFN. Within the simulation area is considered a grid of 100×100 m resolution corresponding to a set of L = 1.173×10^6 receiver locations. The initial stage, prior to optimization assumes that the power transmitted by each transmitter is equal to 1KW EIRP. The frequency band used in this SFN is the Ultra High Frequency (UHF) band V, channel 60–786MHz for both DVB-T and DVB-T2.

Furthermore, according to (7), any receiver location, l, is marked as served or covered only if both conditions $C > C_{min}$ and $C/I > (C/I)_{min}$ are satisfied considering the threshold levels in Table II [18], [23].

DVB-T2 Threshold levels		
	Cmin	*(C/I)min*
DVB-T2 8K	−75 dBm	13.4 dB
DVB-T2 4K	−85.7 dBm	10.2 dB

Table II - Threshold values for DVB-T2

In Table V, we considered the DVB-T2 8K mode for a simplified version of the approach, considering omnidirectional antennas and antennas far-field pattern for transmitters and directional antennas oriented towards the nearest transmitter for receivers (Scen1). DVB-T2 4K mode considers omnidirectional antennas at both transmitter and receiver sides (Scen2) in order to save computational cost for this parametric comparison. Subsequently, in Table VI two different scenarios have been considered for DVB-T 8K mode, Scen2 and Scen3 previously described. ASA has been tested with different COFDM modes: DVB-T 8K (useful symbol T_u = 896μs and 1/4 GI = 224μs) and DVB-T2 4K (useful symbol T_u = 2220μs and 1/4 GI = 555μs).

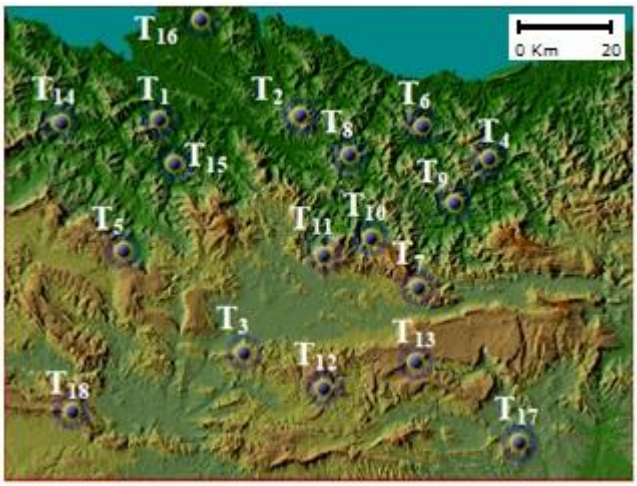


Figure 6 - Details of the Basque Country SFN with T=18 transmitters in which the analysis has been carried out.

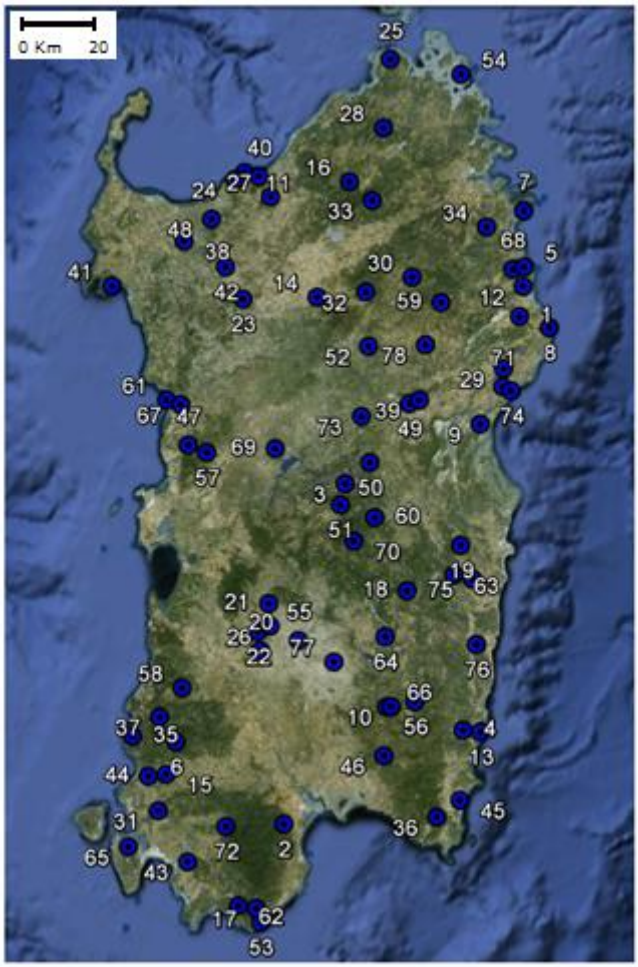


Figure 7 - Sardinia study case, 78 transmitters SNF network

3.5.2 SARDINIA STUDY CASE

The SFN Sardinia network given in Figure 7 consists of 78 transmitters. This region is around 260 km long and 120 km wide in expanse. The number of receiver locations is equal

to L = 3.129× 10^6. Broadcast band and frequency related this SFN network are Very High Frequency (VHF) band III, channel E9 - 205.5 MHz. These network planning data were obtained from [24]. Power transmitted for all sites is equal to 1kW EIRP. Modulation used is 16 QAM, code rate 1/2, GI 1/4 and bandwidth 1.7 MHz. Scattered pilot pattern is equal to PP1. At receivers side, a mobile reception at no less than 1.5 m above ground level is considered. All simulations were performed using the 4K mode [23]. Furthermore, the useful symbol length T_u is equal to 2, 220 μs and 1/4 GI = 555μs. $(C/I)_{min}$ and C_{min} are considered according to Table II as the minimum receiver signal field strength considered in the simulations.

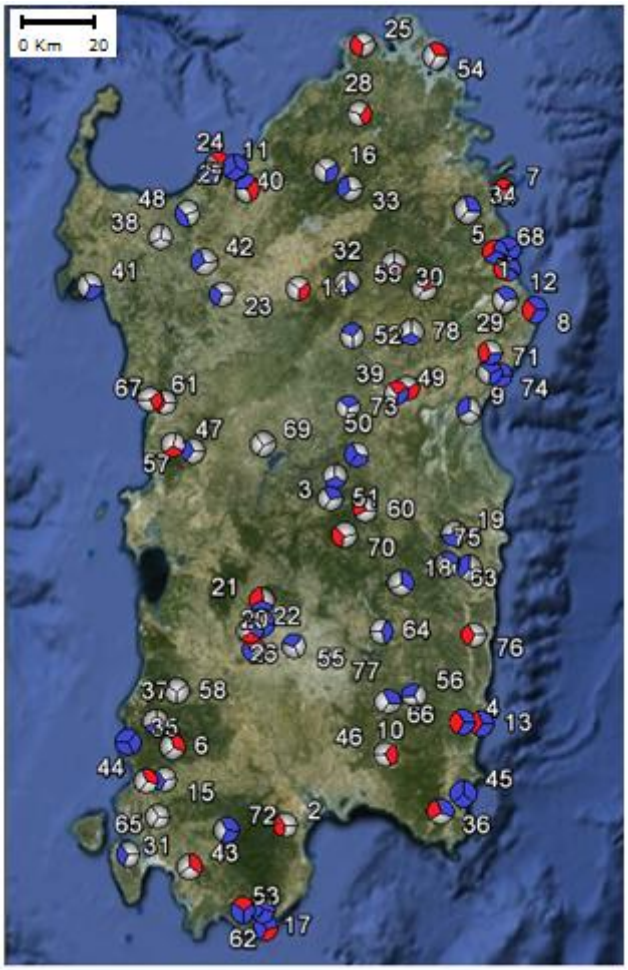


Figure 8 - Optimized SFN Sardinia network (Scen2): gain sectors and relative orientation. The white, blue and red sectors represent the gains in the ranges [-0.99, 0], [-2.99, -1] and [-5, -3] dB, respectively.

3.6 ITU-R PATH LOSS COMPARISON

In this section, the contribution of the N transmitters at each receiving location has been estimated via two prediction methods for point-to-area terrestrial services in the VHF and

UHF bands: the ITU-R P.1546 [15] and the ITU-R P.1812 [25]. ITU-R P.1546 forecast a simplified study of clutter obstruction and terrain clearance, it considers minimum diffraction effects and takes account a limited analysis of the geography of the territory. Furthermore, ITU-R P.1546 is not suitable for hilly regions in which relevant obstacles are close to either transmitter or receiver side between distances of 3 to 15 Km in the direction of the receiving/mobile antenna. On the other hand, ITU-R P.1812 defines a real physical model analysis. A more accurate diffraction effects, ducting and surfaces reflection are considered. This recommendation treats the type of soil and a comprehensive analysis of the digital elevation map.

Considering that there is a need to give guidance in the planning of terrestrial radio communication services in the VHF and UHF bands and, noting that ITU-R P.1546 provides guidance on the prediction of point-to-area field strengths in the VHF and UHF bands based principally on statistical analysis of experimental data, ITU-R P.1812 integrates and complements ITU-R P.1546 introducing a path-specific propagation method [26]. The complete approach is shown in Figure 9 in which three interconnected blocks work together to improve coverage area over a region of interest. The method models the propagation, the terrestrial reception and the optimization algorithms to achieve the optimum configuration for the SFN transmitters.

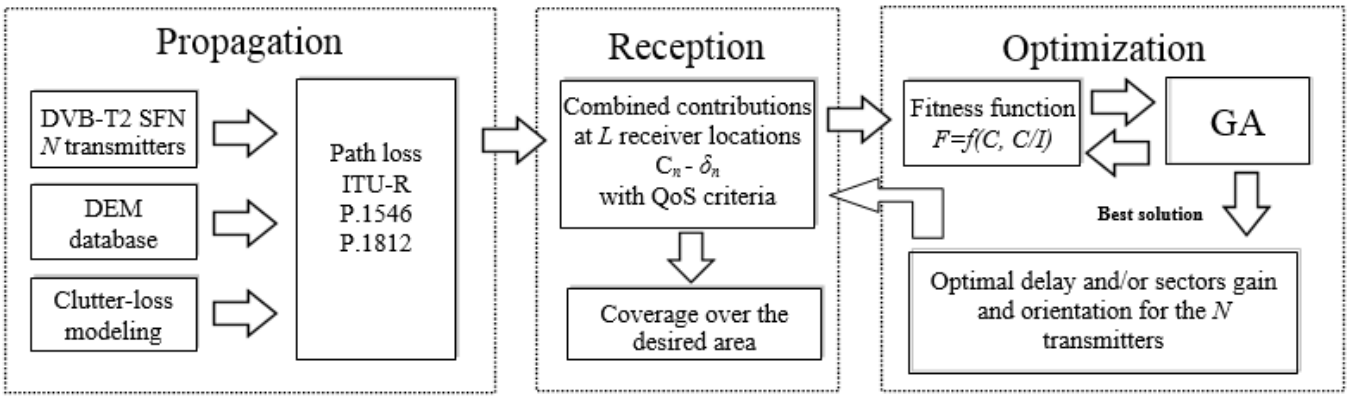


Figure 9 – Block diagram of the SFN planning tool using ITU-R P.1546 and P.1812

3.6.1 THE GA BASED APPROACH

For this purpose different operator values from region to region were identified and used as shown in Table III.

The task of Selection operator is to select and allocate reproductive units (i.e., parents) so that these, reproducing, can give birth to the next generation. The general criteria for

Selection is the comparison between the fitness function values of all individual candidates: individuals with better fitness function value are more likely to be copied.

The Crossover operator performs the task of communicating the genetic heritage of two individuals, in order to donate it to the next generation. Finally, the Mutation function specifies how the GA performs small random changes between individuals of the population so that they are generated mutated children. The Mutation function leads to genetic diversity and GA allows searching the solution in a large space.

Operators	**Region of Interest**		
	Sardinia	**Belgium**	**Basque Country**
Population	200	90	100
Generation	240	110	120
Selection function	Tournament to 8 individuals	Tournament to 4 individuals	Tournament to 4 individuals
Crossover function	Crossover 2 point	Crossover 2 point	Crossover 2 point
Mutation function	Uniform mutation-rate 0.3	Uniform mutation-rate 0.1	Uniform mutation-rate 0.2

Table III - Genetic Algorithm Parameters

3.6.2 REGION OF INTEREST: BASQUE COUNTRY, SARDINIA ISLAND, FLANDERS REGION

In this section, three different regions that exhibit a variable complexity have been analyzed.

The first SFN topology analyzed for A-GA, concerns the geographic region of the Basque Country, in northern Spain. The region shown in Figure 10 exhibits a very complex orography.

As it can be seen for Basque Country region, Sardinia region exhibits a very complex and no-homogeneous orography.

Flanders region orography of territory has been depicted in Figure 11 whereas Figure 13 illustrates the SFN Flanders network consisting of 16 transmitters. This region is around 215 km long and 65 km wide in expanse. The number of receiver locations is equal to $A=1.3975x10^6$ considering a 100x100 m resolution.

In this section, three real SFNs for DVB-T2, different OFDM modes (i.e., 4K, 8K, 16K, 32K) and GI (i.e., 1/4, 1/8, 1/16), involving mobile and portable reception (i.e., 1.5 m receivers height) and optimizing the coverage by efficiency of the GA are been treated.

The initial stage, prior to optimization assumes that the power transmitted by each transmitter is equal to 1 KW EIRP. The frequency band used in this SFN is the Ultra High Frequency (UHF) band V, channel 60 - 786MHz. According to [23] different GI and (C/I)min are been considered. In Table II, OFDM modes, bandwidth, QAM modulations, GI and

relative threshold value (C/I)min are outlined to decide whether a receiver location within the simulation area is covered or not.

According to (3), any receiver location, *l*, is marked as covered only if both conditions $C>C_{min}$=-89 dBm and $C/I>(C/I)_{min}$ are satisfied considering the threshold levels [22] [23] detailed in Table IV.

OFDM modes	Bandwidth	Guard Interval	Modulation	$(C/I)_{min}$
4K	1.7 MHz	1/4 (555 μs)	16 QAM	10.2 dB
8K	8 MHz	1/4 (224 μs)	16 QAM	10.2 dB
16K	8 MHz	1/8 (224 μs)	16 QAM	9.8 dB
32K	8 MHz	1/16 (224 μs)	64 QAM	17.9 dB

Table IV - Threshold Values for DVB-T2

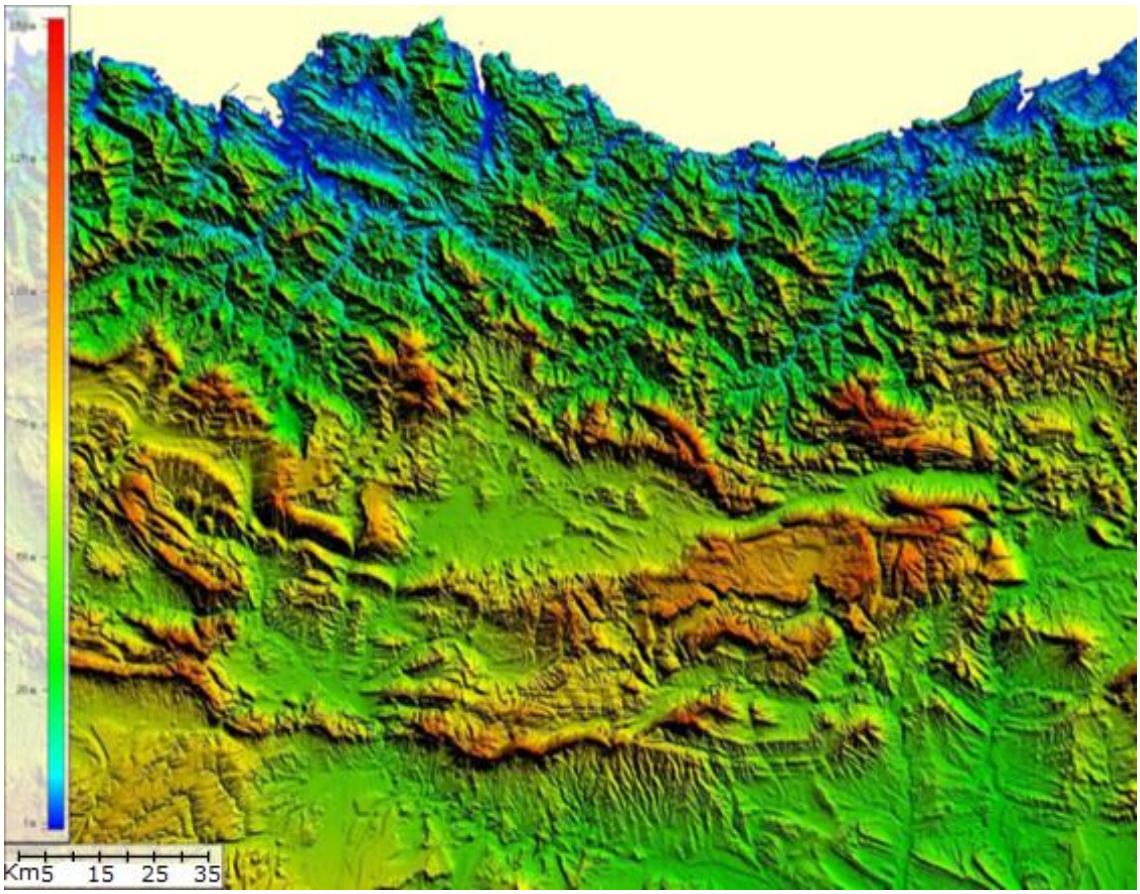


Figure 10 – Basque Country orography of territory

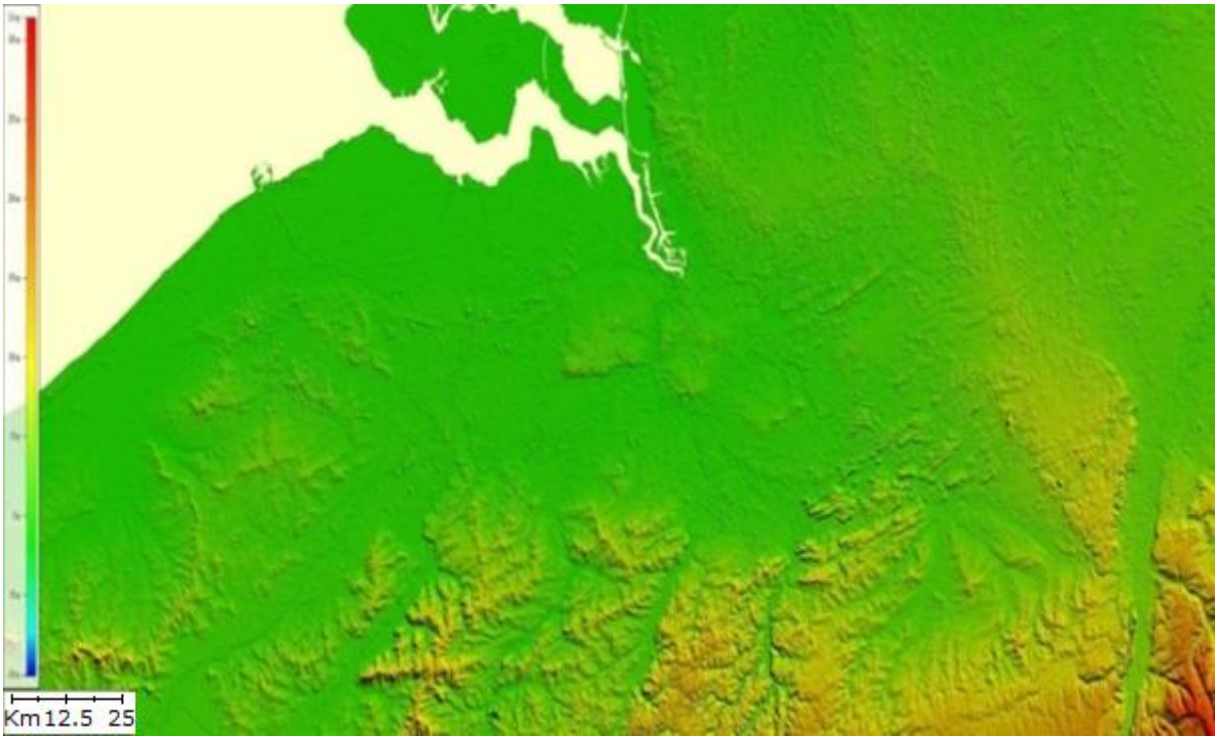


Figure 11 – Flanders orography of territory

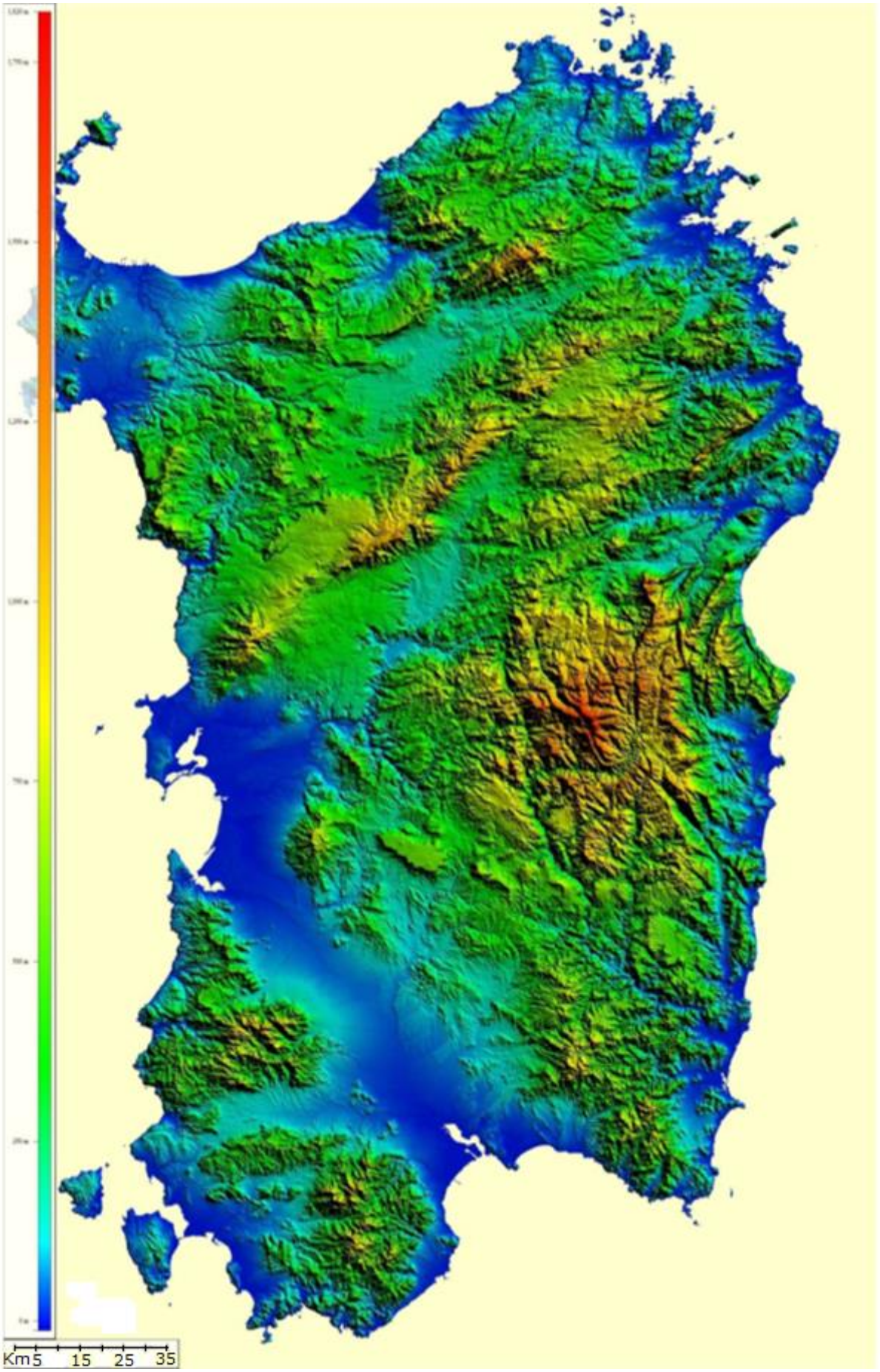


Figure 12 – Sardinia orography of territory

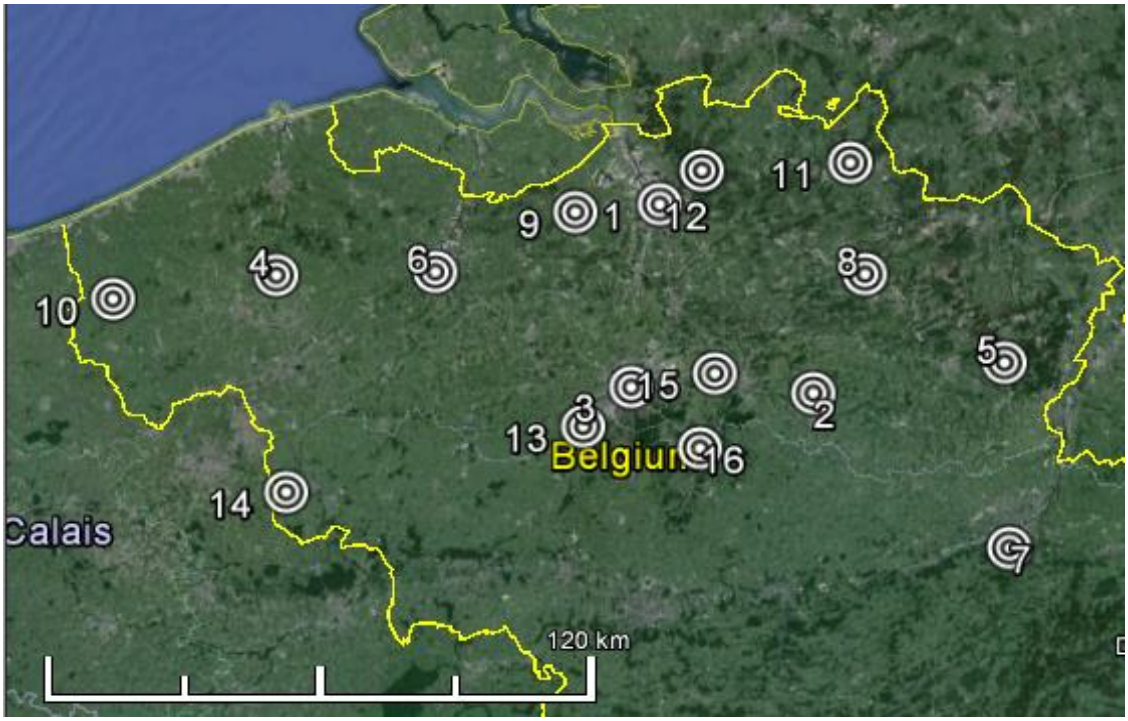


Figure 13 – Flanders SFN 16 transmitters

3.7 DVB-T2 SIMULATION RESULTS

Table V details a study regarding Basque Country original SFN and ASA in term of $C > C_{min}$, $(C/I) > (C/I)_{min}$ and Coverage [12], [13], [16], [18]. Coverage indicates the percentage of receiver locations in which either the $C > C_{min}$ or $C/I > (C/I)_{min}$ criteria, or both, *Coverage*, are accomplished. The initial network without static delay optimization presents a coverage of 77.85% for DVB-T2 8K mode and 68.30% for DVB-T2 4K mode. The estimated coverage with static delays and ASA optimization increases up to 77.96% and 69.83% for 8K mode and 4K mode, respectively. Therefore, the high complexity of the Basque Country territory and related radio-wave propagation affects the fitness function and makes it so hard that SA and ASA schemes achieve significant improvements. Table VI summarizes a detailed study regarding original SFN without static delay, SA and ASA in term of $C > C_{min}$, $C/I > (C/I)_{min}$ and *Coverage* for both Scen2 and Scen3. In Scen2, 8K mode, GI 1/4 for DVB-T transmitters and omnidirectional antennas at both transmitters and receivers sides are implemented. In Scen3, 120 degrees sectors for the transmitters are considered and the initial coverage (82.47%) increases up to 87.36% using ASA. ASA optimization proves a gain of approximately 1% with respect to SA both for Scen2 and for Scen3. Better results, comparing initial coverage without static delay optimization and ASA simulations, are observed. Coverage increases by up to 2.22% and 4.89% for Scen2 and Scen3, respectively. In Table

VII, the difference between SA and ASA has been evaluated when Scen1 in Sardinia island is considered.

DVB-T2 8K	SFN without static delays	$C>C_{min}$	78.0 %
		$(C/I)>(C/I)_{min}$	99.75 %
		Coverage	77.85 %
	ASA	$C>C_{min}$	78.39 %
		$(C/I)>(C/I)_{min}$	99.78 %
		Coverage	77.96 %
DVB-T2 4K	SFN without static delays	$C>C_{min}$	77.85 %
		$(C/I)>(C/I)_{min}$	80.46 %
		Coverage	68.30 %
	ASA	$C>C_{min}$	77.86 %
		$(C/I)>(C/I)_{min}$	84.83 %
		Coverage	69.83 %

Table V - Static delay optimization in Basque Country

	Coverage (%) DVB-T		
	$C>C_{min}$	**$C/I>(C/I)_{min}$**	**Coverage**
No optimization (without static delays)	92.72	87.04	82.47
Scen1 (SA)	91.98	90.23	83.72
Scen1 (ASA)	91.49	91.17	84.69
Scen2 (SA)	91.15	93.56	86.24
Scen2 (ASA)	90.93	92.19	87.36

Table VI - Comparison of the SA and ASA optimization schemes under different conditions of reception and transmission in Basque Country

	Coverage (%) DVB-T2		
	Min	***Max***	***Average***
No optimization (without static delays)	--	71.13	--
SA (Scen1)	77.33	77.39	77.36
ASA (Scen1)	78.12	78.14	78.13

Table VII - Comparison of the SA and ASA optimization schemes in Sardinia

DVB-T2 Mode			**CoG (%)**
DVB-T2 4K	SFN without static delays	$C>C_{min}$	73.58
		$(C/I)>(C/I)_{min}$	82.83
		Coverage	61.87
	SA	$C>C_{min}$	72.96
		$(C/I)>(C/I)_{min}$	83.28
		Coverage	67.38
	ASA	$C>C_{min}$	73.09
		$(C/I)>(C/I)_{min}$	83.56
		Coverage	68.12

Table VIII - Static delays optimization for cog in Sardinia

	Coverage (%) DVB-T2		
	$C>C_{min}$	$C/I>(C/I)_{min}$	*Coverage*
No optimization (without static delays)	82.27	80.13	78.36
Scen1 (SA)	84.76	82.98	80.09
Scen1 (ASA)	84.97	83.21	80.83
Scen2 (SA)	87.46	85.62	82.30
Scen2 (ASA)	88.25	86.17	83.67

Table IX - Comparison of the SA and ASA optimization schemes under different conditions of reception and transmission in Sardinia

Table VII shows the coverage improvement by using static delays for DVB-T2 8K mode, Scen1 configuration, SA and ASA optimization. Table VII summarizes the results obtained in terms of estimated coverage for the SA and ASA, including for the two schemes independent runs the minimum, maximum and average coverage achieved in each case. Improvements by up to 6.26% and 7.01% for SA and ASA, respectively, have been reached. It appears evident that more appreciable results for Sardinia island with respect to the Basque Country both for SA and ASA schemes have been obtained. This enhancement depends on different factors, such as, for instance, the size and orography of territory, the physical location of DVBT/T2 transmitter antennas, and the optimal mapping of the static delay optimized by SA and ASA. Table VIII shows coverage for DVB-T2 4K mode when ominidirectional antennas are considered for both transmitters and receivers sides (Scen2) and CoG synchronization strategy is evaluated. Once more, SA and ASA improve the initial coverage fulfilling both the conditions $C > C_{min}$ and $(C/I) > (C/I)_{min}$. The proposed optimization increases the coverage by up to 5.51% and 6.25% for SA and ASA schemes, respectively. Finally, in Table IX the outcomes related to SA and ASA for both Scen2 and Scen3 are presented. Also in this case, ASA scheme proves better performance rather than SA for both scenarios. A coverage of 80.83% for ASA(Scen2) scheme has been reached, whereas the initial SFN coverage of 78.36% increases by up to 5.31% (83.67%) using ASA(Scen3). The Scen3 configuration is shown in Figure 8.

In Table X, the DVB-T2 4K, 8K, 16K, 32K modes, considering omnidirectional antennas at both transmitter and receiver sides for Sardinia scenario have been considered. Simulation results show a better coverage using ITU-R P.1812 with respect to ITU-R P.1546. ITU-R P.1812 improves coverage by up to 8.04%, 0.41% and 2.05% for 4K, 8K and 32K modes, respectively. On the other hand, ITU-R P.1546 increases coverage by up to 1.45% with respect to ITU-R P.1812 for 16K mode.

Table XI details a study regarding Basque Country. Better results for 8K, 16K and 32K modes using ITU-R P1812 with respect to ITU-R P.1546 are observed. Otherwise, ITU-R P.1546 for 4K mode increases coverage by up 1.37% with respect to ITU-R P.1812.

Finally, Table XII shows coverage optimization for DVB-T2 in Flanders region. In this study case, GA proves better performance evaluating ITU-R P1812 rather than ITU-R P.1546 for all OFDM modes.

DVB-T2	Sardinia			
	4K	8K	16K	32K
ITU-R P.1546 Original SFN	61.87 %	78.36 %	89.04 %	94.85 %
ITU-R P.1546 Optimized SFN	68.12 %	80.73 %	**97.35 %**	96.45 %
ITU-R P.1812 Original SFN	75.77 %	80.83 %	88.35 %	95.25 %
ITU-R P.1812 Optimized SFN	**76.16 %**	**81.14 %**	95.90 %	**98.5 %**

Table X - Static Delay optimization in Sardinia Region

DVB-T2	Basque Country			
	4K	8K	16K	32K
ITU-R P.1546 Original SFN	68.30 %	77.85 %	80.05 %	82.16 %
ITU-R P.1546 Optimized SFN	**74.56 %**	78.46 %	80.21 %	82.31 %
ITU-R P.1812 Original SFN	67.70 %	79.32 %	80.11 %	82.17 %
ITU-R P.1812 Optimized SFN	73.26 %	**79.83 %**	**80.29 %**	**82.32 %**

Table XI - Static Delay optimization in Basque Country Region

DVB-T2	Flanders			
	4K	8K	16K	32K
ITU-R P.1546 Original SFN	73.89 %	77.13 %	83.82	85.25 %
ITU-R P.1546 Optimized SFN	77.26 %	78.47%	84.09 %	85.66 %
ITU-R P.1812 Original SFN	74.19 %	78.03 %	84.02 %	86.23 %
ITU-R P.1812 Optimized SFN	**79.83 %**	**81.82 %**	**85.12 %**	**86.25 %**

Table XII - Static Delay optimization in Flanders Region

3.8 CONCLUSIONS

A planning tool combining the heuristics SA, ASA and a propagation prediction tool is proposed with the final objective of optimizing SFN network coverage by reducing self interfered areas in Basque Country (DVB-T/T2) and the Sardinia island (DVB-T2). Results show that coverage increases by up to 4.9% in Basque Country (DVB-T 8K) and up to 7% in Sardinia island (DVB-T2 8K) for fixed receivers. Significant results are also obtained in Sardinia island for mobile reception: DVB-T2 4K mode increases coverage of 6.25% considering omnidirectional antennas at both transmitters and receivers sides. The coverage is improved by minimizing the areas suffering self-interference in the Basque Country and Sardinia. The study has demonstrated the efficacy of heuristic algorithms in coverage

optimization. Sectorization of antennas and synchronization strategies are also effectively considered. The SA has improved the DVB-T/T2 service of two existing networks without changing the physical location of the transmitters. The method can be used in the design phase in order to find the best location of the transmitters or in the event of existing networks expansion. ASA responds well to different configurations of hilly terrain and different density of transmitters introduced in the second network; results obtained shown effects of amendments. The increase of the computational complexity does not affect the effectiveness of the method and the quality of the solution The strategy is validated by means of simulation, showing good performance in front of typically done manually iterative adjustments by the radio planning engineer (i.e., static delays, gain and orientation of antennas). Any amendments via software require low economic resources. These amendments can be made independently by each broadcaster; each SFN network operates at a certain frequency and can have its own static mapping of delays. All this is possible by identifying the best set of parameters of SA. The efficacy of SA and ASA to optimize a complex system of network planning is verified in this work.

The obtained results have allowed assessing the level variation of the electromagnetic field strength in three different real scenarios (i.e., Sardinia, Basque Country, Flanders) with several OFDM modes and GI for mobile and portable reception.

A planning tool combining the heuristics GA and propagation prediction methods are proposed with the final objective of optimizing SFN network coverage by reducing MPI. Results demonstrated the efficacy of heuristic GA in coverage optimization and show an increased coverage for both ITU-R Recommendations.

This study has been focused on evaluating differences between ITU-R P.1812 and ITU-R P.1546 for DVB-T2 services. Results show a better propagation prediction in terms of covered receiver locations using ITU-R P.1812 rather than ITU-R P.1546. However, in certain cases the ITU-R P.1546 based principally on statistical analysis of experimental data, exhibits improvements respect ITU-R P.1812 (i.e., 4K mode and 16K mode for Basque Country and Sardinia, respectively).

4 COVERAGE OPTIMIZATION OF SATELLITE BROADCASTING

DVB-S2 and DVB-SH satellite broadcasting standards currently deploy 16- and 32-APSK modulation using the consultative committee for space data systems (CCSDS) mapping. Such standards also include hierarchical modulation as a mean to provide unequal error protection in highly variable channels over satellite. Foreseeing the increasing need for higher data rates, this work tackles the optimization of 64-APSK constellations to minimize the mean square error between the original and received symbol. Optimization is performed according to the sensitivity of the data to the channel errors, by means of genetic algorithms, a well-known technique currently used in a variety of application domains, when close form solutions are impractical. Test results show that through non-uniform constellation and asymmetric symbol mapping, it is possible to significantly reduce the distortion while preserving bandwidth efficiency. Tests performed on real signals based on perceptual quality measurements allow validating the proposed scheme against conventional 64-APSK constellations and CCSDS mapping.

4.1 INTRODUCTION

AMPLITUDE Phase Shift Keying (APSK) modulation with pre- and post- compensation schemes is deployed in satellite broadcasting [27], [28], [29] for its power and spectral efficiency over nonlinear satellite channels. Nevertheless, for multimedia broadcasting applications, further improvements by means of non-uniform constellations could be obtained [30]. As a matter of fact, multimedia streams employed in digital broadcasting are hierarchical by nature, so that bits associated with transmitted symbols present different error sensitivities. In particular, faults on most significant bits (MSBs) affect the transmission more than errors on the least significant bits (LSBs). For this reason, in the last years several works tackling the problem of guaranteeing scalable quality in satellite communications through hierarchical modulation have been presented [31]–[35] and integrated in standards such as DVB-S2 and DVB-SH. A hierarchical modulation carries two separate and independent bit streams. The primary (i.e., coarse) bit stream and the secondary (i.e., fine) stream. The former is intended for users with poor channel quality, whereas the latter stream refines the first one, but requires a larger signal-to-noise ratio (SNR) to be decoded

error-free. Hierarchical modulation can be used effectively to upgrade a digital broadcast system in response to both the demand for higher bit rate that is made possible due to advances in technology and coding algorithm development, and the need to be backward compatible to the already deployed old receivers.

In hierarchical modulation fine and coarse streams present different sensitivity to channel error and different techniques can be implemented to assure robustness. Channel coding techniques (i.e., FEC coding) to implement unequal error protection (UEP) have been studied in [36] for quadrature amplitude modulation (QAM) and in [37] for APSK, even though this introduces overhead and reduces bandwidth efficiency, which is a critical issue for satellite applications [38], [39].

In [40] modulation with unequal power allocation (MUPA) was proposed as a mean to improve the performance of conventional modulation schemes. In fact, saving bandwidth is verified in case of digital wireless communication systems that do not include channel coding for some reason. MUPA achieves UEP hierarchically distributing the available budget power over the symbols according to their sensitivity to channel errors, whereas the average transmission power per symbol remains unchanged. The resulting quality on received data was measured by means of the mean square error (MSE) between the original and decoded symbol. By increasing the robustness of MSBs MUPA reduces the average distortion (i.e., MSE) between transmitted and decoded data and the quality achieved at receiver side is improved without any increase of transmission bandwidth. MUPA performance evaluation does not consider conventional bit error rate (BER) measurements, because the final goal is to achieve minimal distortion among symbols which are considered as composed by bits with different importance and which are therefore unequally protected by an opportune optimized power distribution process [40].

The MUPA concept was applied to APSK constellations for satellite applications in [41] and [42] where 16 and 32-APSK modulations have been treated in order to achieve UEP through asymmetric layout of the constellation symbols. The optimization problem (OP) aimed at selecting the opportune radius of the constellation circles and the phase of each symbol that minimize MSE (i.e., the average distortion) between transmitted and decoded data. Due to the complex nature of the OP which does not allow close form solution, genetic algorithms (GA) [43] have been deployed. GA is a search technique used in many fields to solve complex problems which do not allow analytical derivation [44], [45]. GA fall within the class of gradient search techniques that need opportune setup to avoid sub optimal solutions due to the presence of local minima. In this work to set up the GA evolution the result of a previous study performed in [42] which analyzed the influence of the GA parameters on the accuracy and convergence performance of the algorithm has been considered.

The novelty of this work consists in studying the case of 64-APSK based on the results obtained in [41] and [42] foreseeing near future deployment of high order modulations in satellite broadcasting: new mappings derived starting from the one released by the consultative committee for space data systems (CCSDS) [46] and novel non-uniform constellations are proposed and compared with standard 64-APSK schemes. The outcome of this work demonstrates that the use of such modified mapping and non-uniform 64-APSK constellations through GA optimization yields to a reduction of the perceived distortion at the receiver. In fact, GA optimization discriminates the most significant piece of information in the modulated symbol obtaining reduced MSE for the proposed system with respect to the state-of-the-art conventional 64-APSK with CCSDS mapping. Correspondence between MSE and subjective perception has been validated by test performed on real signals. Digital audio and image signals and have been generated and transmitted both with the proposed optimized constellation and with conventional 64-APSK. In case of audio, performance evaluation has been measured by means of the perceptual evaluation of audio quality (PEAQ) algorithm [47]. PEAQ is an objective measurement technique recommended by the International Telecommunication Union – Telecommunication Standardization Bureau (ITU-T), which evaluates the quality of an audio signal by a single number, called objective difference grade (ODG), which varies within a range [-4÷0], with 0 the highest quality score. PEAQ has proven to achieve higher performance than conventional metrics based on MSE on the evaluation of the performance of the conventional audio codecs [48]. In case of image transmission, direct subjective evaluation is given to the reader through displaying the result by figures.

4.2 SYSTEM MODEL

In Figure 14, the model used for the optimization is introduced by schematically describing the various building blocks of the communication system.

Considering M to be the alphabet size of the constellation, each symbol represents a stream u(τ) of $\log_2 M$ bits generated by a memory less source and is put in parallel by the serial-to-parallel (S2P) block.

These symbols are then modulated by the APSK modulation block giving place to a complex number that represents constellation symbols (8):

$$v_{n,m}(\tau) = \rho_n \cdot e^{i\theta_m} \tag{8}$$

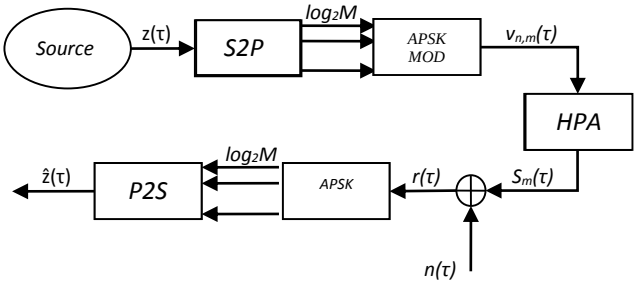


Figure 14 - System model

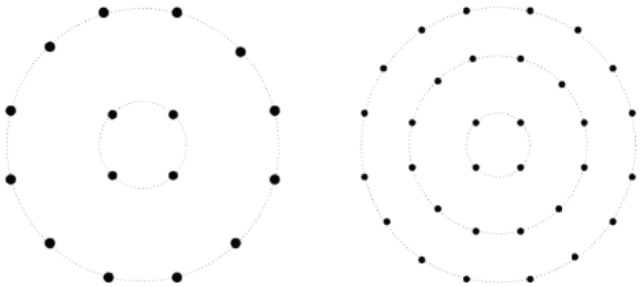


Figure 15 - 16-APSK constellation (left) and 32-APSK constellation

where ρ_n is the radius of the n-th circle of the constellation and θ_m is the phase of the m-th symbol. Figure 15 represents two examples for the 16 and 32-APSK.

Once constellation symbols are modulated through a squared root raised cosine (SRRC) filter, they are amplified by the high power amplifier (HPA) prior to transmission, thus being subject to the non-linear behavior of the amplifier, whose effects can be modeled using the Saleh model [49] resulting in the output $s_m(\tau)$. This model distinguishes two effects:

- the AM/AM non-linear effect that models amplitude distortions on the input signal;
- the AM/PM non-linear effect that models phase distortions on the input signal.

This impairment can be efficiently reduced by ad-hoc pre-compensation at the transmitter. Despite its hardware complexity impact, commercial satellite modems have already adopted advanced dynamic pre-compensation techniques for standardized 16- and 32- APSK modes [50]. The dynamic pre-distortion algorithm takes into account the memory of the channel, conditioning the pre-distorted modulator constellation not only to the current symbol transmitted but also to the *(L-1)/2* preceding and *(L-1)/2* following symbols (*L* being the number of symbols in total) [51].

Considering the channel to be of the additive white Gaussian noise (AWGN) type, transmitted symbols are affected by the addition of a random nuisance signal with zero

mean and variance $N_0/2$. Therefore, the received signal at the destination is $r(\tau) = S_m(\tau)+n(\tau)$. At the receiver, the signal is first passed through the SRRC filter. Note that due to the non-linear characteristics of the channel, this filter is no longer matched, and therefore ISI is introduced [49]. ISI appears at the receiver as the HPA, although memoryless, is driven by a signal with controlled ISI due to the presence of the modulator SRRC filter. This leads to an overall nonlinear channel with memory. As a consequence, the demodulator SRRC is not matched anymore to the incoming signal.

The resulting signal is passed to the demodulator that applies the maximum likelihood (ML) criterion in order to estimate the received symbol and it is then converted back from parallel to the serial signal $\hat{z}$.

In this work, to simplify the optimization at first the effect of SRRC filtering and therefore the ISI has been neglected. Moreover, only the AM/AM non-linear amplitude distortion A(ϱ), which represents the memoryless effect of the non-linearity, is taken into account through the formula

$$A(\rho) = \frac{\alpha\ \rho_n}{[1+\beta]\rho_n^2} \tag{9}$$

where α = 2.1587 and β = 1.1517 are standard values of the constants gathered from the literature [52] and obtained by means of curve-fitting techniques. In section 4.4 the above assumptions will be validated by testing the resulting optimized constellations in realistic scenario, that is considering real SRRC filters and opportune dynamic pre compensation techniques to mitigate ISI [37]. Dynamic symbol level pre-distortion considers the impact of memory introduced by the channel in order to compensate for the ISI.

The difference between the symbol generated by the memoryless source and the one estimated at the receiver can be computed as

$$d(\tau) = \left\| z(\tau) - \hat{z}(\tau) \right\| \tag{10}$$

and it is called distortion. The aim of the optimization presented in the next section is to minimize the average distortion defined as the minimum square error

$$MSE = E\left\{ [d(\tau)]^2 \right\} \tag{11}$$

which represents the optimization criterion. From (10) and (11) it can be noticed that errors in MSBs produce higher distortion, thus, higher MSE with respect to errors occurring in LSBs.

4.3 GENETIC ALGORITHMS OPTIMIZATION

OP is tackled by means of GA, a well-known algorithm used in a variety of fields for all those non-convex problems that require a technique able to tentatively crawl the solution space in order to find the best one.

At each iteration n, a GA gives birth to a generation G_n of potential solution vectors (also called chromosomes γ_i) that constitute the population of size p of the OP:

$$G_n = \left\{\gamma_1^{(n)}, \gamma_2^{(n)}, \dots, \gamma_i^{(n)}, \dots, \gamma_p^{(n)}\right\} \tag{12}$$

A vector of fitness scores S_n is also calculated for each generation using the objective function R:

$$S_n = \left\{R(\gamma_1^{(n)}), R(\gamma_2^{(n)}), \dots, R(\gamma_i^{(n)}), \dots, R(\gamma_p^{(n)})\right\} \tag{13}$$

The chromosomes with the highest fitness score are meant to be the closest to the desired solution and are thus selected for surviving and giving place to the next generation.

The next generation is created in three steps:

- *Selection* of the part of population with the best fitness score that will be parents for the next generation;
- *Crossover* of selected parents according to a mixing criterion in order to give birth to a number of children from each couple that will constitute the next generation;
- *Mutation* of a percentage of the offspring in order to spread the optimum solution search and avoid local optimal solutions.

The computation is stopped when the population has converged to the same fitness value which is supposed to be the optimal solution.

In the case of APSK the chromosomes are the radii and the phase of each symbol. The optimization criterion chosen is the minimization of the MSE, i.e., the expected minimum squared error between the transmitted and the received data.

Therefore, the function used to calculate the fitness scores is *R=1/(MSE)*. The optimization is constrained by the assumption that the peak-power limit imposed over the constellation cannot exceed the saturated power of the HPA.

For a more in-depth treatise about the use of GA in the context of constellation optimization, the reader can refer to [41] and [42].

4.4 MAPPING ANALYSIS

The goal of this work is to demonstrate that the performance of 64-APSK modulation can be improved in two successive steps: at first choosing opportune alternative mappings with respect to conventional CCSDS ones and then optimizing the position of the constellation symbols in terms of amplitude and phase, accordingly. These two options are strictly related, since in the optimization step the GA could modify the position (i.e., radius and angle) of the symbols in such a way that symbols on the same ring result interchanged thus modifying the mapping. In this section the influence of the mapping in term of inter symbol distortion will be analyzed. Consider the constellation described as a+b+c+d, where a is the number of symbols in the inner circle, d is the number of symbols in the outer circle and so forth. The transmission of equiprobable symbols using the 64-APSK scheme proposed by CCSDS [46] (that is of the type 4+12+20+28) is first presented. Afterwards, a new mapping of the 4+12+20+28 type and one of the 4+12+16+32 type specifically designed for the case under investigation are presented and compared with the CCSDS one.

4.4.1 CCSDS MAPPING FOR THE 4+12+20+28 CONSTELLATION

As previously said, the reference mapping used for the 4+12+20+28 constellation is the CCSDS mapping [46]. The considered mapping is symmetric with respect to the x and y axes. The advantage of such a mapping resides in its simplicity, since the coordinates of the symbols from a single quadrant are sufficient to describe the entire constellation. Therefore, in this case, the chromosomes used for the successive optimization will be as follows:

$$\gamma = [\rho_0, \rho_1, \rho_2, \theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}, \theta_{12}, \theta_{13}, \theta_{14}, \theta_{15}] \tag{14}$$

with ρ_0, ρ_1, ρ_2 indicating the three inner radii (the outer one is always equal to one thus not taking part in the optimization phase), and values from θ_0 to θ_{15} indicating the angles for the symbols with value from 0 to 15. Although symmetric mapping reduces the number of variables in the optimization problem, this limitation yields to a bigger distortion. In fact, in Figure 16(a) it is noticeable that the regions with a higher distortion value are those close to both the horizontal and the vertical axis. The meaning of numbers between the connection in each constellation figures represents such distortion. To reduce this effect, it is necessary to eliminate those critical zones dropping some of the symmetry constraints thus adding more variables and more complexity to the optimization problem. Considering that solving the optimization problem using 64 independent symbols is computationally impractical, exploiting the symmetry the number of variables involved is equal to 35.

4.4.2 PROPOSED MAPPING FOR THE 4+12+20+28 CONSTELLATION

The presented mapping has only a vertical symmetry and 35 variables: 3 regarding the radii and 32 regarding the angles. Exploiting the symmetry drop it is possible to radically change the mapping so that the distortion of the symbols of the outer radius that are the closest to the horizontal axis is minimized.

The resulting mapping is represented in Figure 16(b), where it is noticeable that symbols near the vertical axis keep high distortion levels as a consequence of the symmetry choices made in the design phase. In the horizontal axis instead, the distortion among symbols varies from 1 to 10, while in the case described in the previous section it was constantly equal to 16. This variety of distortion zones will be beneficial in the constellation optimization phase, since it will allow to keep higher distance between symbols which potentially can cause higher levels of distortion.

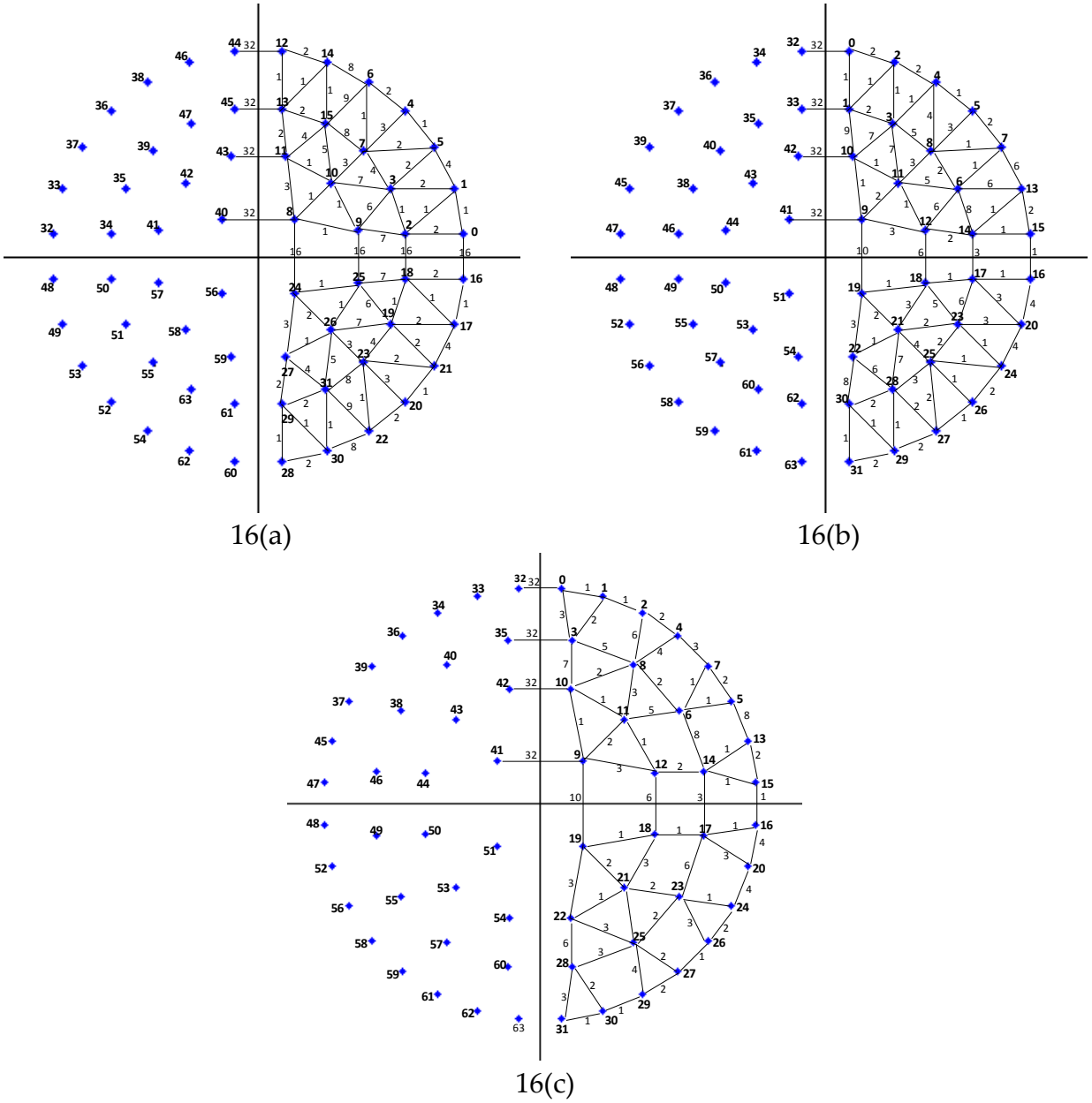
In Table XIII it can be seen that the average distortion for symbols on the same ring is smaller than in the CCSDS mapping case. Another important aspect is the distortion among symbols on different rings. This parameter is obtained averaging the distortion among all the neighbor symbols on two rings.

	CCSDS Mapping 4+12+20+28	Proposed Mapping 4+12+20+28
Outer ring vs. Ring 3	2.27	1.77
Ring 3 vs. Ring 2	4.85	5.42
Ring 2 vs. Inner ring	2	2

Table XIII - Comparison of the Average Distortion among Neighbour Symbols on Different Rings

	CCSDS Mapping 4+12+20+28	Proposed Mapping 4+12+20+28	Proposed Mapping 4+12+16+32
Outer ring	6	4.5	4.18
Ring 3	7.8	6.5	7.65
Ring 2	8.66	7.33	7.16
Inner ring	24	21	21

Table XIV - comparison of the average distortion among neighbor symbols on the same ring



16(a) 16(b)

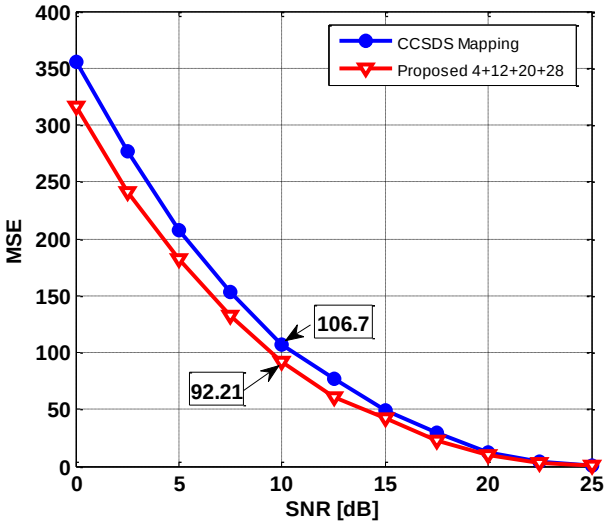
16(c)

Figure 16 - Distortion of the CCSDS mapping for the 4+12+20+28 (a), the proposed mapping for the 4+12+20+28 constellation (b) and the 4+12+16+32 composed 64-APSK constellation (c). Symbols and relative mutual distortion are expressed in decimal representation.

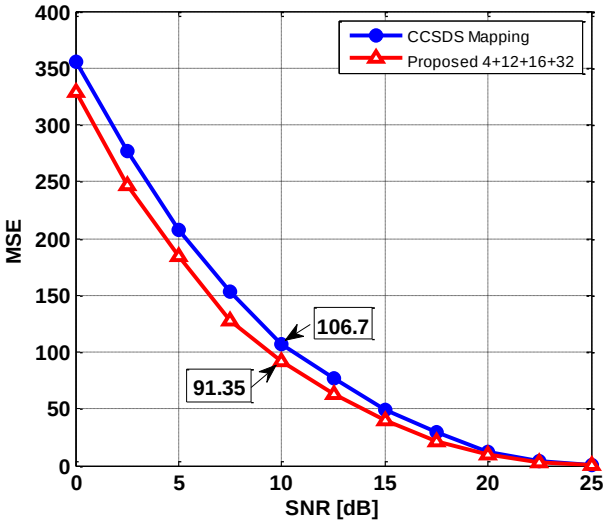
4.4.3 PROPOSED MAPPING FOR THE 4+12+16+32 CONSTELLATION

Sometimes, as shown in Figure 16(c), the 64-APSK constellation operates with 4 more symbols in the outer ring. This yields to smaller regions of decision in the outer part of the constellation, but leaves more space in the inner circles so that distortion can be more efficiently reduced in the inner zone. As in the case treated in section 4.4.2, also here only the vertical symmetry has been kept and the intra-ring distortion has been reduced especially in the outer symbols. Table XIV shows the distortion among symbols in the same

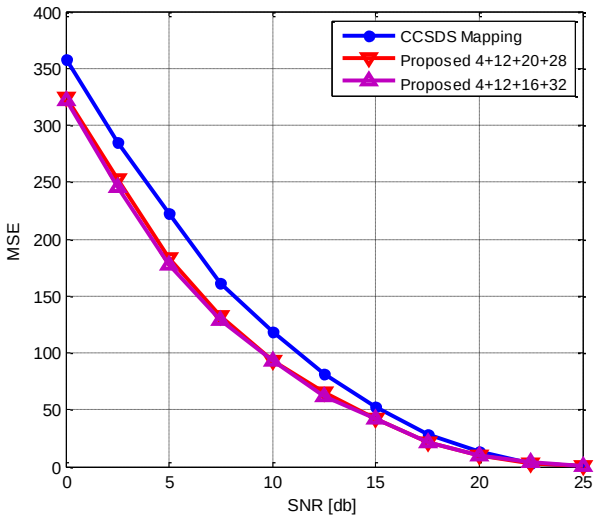
ring. As it can be seen, moving 4 symbols to the outer radius the resulting distortion is reduced while the one on the third ring is increased. This allows to better optimizing the constellation through the GA, since now the third ring has more space to put more distance between those symbols that create critical levels of distortion. In this case, comparing inter-ring distortion would lose its meaning, since the number of symbols between the two outer rings is changed.



17(a)



17(b)



17(c)

Figure 17 - MSE results for the CCSDS mapping and the 4+12+20+28 (a), 4+12+16+32 (b) and overall comparison of the presented constellations without GA optimization, (c).

4.5 Optimization Results

4.5.1 Results without constellation optimization

In the case of CCSDS presented in Figure 17(a), the first thing to notice is that the MSE = 106.7 at SNR = 10 dB while the average distortion is 10.32, that is approximately 16% of the maximum distortion level for 64-APSK. As already stated, the reason of this high level of distortion can be found in the presence of critical zones close to the axes. In the case of 28 symbols in the outer ring and proposed mapping, the mitigation of those high distortion zones already yields to better results for conventional uniform constellation. In particular, when SNR = 10 dB the MSE = 92.21 that is 13.6% less than in the CCSDS mapping case. In Figure 17(b), the distortion in the outer ring for SNR = 10 dB of the proposed constellation with 32 symbols, is 91.35 that means a reduction of 14.3% with respect to the CCSDS constellation.

Finally, in Figure 17(c) a comparison among the three mappings is given demonstrating that regardless of the number of symbols in the outer rings, the two proposed mappings have similar performances, with a slight predominance of the 4+12+16+32 case.

4.5.2 Results with constellation optimization

At first, to set up the GA evolution we refer to a previous study performed in [42] which analyzed the influence of the GA parameters on the accuracy and convergence performance of the algorithm. Based on [42] only 10 % of a chromosome could vary and a maximum of 40% of the chromosome parents are allowed to appear on the next generation. In the results that follow, the GA is run with a population of 100 chromosomes characterized by genes with values uniformly distributed on the constellation circles and probabilities of mutation and crossover $P_{mut} = 0.3$ and $P_{cross} = 0.5$, respectively.

Using these values, transmission over satellite is simulated according to the model presented in Section 4.2. Then, the fitness function is evaluated thanks to the function R already discussed. Once this step is completed, the GA modifies the population in accordance with the fitness results using the policies defined by the specific selection and crossover functions taken into consideration, that in this case are the tournament selection function and the two-point crossover function, chosen using the procedure already presented in [42]. The transmission is then repeated using the new generation until convergence or the maximum number of generations (in this case set to n = 200) is reached.

The considered chromosomes are respectively

$$\gamma = [\rho_0, \rho_1, \rho_2, \theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}, \theta_{12}, \theta_{13}, \theta_{14}, \theta_{15}] \tag{15}$$

for the case of the CCSDS mapping;

$$\begin{aligned}\gamma = [&\rho_0, \rho_1, \rho_2, \theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}, \theta_{12}, \theta_{13}, \theta_{14}, \theta_{15}, \theta_{16}, \theta_{17}, \\ &\theta_{18}, \theta_{19}, \theta_{20}, \theta_{21}, \theta_{22}, \theta_{23}, \theta_{24}, \theta_{25}, \theta_{26}, \theta_{27}, \theta_{28}, \theta_{29}, \theta_{30}, \theta_{31}]\end{aligned} \tag{16}$$

for the proposed new mapping.

The three radii indicate the three inner layers (since the outer one is normalized to 1, while the phases have been defined so that the subscripts of each theta corresponds to the alphabet value assigned to that symbol.

	CCSDS Mapping 4+12+20+28	Proposed Mapping 4+12+20+28
Non-linear HPA	-3.47	-3.34
Ideal HPA	-3.57	-3.37

Table XV - Comparison of the PEAQ Using Non-Linear and Ideal HPA Filter

The optimized constellation using the mapping proposed by CCSDS and described in section 4.4.1 is shown in Figure 18(a). Figure 18(b) shows the optimized constellation for the optimized 4+12+20+28 mapping proposed in section 4.4.2 while in Figure 18(c) the optimized 4+12+20+32 proposed constellation is drawn.

In Figure 19(a), for SNR = 10 dB the MSE = 34.28 with a reduction of 67% with regard to the case of non–optimized constellation described in Figure 17(a). The optimization is more evident for low SNR while the two curves converge for higher values.

In Figure 19(b), a comparative analysis at SNR = 10 dB shows that for optimized 4+12+20+28 mapping the MSE =24.1 that is 77% less with regard to the non-optimized case. Similarly, in Figure 19(c), when SNR = 10 dB the average distortion for optimized 4+12+16+32 mapping is 22.76 with a reduction of the 78% with regard to the non-optimized mapping.

Figure 19(d) puts compares the results for the 3 constellations optimized through GA presented in Figure 18(a), (b), and (c), showing once again that placing 28 or 32 symbols on the outer ring yields to similar results.

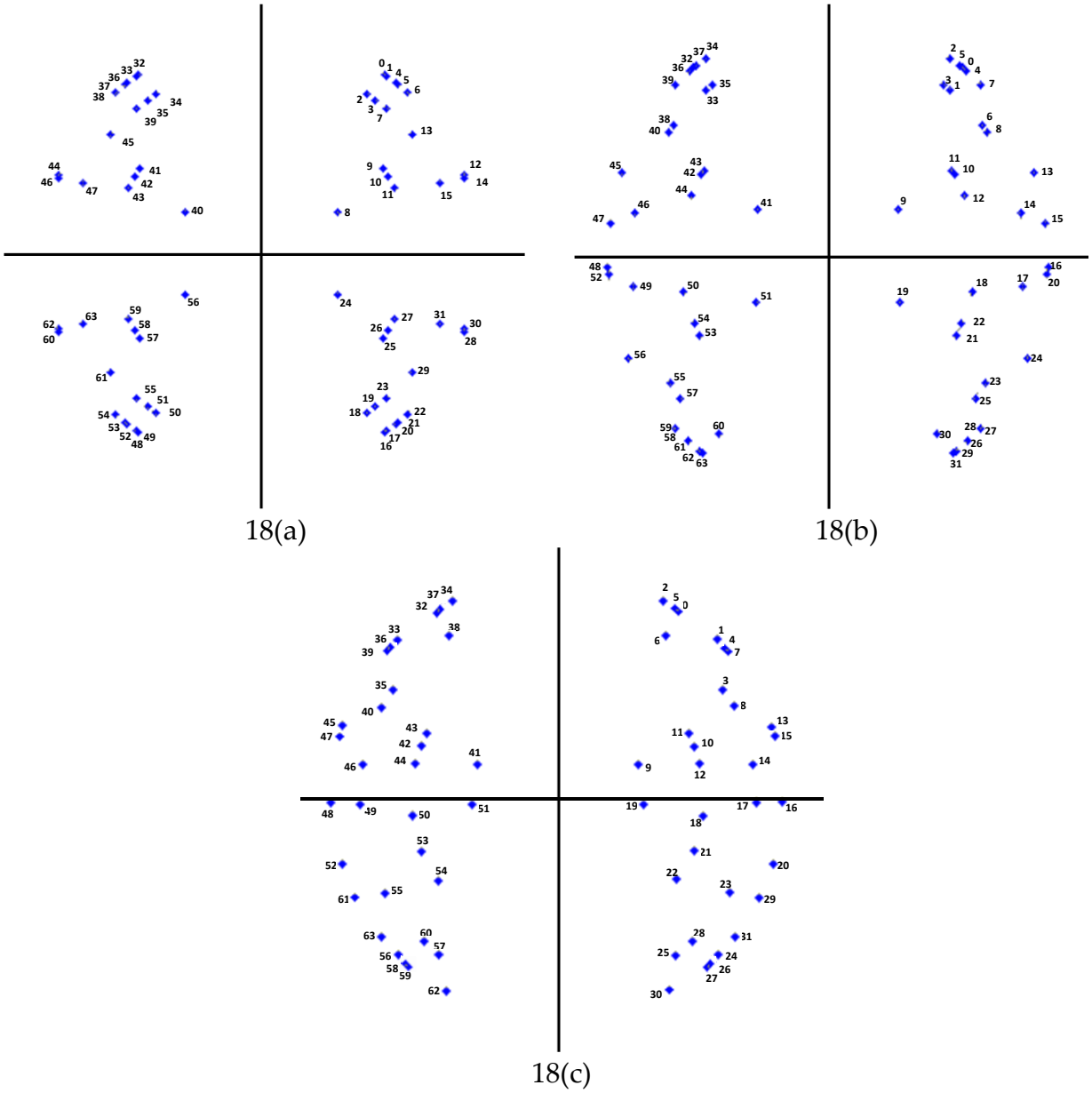


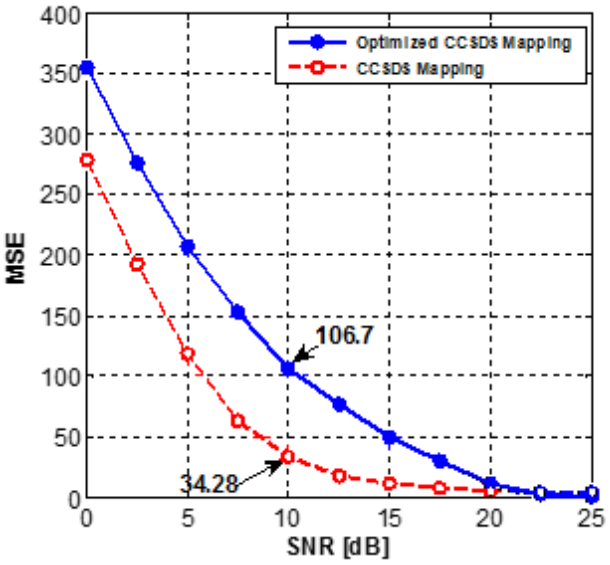
Figure 18 - Optimized constellation for the 4+12+20+28 CCSDS mapping (a), 4+12+20+28 proposed mapping (b), and 4+12+16+32 proposed mapping (c). Symbols are expressed in decimal representation

4.5.3 VALIDATION IN REALISTIC SCENARIO

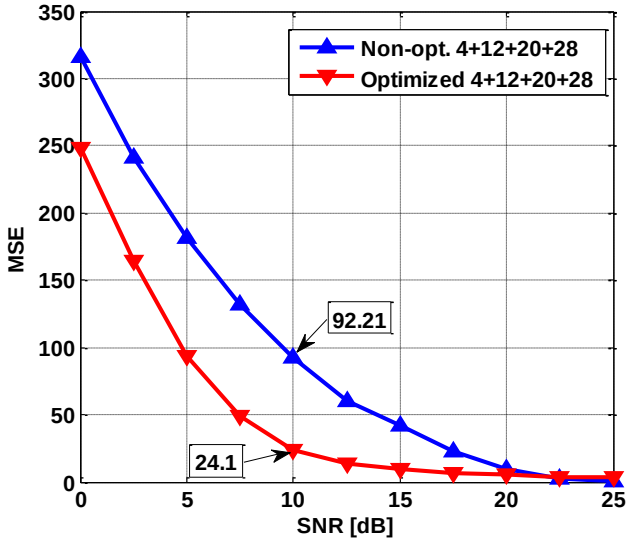
In this section the performance of the proposed optimized constellations over a realistic system model is investigated. The roll-off factor of the SRRC filter is fixed to 0.2 and dynamic pre-distortion have been considered using L = 3 symbols simultaneously for reducing the ISI introduced by the non-linear HPA [37].

Furthermore, to assess the suitability of the proposed approach with real applications. At first a 5 seconds from a stereo audio CD signal sampled at 44.1 KHz coded at 16 bps has been considered. The measurements have been made using the computer measuring system

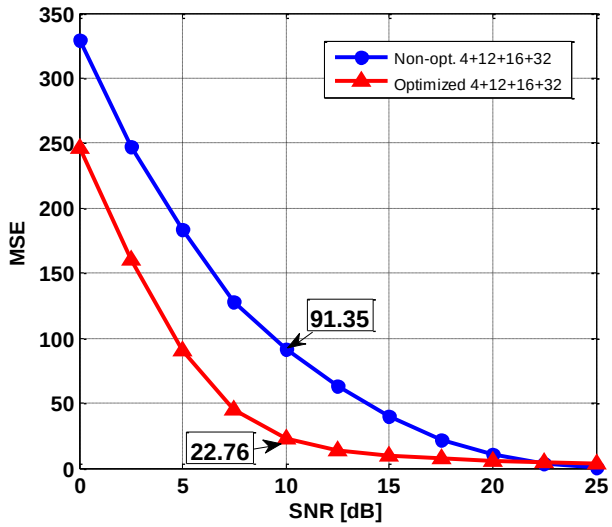
OperaTM comparing them with the referential clip. Again for SNR = 10 dB, the PEAQ measured on the received audio sequence transmitted with the proposed optimized 64-APSK constellation with proposed modified 4+12+20+28 mapping is -3.34, against the conventional 64-APSK with CCSDS mapping which achieved a lower score of -3.47. The obtained PEAQ score is just below the typical performance of low bit rate audio codecs: for the transmission of audio sequences at a rate of 64 Kbit/s, MP3/MPEG-4 codecs achieve PEAQ of around -3.36 [48]. As detailed in Table XV, notice that the same transmission performed with ideal HPA under the simplified hypothesis specified in section 4.2 and used in section 4.5.1 and 4.5.2 produced a score of -3.37 for the proposed optimized 64-APSK constellation with proposed modified 4+12+20+28 mapping and a PEAQ of -3.57 for conventional 64-APSK with CCSDS mapping. In fact, this shows that dynamic pre-distortion can outperform the ideal HPA since dynamic pre- distortion also reduce ISI while the ideal HPA is a memory less device that cannot reduce ISI. It is worth noting also that the gain obtained when dynamic pre-compensation is used is higher for conventional 64-APSK with CCSDS mapping against the proposed optimized 64-APSK constellation with proposed modified 4+12+20+28 mapping. Indeed, in the proposed optimized constellation the symbols tend to be approaching the outer ring, thus being more likely to cause performance degradation due to ISI. To further assess the correspondence between MSE and the subjective perception the transmission of a standard test grey-level image "Lena" of size 512x512 coded at 8 bpp has been considered. Figure 20(a) shows the original image and the images received in case of a transmission for SNR = 5 dB. Figure 20(b) is obtained by adopting the proposed optimized 64-APSK constellation with proposed modified 4+12+20+28 mapping, whereas Figure 20(c) is obtained by using the conventional 64-APSK with CCSDS mapping.



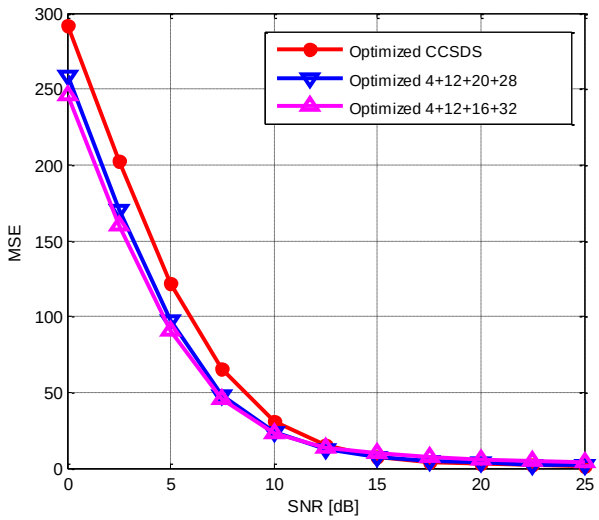
19(a)



19(b)



19(c)



19(d)

Figure 19 - Performance comparison between optimized and non-optimized CCSDS mapping (a), performance comparison between optimized and non-optimized 4+12+20+28 proposed mappings (b), MSE results obtained optimizing the CCSDS and the 4+12+16+32 proposed mappings (c), and overall comparison of the presented constellation with GA optimization (d)

It can be observed that in the second case the image is completely incomprehensible, whilst in the first case it can be easily understood in spite of the errors due to the highly noisy channel.

(a) (b) (c)

Figure 20 - Original "Lena" 512x512, 8bpp (a), "Lena" transmitted for SNR = 5dB, with the proposed modified 64-APSK 4+12+20+28 mapping (b), and conventional 64-APSK with CCSDS mapping (c).

4.6 CONCLUSIONS

In this chapter, the use of modified constellations and mapping for transmission in the satellite broadcasting scenario has been proposed for 64-APSK to achieve unequal error

protection of data which exhibits different sensitivity to channel errors (e.g., multimedia signals). In particular, new mappings derivation starting from conventional CCSDS mapping together with constellation optimization using genetic algorithms have been presented. Based on the assumption that errors in most significant bits (MSB) produce higher distortion than errors in the least significant bits (LSB), the proposed optimization aggregates symbols that have lower inter-symbol distortion and separates symbols which have higher inter-symbol distortion. Found results demonstrate that it is possible to improve the performance of conventional 64-APSK in terms of average distortion between transmitted and received data (i.e., MSE) using modified mappings where some symmetries are dropped and the symbols on the same circles are not equally spaced anymore, thus producing asymmetric constellations. Tests using real multimedia signals in realistic transmission scenario allow assessing the suitability of the proposed scheme for digital multimedia broadcasting.

PART II - HETEROGENEOUS WIRELESS NETWORKS OPTIMIZATION

5 Adaptive Real-time Multi-user Access Network Selection Algorithm for Load-balancing over Heterogeneous Wireless Networks

The coexistence of several wireless technologies gives users a wide connectivity choice. However, in this heterogeneous context with the even growth of multimedia data demand and continuous amendments of network conditions, the best network selection represents a crucial issue. In this chapter the authors propose a novel Adaptive Real-time Multi-user Access Network Selection (ARMANS) load balancing algorithm, taking into account not only the real-time global traffic load on each network, but also considering the different classes of traffic. A novel Prioritized Adaptive Real-time Multi-user Access Network Selection (P-ARMANS) load balancing algorithm over heterogeneous wireless networks (HWNs), 4G-LTE and 802.11n, is also proposed. Markov decision process (MDP)-based multi-agent framework was introduced to perform the bandwidth allocation on users with different typology and priority and also considering and classifying the type of traffic involved in a multimedia data user demand. The proposed approach performs the available bandwidth on users with different screen resolution and priority. The simulation results show that the proposed solutions improve both QoS and load balancing in comparison with the case when a classic network selection with no traffic type load balancing is employed. Moreover, P-ARMANS framework ensures high QoS distributing different type of traffic load among typical and business users compared to a classic no-priority model.

5.1 INTRODUCTION

The increasing number of wireless devices that are accessing mobile networks worldwide is one of the major contributors to global mobile traffic growth. Each year several new devices with different form factors and increased capabilities and degree of innovation are introduced to the market. Nearly half a billion (563 million) network connected mobile devices were added in 2015 alone. Global mobile devices and connections in 2015 grew to 7.9 billion, up from 7.3 billion in 2014 and they are expected to reach 11.5 billion by 2019. The overall mobile data traffic is expected to grow to 24.3 exabytes per month by 2019,

nearly a tenfold increase over 2014 [53] due also to the extensive exchange of high quality multimedia content [54].

Furthermore, according to the same study reported by Cisco, mobile offload exceeded cellular traffic for the first time in 2015. In particular, 51% percent of total mobile data traffic was offloaded onto the fixed network through Wi-Fi or femtocell in 2015 and in total, 3.9 exabytes of mobile data traffic were offloaded onto the fixed network each month.

Moreover, at present various access technologies, such as UMTS [55], LTE [56] and IEEE 802.11 [57], coexist in the deployed networks of existing network operators. In this context, it is expected that the operators' limited network resources should be fully utilized to provide best multimedia content to users at high Quality of Service (QoS) levels, while also securing the best profits.

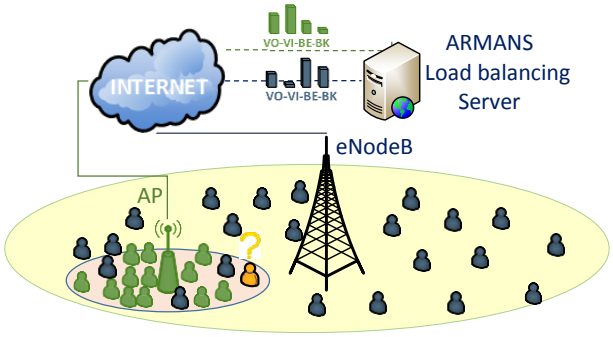


Figure 21 - Load balancing ARMANS server in heterogeneous network

Currently, one way to achieve this complex goal is load balancing, through which the spatial varied traffic load can be scheduled dynamically among cells to avoid traffic blocking in highly loaded areas, while the resources in low loaded areas are underutilized. However, the recent evolution of access network environment, the scope of load balancing is being extended to heterogeneous networks, where different radio access technologies (RAT) coexist [58] [59] [60].

This chapter proposes a novel Adaptive Real-time Multi-user Access Network Selection (ARMANS) algorithm for load balancing for multimedia streaming over heterogeneous networks, as illustrated in Figure 21. The proposed ARMANS selects dynamically the best candidate network considering the mobile station traffic typology request and real-time traffic type. Based on a novel priority-based approach, ARMANS also considers reallocating network resources in order to balance high quality user satisfaction with several device-related characteristics and running application type. The ARMANS mechanism ensures the best multimedia experience to the users and obtains very good load balancing for

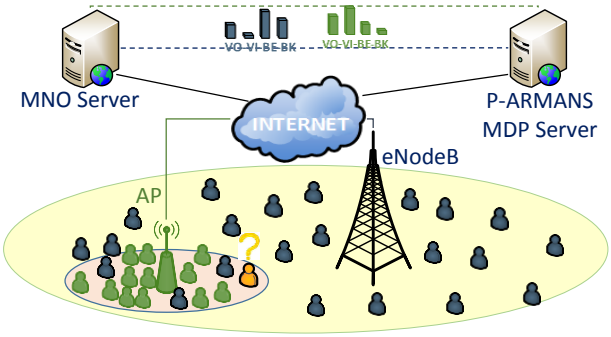


Figure 22 - Load balancing P-ARMANS server in heterogeneous network

multimedia deliveries over heterogeneous networks. The ARMANS key performance benefits are as follows: better fairness in terms of load balancing, higher QoS than load balancing with no consideration for traffic type, and prioritization of access network selection in heterogeneous networks.

Considering data network mixed traffic, it is evident that it will be partially high and partially low priority. High-priority data enforce specific requirements due emergency (e.g., wireless or wired public safety network control data processing) or due reconstruction constraints (e.g., real time and streaming applications). On the other hand, so-called low-priority data can bear longer latency and delays.

Nowadays, within the heterogeneous wireless networks (HWNs), there is the need to consider not only the data priority but also the type of user. Taking into account heterogeneous communication and mixed users, it should be notice that mobile network operators offer different traffic planes; standard individuals or business agreements are a non-exhaustive list of these options.

Optimal network selection according to user type of traffic, load balancing, user priority and screen resolution is a challenging task, which can be done by multiple attributes decision-making (MADM) methods [61], [62]. MDP-based optimization process has been used in this study with the final goal to ensure load balancing and user QoS among user and network conditions with different priorities. This work also proposes a novel Prioritized-Adaptive Real-time Multi-user Access Network Selection (P-ARMANS) MDP-based algorithm for load balancing for multimedia streaming over HWNs, as illustrated in Figure 22.

5.2 Related Works

5.2.1 Load balancing state of the art

In [63],[64],[65] the authors propose reputation-based load-balancing network selection strategies for heterogeneous wireless environments. These reputation-based mechanisms select the most appropriate set of networks for the mobile user and a load balancing mechanism to distribute the traffic load among the networks by making use of the different protocols, including TCP, UDP and Multipath TCP (MPTCP).

In [66], Luo et al. propose a Strenghten-Gray Relative Analysis (S-GRA) access selection algorithm based on load balance in HWNs. This algorithm not only takes network state into account, but also includes some load balance process when guarantee low block probability in a wireless local area network (WLAN) and UMTS scenario. In [67] Militano et al. propose a radio resource management approach in order to increase the aggregate data rate of the cell, while also maintaining the conventional multicast scheme short-term fairness. In [68], Yang et al. present a novel load balancing scheme conceived for the OFDMA LTE cellular and Wi-Fi coexisted network. The load balance entity manages the user equipment (UE) load in coexistent network dynamically. A load balancing algorithm is developed to balance the network load measured by the ratio of the average Wi-Fi load to the LTE cell load as well as the Jain's fairness index [69]. In [70], the authors extend the concept of Load Balancing to the Spatial domain developing two approaches – Network Load Balancing and Single-Carrier Multi-Link - for Spatial Load Balancing. Both these methods apply when the device has more than one candidate server and determine the server(s) using not only the channel quality from the server to the device but also the current load on each server.

However, most of the previous works do not consider the characteristics of application running at the end-user side and the type of traffic over both IEEE 802.11 and LTE networks. Therefore, this work proposes a novel Adaptive Real-time Multi-user Access Network Selection (ARMANS) to determine the candidate network that best supports the QoS requirements of the transmitted type of traffic.

5.2.2 Prioritization and radio access network selection

In [71] Wang et al., proposed a priority-based cell selection scheme that divides the total cells into different types, decide their priority order, and selects the type of cells with the highest priority. With this method, when the user moves across cell border, the scheme has

only a small probability to trigger cell selection decision, hence decreasing the number of triggers of cell selection decisions.

Liu et al., presented a random-access algorithm considering data network carrying mixed traffic, partially low and partially high-priority [72]. The algorithm consists of two dynamically coupled window algorithms, one for the high and one for the low-priority packets. The optimization consists of throughput maximization under no specific delay constraints, and throughput maximization subject to expected delay constraints for the high-priority traffic. Each user continuously observes the channel feedback, from the time he generates a packet to the time that this packet is successfully transmitted.

In [73] the authors deal with a dynamic channel selection by each one of the high priority users, towards the reduction of their accessing delays. The considered environment contains a number of transmission channels among which each user may select to direct his/her transmission. Time is divided into slots of length equal to the duration of a packet, and the starting instants of the slots are identical in all channels. The accessing of the high priority users is accelerated significantly at the expense of some throughput reduction. Performance analysis were performed in terms of throughputs and reduced delays.

Several studies were carried out for access network selection and multimedia delivery content. Some of these are related to 802.11 WLANs, others involved heterogeneous networks. The IEEE 802.11e protocol enables QoS differentiation between different traffic types, it requires MAC layer support and assigns traffic with static priority. In order to solve this limitation Yuan et al., proposed an intelligent Prioritized Adaptive Scheme (iPAS) to provide QoS differentiation for heterogeneous multimedia delivery over wireless networks (WNs) [74]. iPAS assigns dynamic priorities to various streams and determines their bandwidth share by employing a probabilistic approach-which makes use of stereotypes. The priority level of individual streams in iPAS is variable and considers service types and network delivery QoS parameters (i.e. delay, jitter, and packet loss rate).

5.2.3 Markov Decision Process in Access Network Selection

MDP is one of well-known MADM-based methods used for handoff decision in heterogeneous environment. MADM procedures are characterized by multi-criteria decision problems and are widely used for solving several issues including the network selection problem in HWNs.

In [75] Yang proposes a Semi-Markov Decision Process (SMDP) for active state control of a HWNs. The author deals with the problem of a large number of different base station in idle state and proposes an active state control algorithm based on SMDP with coverage and signal to noise ratio (SNR) constraints. The proposed algorithm properly controls the active

state depending on traffic densities without increasing the number of handovers excessively while providing average user perceived rate (UPR) in a more power efficient way than a conventional algorithm. In this study, active state control has been formulated into a sequential decision problem and one of the most efficient methods for solving sequential decision problem is to exploit the framework of MDP.

In [76] Jakimoski et al., developed an analytical framework using the Markov chain to determine the all possible states of the mobile terminals equipped with three interfaces for WLAN, WMAN, and WWAN networks. Such HWN system is modeled as a Markov chain with the appropriate transitions among the all-possible states of the mobile terminals. The proposed framework aims the network selection and vertical handover decision algorithm for the three access technologies in order to optimize the performances of the heterogeneous mobile and HWNs.

The coexistence of more RAT in the same geographical area offers various advantages. In order to maximize the generated revenue while satisfying the customers' increasing demand the authors in [77] studied MDP-based RAT selection in a cellular/WLAN heterogeneous network where the maximization of the revenue is considered as objective. Performance is evaluated in comparison with two other policies, namely cellular-first policy and load balancing policy.

However, most of the previous works introduced priority-based cell selection scheme or occasional accessing requests by public safety-high-priority users. In some cases, QoS differentiation for heterogeneous multimedia delivery over WNs has been reached assigning dynamic priorities to various streams. Moreover, previous studies do not consider the characteristics of application running at the end-user side and the type of traffic over both IEEE 802.11 and LTE networks.

Therefore, this work proposes a novel Prioritized-Adaptive Real-time Multi-user Access Network Selection (P-ARMANS) to determine the candidate network that best supports the QoS requirements of the transmitted type of traffic. Users are classified by device characteristics, mobility, user-to-user priority and screen resolution-to-resolution priority methods. P-ARMANS makes use of MDP to manipulate the optimal transition among the all possible states of the mobile configuration equipped with WLAN and LTE radio interfaces. According to network load, MDP handles the available throughput to assign to each user classified with different attributes.

5.3 ARMANS AND P-ARMANS SYSTEM MODELS

5.3.1 FRAMEWORK OF THE PROPOSED ARMANS

The ARMANS scheduling framework in Figure 23 consists of server-side, client-side and Mobile Network Operator-side (MNO) components. At the mobile client-side, the UEs are characterized by both Wi-Fi and LTE compatible multi-radio devices and they are connected to either an AP or to an Evolved Node B (eNodeB). The UEs integrates several functional blocks, as follows. The Network Discovery block, which provides the user a list of available wireless connectivity options based on their coverage footprint. The Device Characteristics block details information about screen resolution. The QoS Monitor block sends to the server the quality of the received services in terms of Peak Signal to Noise Ratio (PSNR), packet delay and packet loss rate. Finally, the Application Type block provides to the server the type of traffic involved in a current session, with the QoS associated to voice, video, best-effort and background traffic. Four Wi-Fi classes (i.e., VO – voice, VI – video, BE – best-effort, BK – background) [78] and nine QoS Class Identifier (QCI) [79] for LTE have been analysed. A virtual traffic classes-based grouping giving four choices both for WLAN and LTE traffic type is proposed. This information is periodically sent at every Transmission Time Interval (TTI). The MNO-side integrates the real-time network characteristics transmitted by the network operator to the ARMANS server. Both LTE (i.e., eNodeB) and Wi-Fi (i.e., AP) transmitters are geolocalized and communicate to the ARMANS server their position and respective network IDs (i.e., eNodeB-ID and AP-ID, respectively). Each transmitter node also communicates the total amount of traffic load and corresponding distribution within different type of traffic. The ARMANS server-side gathers information from both mobile client-side and MNO-side. The ARMANS algorithm located at the ARMANS server-side will compute a network ranking list based on the estimated quality of the multimedia stream.

The core of the proposed mechanism gives a network ranking and suggests the best network to the UE either according to the traffic load MNO-side and the application type running mobile user-side.

5.3.2 FRAMEWORK OF THE PROPOSED P-ARMANS

As seen in previous section, P-ARMANS scheduling framework in Figure 24 consists of three main blocks: server-side, client-side and Mobile Network Operator (MNO)-side block

The Application and User type-ID block provides to the P-ARMANS MDP-based server the type of traffic involved in a user current session, with the QoS associated to VO, VI, BE and BK traffic. The user type-ID is employed to differentiate the user from custom to business. Each types of traffic required by a business user have higher priority than required by custom user.

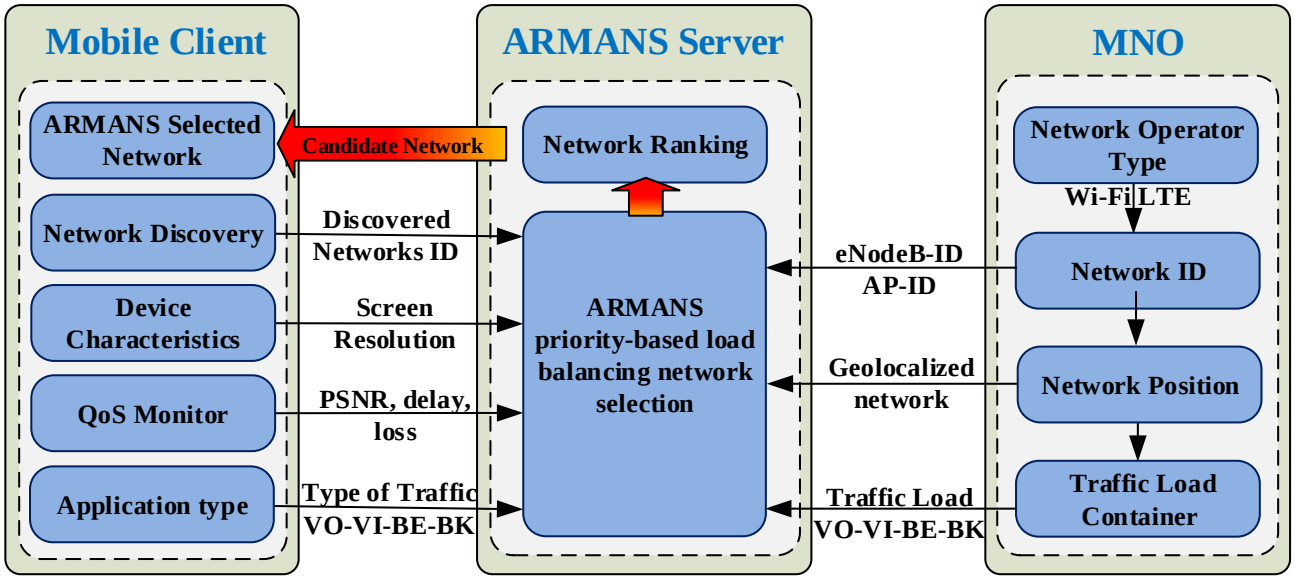


Figure 23 - ARMANS scheduling framework

P-ARMANS load network selection algorithm pseudo code

t=0
WHILE(TRUE)
  FOR each UEs
    FOR each MNO
    calculate average received power $\bar{p}_m^u(t)$
    calculate power standard deviation $\bar{\sigma}_m^u(t)$
    calculate mobile index $v_m^u(t)$
    END FOR
      IF $Th_1 \cdot l_i^{LTE}(t) < l_i^{WiFi}(t)$ **OR** $l_i^{WiFi}(t) < Th_2 \cdot l_i^{LTE}(t)$
        calculate access index $AI_m^u(t)$
        IF $AI_m^u(t) \leq 1$, then *LTE*
        ELSE *Wi-Fi*
  END FOR
  U^{LTE}++ *or* U^{WiFi}--
  U^{LTE}-- *or* U^{WiFi}++
END WHILE

UEs communicate to P-ARMANS MDP server their position and speed in order to assign the best network when moving through different networks.

The P-**ARMANS server-side** gathers information from both mobile client-side and MNO-side. The E-ARMANS algorithm located at the E-ARMANS server-side will compute a network-ranking list based on the calculated quality of the multimedia stream. Always best connected devices is the core of the proposed mechanism either according to the traffic load and load balancing MNO-side, energy consumption, and the application type running mobile user-side.

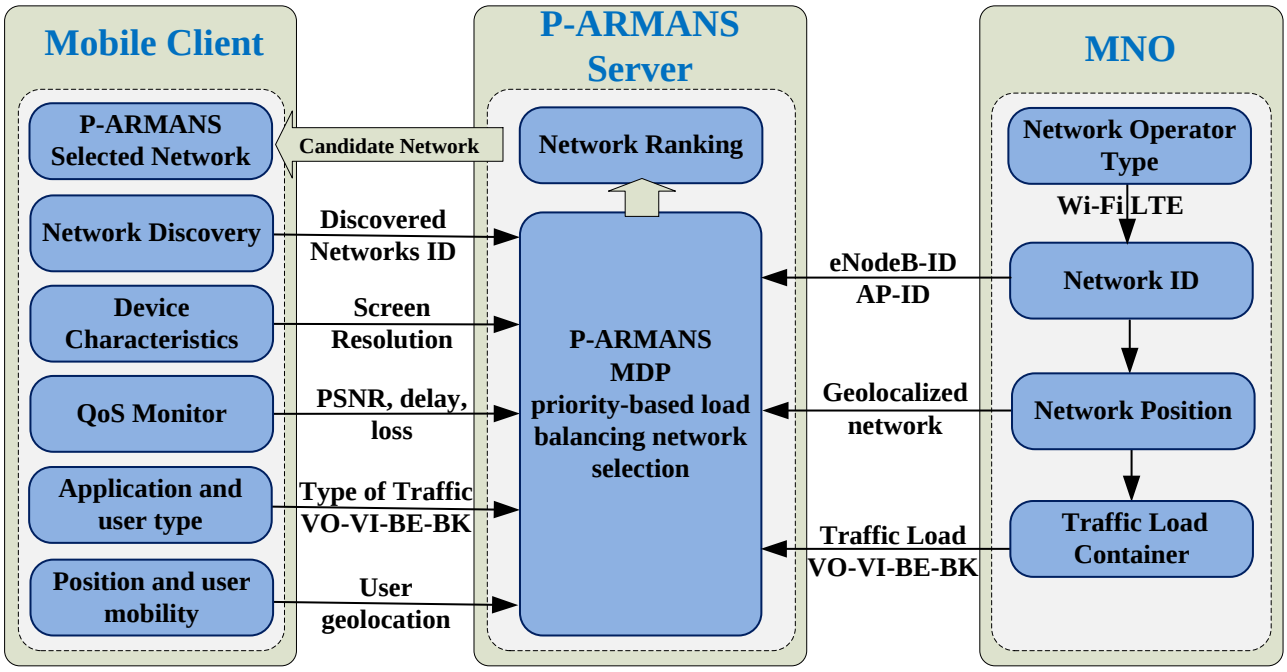


Figure 24 - P-ARMANS scheduling framework

5.4 NETWORK LOAD MODEL

The system model considers *L* LTE cells uniformly distributed in the networks, the LTE base station (BS) eNodeB is located in the cell center, and *W* WiFi APs randomly distributed within the LTE cell. User *UEs* are distributed in each cell randomly and uniformly and are able to access either the LTE BS or WiFi APs.

The LTE traffic load is defined as the ratio of the current total amount of resource blocks scheduled to UEs in a particular cell divided to the total amount of resource blocks available in that cell. In every TTI, each eNodeB will transmit scheduled information to the UEs. The LTE load is defined in equation (17):

$$l_L^{LTE}(t) = \frac{1}{LTE_{r_{MAX}}} \sum_{i=1}^{I} l_i^{LTE}(t) \tag{17}$$

where $l_L^{LTE}(t)$ represents the load of the LTE cell in time slot t which is a normalized positive value between 0 and 1, LTE_{rMAX} is the maximum available resources [Mbps] in LTE cell, $i \in \{1,\ldots,I\}$ defines the type of traffic classification (18) and, $l_i^{LTE}(t)$ represents the allocated resources classified in type of traffic i for LTE cell at the time slot t.

$$i \in \{VI, VO, BE, BK\} \tag{18}$$

Equation (19) defines the WiFi traffic load as the ratio of the current load of UEs access to WiFi AP divided by the maximum available resources those UEs can access via the WiFi AP.

$$l_W^{WiFi}(t) = \frac{1}{WiFi_{r_{MAX}}} \sum_{i=1}^{I} l_i^{WiFi}(t) \tag{19}$$

In this equation, $l_W^{WiFi}(t)$ is a normalized positive value between 0 and 1 and represents the WiFi load in time slot t, $WiFi_{rMAX}$ is the maximum available resources in the WiFi cell [Mbps], $i \in \{1,\ldots,I\}$ defines the type of traffic classification in equation (18) and, $l_i^{WiFi}(t)$ represents the WiFi allocated resources classified in traffic type i for the WiFi cell at the time slot t.

5.4.1 PROBLEM FORMULATION

Considering that new users login to the network requesting a particular type of traffic and old users logout making available new resources, both LTE and WiFi network loads are time-variable. Thus, performing load balancing is an effective method to adjust fairly the distribution of network traffic among cells while maintaining certain quality of service (QoS) levels. In order to describe the degree of fairness there is a need of a mathematical model, index or equation.

Jain's fairness index [80], is used to assess the load balancing for type of traffic classes over heterogeneous networks and to measure the load fairness among L LTE eNodeB and W WiFi APs.

$$J_i(t) = \frac{\left[\sum_{m=1}^{W} l_i^{WiFi}(t) + \sum_{n=1}^{L} l_i^{LTE}(t)\right]^2}{N \cdot \left[\sum_{m=1}^{W} \left(l_i^{WiFi}(t)\right)^2 + \sum_{n=1}^{L} \left(l_i^{LTE}(t)\right)^2\right]} \tag{20}$$

where $J_i(t)$ represents the Jain index for the class of traffic i at the slot time t and $N=W+L$ is the total number of network discovered by the UE under its radius coverage. The large

value of $J_i(t)$ denotes the balanced UEs load among LTE eNodeB and WiFi APs. To reach the best load fairness among type of traffic, the Jain's fairness index is maximized. Nevertheless, according to the system model, the load balancing in heterogeneous network is performed as follows:

$$\begin{aligned} &Th_1 < \frac{l_i^{WiFi}(t)}{l_i^{LTE}(t)} < Th_2 \qquad i \in \{1,\ldots,I\} \\ &0 < Th_1 < 1 < Th_2 \end{aligned} \tag{21}$$

where Th_1 and Th_2 represent load balancing thresholds, and have a variable setting in specific scenarios, such as Th_1 = 0.9 and Th_2 = 1.1, for instance. Equation (5) denotes the degree of load balancing between a LTE eNodeB transmitter and WiFi AP, and the variable load thresholds $Th_1 \cdot l_i^{LTE}(t)$ and $Th_2 \cdot l_i^{LTE}(t)$ represent the tolerated difference between heterogeneous network load.

Another metric to be considered is the WiFi average traffic load compared to LTE traffic load in equation (22):

$$\begin{aligned} r_i^{WiFi}(t) &= \frac{1}{N} \frac{\sum_{i=1}^{I} l_i^{WiFi}(t)}{\sum_{i=1}^{I} l_i^{LTE}(t)} \\ r_i^{LTE}(t) &= 1 - r_i^{WiFi}(t) \end{aligned} \tag{22}$$

According to this equation, it is clear that the load between LTE cell and WiFi AP is balanced when $r_i^{WiFi}(t)$ converges to $1/N$. The traffic type load balance problem formulation is optimized when $J_i(t)$ described in equation (20) is max with the constraints in equation (21). The optimal formula maximizes the Jain's index and have a convergence of $r_i(t)$.

5.4.2 ARMANS AND P-ARMANS PROPOSED LOAD BALANCING SCHEME

In the ARMANS scheme, a load management entity is introduced to manage the traffic type load and obtains load information from the eNodeB and WiFi AP and computes the present cell load state to balance network load dynamically. The ARMANS load management framework obtains current access information from U^{LTE} and U^{WiFi}. UEs in set U^{LTE} have access to the LTE BS and UEs in set U^{WiFi} have access to the WiFi APs.

Given the UEs inputs regarding the application running in the mobile device and the discovered LTE and WiFi networks, the load management entity computes UE access index and determines the access pattern of UEs. The proposed load balancing algorithm is based

on this UE access index. There are different supports for mobility in LTE network and WiFi APs.

The WiFi APs can only support the UEs with low mobility. In this work, the mobile index is proposed to describe the UE mobile state. The load management entity computes the mobile index $v_m^u(t)$ according to the received reference signal strength in each UE. A sliding window method is adopted to obtain the UE average received power with the window size of T. The equation of the average received power and power standard deviation for the user m are as follows:

$$\bar{p}_m(t) = \frac{1}{T}\sum_{d=0}^{T-1} p_m^u(t-d) \tag{23}$$

$$\sigma_m(t) = \frac{1}{T}\sum_{i=0}^{T-1}\left[p_m^u(t-i) - \bar{p}_m^u(t)\right]^2 \tag{24}$$

where, u is the UE number, t denotes the present time slot and $\bar{p}_m^u(t)$ denotes the average value of received power in the latest T time slots from transmitters under its radius coverage.

Mobile index is defined as a function of the power standard deviation.

$$v_m^u(t) = 1 - \frac{1}{1+e^{-\sigma_m^u(t)}} \begin{cases} 0 & \sigma_m^u(t) > \sigma_{Th} \quad \textit{high mobility} \\ 1 & \sigma_m^u(t) < \sigma_{Th} \quad \textit{low mobility} \end{cases} \tag{25}$$

where $v_m^u(t)$ is a positive value between 0 and 1. The mobile index close to 0 indicates the UE has a high mobility, otherwise the mobility user close to 1 corresponds to low user mobility.

The access index (AI) is proposed to measure whether the UE within the heterogeneous coverage area has good access to the WiFi AP or eNodeB LTE. The equation of the AI is as follows:

$$AI_m^u(t) = v_m^u(t) \cdot \frac{\sum_{i=1}^{I} l_i^{LTE}(t)}{\sum_{i=1}^{I} l_i^{WiFi}(t)} \qquad \begin{cases} \le 1 & LTE \\ otherwise & WiFi \end{cases} \tag{26}$$

In equation (26), higher access index indicates UE u is better to access the WiFi, otherwise, the UE should access the LTE network.

In order to obtain optimal $J_i(t)$ and $r(t)$, the ARMANS algorithm is adopted to adjust the traffic type load in existent LTE and WiFi networks. The WiFi APs load set UWiFi consists of the WiFi loads that satisfy $l_i^{WiFi}(t) < \mathrm{Th}_2 \cdot l_i^{LTE}(t)$, which indicates the WiFi AP load is relatively lower than that of the LTE cell. The LTE set U^{LTE} is comprised of the loads that

satisfy $Th_1 \cdot l_i^{LTE}(t) < l_i^{WiFi}(t)$, which indicates the WiFi load is relatively higher than that of the LTE cell.

If the set U^{WiFi} is not empty, choose the lowest WiFi type of traffic load from U^{WiFi}, and mark it as β_{WiFi_min}. The load balancing between LTE cell and WiFi, β_{WiFi_min}, corresponds to choose the UE with the maximum access index AI from set U^{WiFi} of WiFi β_{min}, and mark it as $u_{r\ max}$.

$$\begin{aligned} \beta_{WiFi\_\min} &= \arg\min\ l_i^{WiFi}(t) \\ u_{r\_\max} &= \arg\max_m\ AI_m^u(t) \qquad u \in UEs \end{aligned} \tag{27}$$

When the UE $u_{r\ max}$ is determined, the corresponding data traffic hands over to WiFi β_{WiFi_min}, and update the sets U^{LTE} and U^{WiFi} of WiFi β_{WiFi_min} :

$$\begin{aligned} U^{LTE} &= U^{LTE} - u_{r\_\max} \\ U^{WiFi} &= U^{WiFi} + u_{r\_\max} \end{aligned} \tag{28}$$

If the set U^{LTE} is not empty, the algorithm determines the lowest LTE type of traffic load from U^{LTE}, and marks it as β_{LTE_min}. The load balancing between LTE β_{LTE_min} and WiFi consists of the choice of the UE with the maximum access index AI from set U^{LTE} of LTE β_{LTE_min}, and mark it as u_{r_max}.

$$\begin{aligned} \beta_{LTE\_\min} &= \arg\min\ l_i^{LTE}(t) \\ u_{r\_\max} &= \arg\max_m\ AI_m^u(t) \qquad u \in UEs \end{aligned} \tag{29}$$

Once defined the UE u_{r_max}, its data traffic hands over to LTE β_{LTE_min}, and update the sets U^{LTE} and U^{WiFi} of LTE β_{LTE_min}:

$$\begin{aligned} U^{LTE} &= U^{LTE} + u_{r\_\max} \\ U^{WiFi} &= U^{WiFi} - u_{r\_\max} \end{aligned} \tag{30}$$

The load information of the LTE cell and WiFi APs is updated according to the algorithm above, as well as the sets U^{LTE} and U^{WiFi}. Whether the two sets are both empty, which indicates the load in LTE cell and WiFi APs are balanced, the load balance algorithm stopped; otherwise, restart the load balance algorithm. ARMANS evaluates the total amount of traffic involved in each network and the relative load to each type of traffic. ARMANS algorithm suggests the best connection minimizing delay, and packet loss rate, and maximizing the throughput, PSNR and load balancing over heterogeneous network.

5.5 ARMANS SIMULATION PLAN

Four types of multimedia traffic were used for transmission for both IEEE 802.11n and LTE [81]: voice, video streaming, best-effort, and background. In Table XVI the characteristics of the four traffic classes are shown.

The MNO-side integrates the real-time network characteristics transmitted by the network operator to the ARMANS server. Both LTE (i.e., eNodeB) and Wi-Fi (i.e., AP) transmitters are geolocalized and communicate their respective network ID (i.e., eNodeB-ID and AP-ID, respectively), and position to the ARMANS server. Each transmitter node also communicates the total amount of traffic load and corresponding distribution within different type of traffic. The ARMANS server-side gathers information from both mobile client-side and MNO-side. The ARMANS algorithm located at the ARMANS server-side will compute a network ranking list based on the estimated quality of the multimedia stream.

The core of the proposed mechanism gives a network ranking and suggests the best network to the UE either according to the traffic load MNO-side and the application type running mobile user-side.

The ITU-T G.711 [82] audio codec is one of the most deployed in voice calling applications whereas the H.264/AAC codec [83] video traffic adopted is the is the most popular video codec for IP-based networks. The encoding data-rate of voice data is set to typical values specified in the standard (i.e., 64 kbps and 128 kbps), and. several data rate were reported in Table XVII [84]. ARMANS has been evaluated by using the NS-3 network simulator [85]. The simulation topology shown in Figure 21 includes one ARMANS server communicating with N clients over LTE eNodeB and WiFi AP wireless networks. User in orange are connected to WiFi AP whereas user in dark blue are connected to LTE eNodeB. UEs under WiFi and LTE radius coverage could be access both networks. When a new user (i.e., user in yellow) starts its traffic, ARMANS selects the best network. For all the tests, it was configured that the number of stations transmitting each traffic type is the same: one LTE BTS and a WLAN AP.

The number of mobile stations has been progressively increased from 20 to 170 in steps of 1 every 2 seconds in order to collect the statistics under stable conditions, and increasing the overall offered load.

As the number of mobile stations increased, the corresponding normalized offered load has increased from 15.62 % to 156 % (both channels are overloaded). The experimental time duration was set to 300 seconds. Specifically, the normalized offered load achieved 100 % when the number of mobile stations exceeded 130 at around 160 seconds. The starting point

scenario consists of 20 users making HD video streaming allocated as follows: 6, 12, 1 and 1 users filling a 5.76, 2.56, 1.60 and 1.28 Mbps, respectively. The starting total amount of traffic is equal to 93.76 Mbps unequally distributed between LTE and WLAN cells.

	VO	VI	BE	BK
Traffic	Opus or G.711	H.264/ AAC	Pareto distribution traffic model	Pareto distribution traffic model
Protocol	RTP/ UDP/IP	RTP/ UDP/IP	TCP/IP	RTP/ UDP/IP
Encoding Data-rate	TABLE XVII	TABLE XVII	128 Kbps	100 Kbps
Packet size	100 Bytes	1024 Bytes	512 Bytes	512 Bytes

Table XVI - Type of Traffic Characteristics

Name	Resolution and Priority	Link (Mbps)	Bitrate (Mbps)	Video (kbps)	Audio (kbps)
1080p HQ	1440x1080	9.0	5.76	5632	128
1080p	(P1)	6.0	3.84	3712	
720p HQ	960x720	4.0	2.56	2432	128
720p	(P2)	3.0	1.92	1856	64
576p HQ	768x576	2.5	1.60	1536	64
576p	(P3)	2.3	1.47	1408	
480p HQ	640x480	2.0	1.28	1216	64
480p	(P4)	1.5	0.96	896	

Table XVII - Encoding video data rate

5.6 P-ARMANS MDP-BASED PRIORITY MODEL

5.6.1 MARKOV DECISION PROCESS

A **Markov process** (MP) is basically a **discrete-valued** stochastic process in which the state space of possible values of the Markov chain is finite or countable. The past history of the MP is irrelevant when the current system state is known: all information about the past and present that would be useful in saying something about the future is contained in the current state.

A MDP is a Markov chain (MC) that includes an agent that makes decisions that affect the evolution of the system over time [86]. MDP is one of well-known frameworks for decision-making employed for optimization problems using dynamic optimization. This method is used for solving complex problems breaking it in a set of sub problems. Hence, storing their solutions using a memory-based data structure. Thanks to this technique, when

the same sub problem occurs next time, the solution only consists on look at the previous stored one.

A Markov decision process provides a widely used mathematical framework for modeling decision-making in a stochastic environment. The decision (also known as the action) taken at time t affects the evolution of the future system. The goal of the decision maker is to choose a sequence of actions to optimize a predetermined criterion. For the purposes of this work, we assume that the decisions are made at discrete times $t \in Z^+$.

At each decision time *t*, the system occupies a *state*. We denote the set of all possible states by a finite set X. At each time t, the decision maker choses a decision from
a finite set denoted by U.

Formally, a MDP is a tuple defined as follows:

$$MDP =< S, A, T, R, \gamma > \tag{31}$$

where S is a finite set of states, $s \in$ S is the current state, known to users. In this work, the state includes the current time and the requested throughput according to maximum screen resolution supported by the device. A is a finite set of actions, where $a \in A$ is the action taken based on the current state. In this work we consider as actions whether to execute a transition toward another state or remain idle. T is a state transition probabilities matrix (i.e. the probability that action a in the state s at time t will lead to state s' at time $t+1$. As defined in equation (32) for stochastic actions it represents the distribution $P(s' \mid s, a)$:

$$\begin{gathered} T_{ss'}^{a} = P[S_{t+1} = s' | S_t = s, A_t = a] \\ T: S \times A \rightarrow Prob(S) \end{gathered} \tag{32}$$

It is important note that this transition is partly random (e.g. due to the random arrival of video streaming) and partly under control since it is based on action a.
R is a reward function (given a certain state, and possibly action) as defined in equation (33):

$$R_s^a = E[R_{t+1} | S_t = s, A_t = a] \tag{33}$$

and γ is a discount factor which represents the difference in importance between future rewards and present rewards. It is used to reduce the importance of the of future rewards, as defined in equation (34):

$$\gamma \in [0,1] \tag{34}$$

The main core problem of MDP is finding a policy for the decision maker, i.e., a function π that specifies the action $\pi(s)$ the decision maker will choose when in state s.

5.6.2 MDP STATES

P-ARMANS algorithm MDP-based considers eight states, $s_i \in S$ with $i=0,\ldots,7$. Three states (i.e., $i=1,\ldots,3$) identify the maximum received data rate by each user during multimedia VI streaming or real-time contents. The finite set of states S considers three resolution standards (i.e., $i=1,\ldots,3$), the high definition (HD) state (i.e., i=4) and the disconnected state (i.e., i=0) used to identify when a user returns all allocated resources making them available to other users. The rest of the states represents the VO (i.e., $i=5$), BE (i.e., $i=6$) and BK (i.e., $i=7$) user requests. The complete MDP scheme is shown in Figure 25 and the set of states is defined as follows:

- S_0 represents the "zero state". When all sessions, even the BK sessions are terminated (e.g., the user has no network access or switches off the mobile device), all the allocated resources are returned, and become available to new users (disconnected state);
- S_1 is the state associated with mobile users holding ultra high definition screen devices (4K UHD) 2160p (e.g. 3840x2160 pixels), and requesting 4K UHD VI content.
- S_2 is a state in which users with mobile devices with Quad HD (2K) 1440p screen resolution (i.e., 2560x1440 pixels) requesting VI content are placed;
- S_3 is is a state which corresponds to users with full high definition (FHDi) 1080p (i.e., 1920x1080 pixels) mobile device screen resolution accessing VI content;
- S_4 is is a state associated with high definition (HD) 720p, (i.e., 1280x720 pixels) VI content delivery;
- S_i with $i=5,\ldots,7$ are states in which users request VO, BE and BK contents, respectively.

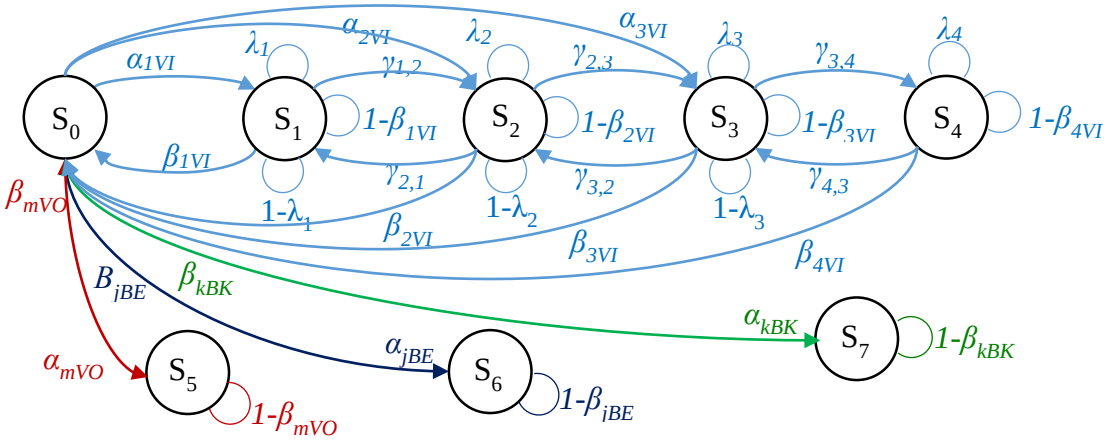


Figure 25. The policy π of the MDP algorithm

5.6.3 MDP ACTIONS

The actions $a \in A$ involved in the P-ARMANS MDP algorithm represent transitions from one state to another when various events affect the delivery process e.g. network overload occurs, a user leaves the network, a user changes the desired content. Transitions from one state to another are bidirectional with the unique VI content constraint that a transition from a current state to better one is limited to the native device screen resolutions. For instance, during a VI session a user with a 2K resolution could not make a transition to a UHD state. However, in overloaded network conditions, a 4K user could move to the 2K state and subsequently return to the previous state when enough resources could be again re-allocated. For all actions and VI resolutions we consider there is a solution to generate the data rate associated with the next state following the transition. The set of actions is defined as follows:

- α_{iVI} represent the VI user request for VI actions, from the state s_0 to the state s_i with i=1,…,3 and according to the native screen resolution device.
- β_{iVI} with i=1,…,4 represent the VI resource release actions. The β_{iVI} actions represent the transition from the current VI *i*-state to the s_0 state. Furthermore, 1-β_{iVI} represents the action to remain in the same state when network will update in terms of user and traffic requests.
- λ_i with i=1,…,4 represent the actions to stay at the same state decreasing the available VI data rate from a maximum to minimum available.
- $\gamma_{i,j}$ represents the action to move from state *i* to state *j*. In particular, we can identify two γ_i actions. $\gamma_{i,i+1}$ identifies the action to move from one i-state to the next *i+1* state. This action is verified when the available resources are not able to satisfy current connected users. The MDP algorithm shares the available resources among users decreasing the data rate while ensuring an acceptable QoS. On the other hand, the action $\gamma_{i,i-1}=1-\gamma i,i+1$ characterizes the user action (without screen resolution constraint) to move from i-state to the previous i-1 state with higher data rate provided.

Actions β and γ can be undertaken only by a certain category of users in particular conditions as indicated in section 5.6.4

Apart from the VI-related actions, this paper considers the user request of other types of traffic and consequently the requests for VO, BE and BK traffic have associated actions α_{mVO}, α_{jBE}, and α_{kBK}, respectively. Furthermore, β_{mVO}, β_{jBE}, β_{kBK}, represent the resources release actions and VO, BE and BK traffic. On the other hand, 1-β_{mVO}, 1-β_{jBE} and 1-β_{kBK}, represents di actions to remain in the same state, related to the indicated traffic type, rspectively.

Depending on device characteristics, all these actions can coexist with each other at the same time. Hence, each user could select a BE download (i.e., α_{jBE} action) together with a VI request (α_{iVI} action). Furthermore, the VI actions and following network connections have priority over other actions. All actions and connection were dragged by VI request and P-ARMANS MDP algorithm optimization. This means that a user could start a BK action and relative access network connection but at the same time it is not possible a VI action and connection to a different network. When P-ARMANS selects the best network for a VI session, VO, BE and BK type of traffic are moved to the VI connection.

5.6.4 User Priority Model

Nowadays, users are classified not only according to their device, but also considering their network connection type (i.e., *payment free* or *payment due* networks). In the first case, all users are consider equal, but for the case of *payment due* connectivity, a wide range of options, MNOs offer. This study we only consider two users' profile: typical (TU) and business user (BU). P-ARMANS algorithm manages and shares the available resources in order to ensure user receive content at QoS levels according to user profile priority. In this context BU has higher priority than TU.

5.6.5 MDP States Transition Probability

T represents the probability the action a in the state s at time t will lead to state s' at time t+1 as defined in equation (16) and represented by the distribution P(s' | s, a). Some of these probabilities are not defined a priori but depend on several factors. In this study, we assumed that a generic connection is initialized with a α_{nBK}, action with a probability $P(\alpha_{nBK})=1$. The probabilities of α_{mVO}, α_{jBE} and α_{iVI} actions are defined according to the proposed scenario and the percentage of VO, BE and VI requests, with the constraint $P(\alpha_{mVO})+ P(\alpha_{jBE})+ P(\alpha_{kVI})=1$. The probabilities $P(\alpha_{mVO})$ and $P(\alpha_{jBE})$ are limited to a very small amount of traffic load and have a small influence in comparison with α_{iVI} actions, hence $P(\alpha_{iVI}) >> P(\alpha_{mVO})+P(\alpha_{jBE})$.

Particular attention have α_{iVI} actions. Firstly, whenever a user discovers a single network, P-ARMANS algorithm is disabled in terms of network selection and it can only update its users and requested data rate counters. The discovered network will have TU and BU probabilities $P_{TU} + P_{BU}=1$ and all α_{iVI}, α_{mVO}, α_{jBE}, α_{kBK} actions will involve updating the type of traffic and load in the network. In case of low networks load and users under both LTE and

Wi-Fi coverage conditions, P-ARMANS MDP algorithm with high probability could make free the network selection for a very slow mobility user. In case of medium-high mobility, the algorithm selects the network reducing vertical handover and preferring high network coverage rather than small cell networks. In case of high network load or overloaded network, a random user requesting access network has a probability PTU to be a typical user and a probability P_{BU}=1-P_{TU} to be a business user. All users are either TU or BU with the probabilistic constraint P_{BU}+P_{TU}=1.

Obviously, P_{TU} and P_{BU} depend on the percentage of TU and BU in the total number of users belonging to a defined MNO. Furthermore, we initially assume that there is a fixed and determined percentage of TU and BU according to MNO statistics on stipulated tariff plans. The transition probability also depends on the number of devices with a particular screen resolution associated to both TU and BU users. In MDP, the conditional probability is considered (i.e. the conditional probability is a measure of the probability of an event given that another event has occurred. Assuming P_{TU} and P_{BU}, we evaluate the conditional probability of a 4K, 2K, FHDi or HD state under the condition TU and BU. Hence, P(4K∩TU), P(2K∩TU), P(FHDi∩TU) and P(HD∩TU), represent the probabilities a request from a user of type TU to fall in the state S_1, S_2, and S_3, respectively, and at the same time TU is the type of user. On the other hand, P(4K∩BU), P(2K∩BU), and P(FHDi∩BU), represent the probability to fall in the state S_1, S_2, and S_3, respectively.

$P(4K \cap BU)$, $P(2K \cap BU)$, $P(FHDi \cap BU)$ and $P(HD \cap BU)$ are the probabilities a request from a user of type BU to fall in the state S_1, S_2, and S_3, respectively. Hence, the probability for a user request to be placed in state S_i is denoted as $P(TUorBU \cap S_i)$ and is defined as in equation (35):

$$P(TUorBU \cap S_i) = P(S_i|TUorBU)P(TUorBU) \tag{35}$$

where i=1,…,4 are the considered states. Equation (35) enable introduction of percentage of TU and BU in the initial states S_1, S_2, S_3 and S_4, when α_{iVI} actions have occurred. The remaining λ_i and $\gamma_{i,j}$ actions and their probabilities depend upon the political that governs MDP load distribution and retrieval of new resources to satisfy new user requests. For instance, a BU λ_i action has a probability to be verified equal to $P(\lambda_i \cap BU)$=0 until the total amount of TUs are treated with $\gamma_{i,j}$ actions, due the higher BU priority compared to TUs. P-ARMANS MDP algorithm could be set and proposed in different configurations in such a way that BU λ_i actions will verify before $\gamma_{i,j}$ actions. Both $\gamma_{i,j}$ and λ_i actions have a probability according to the number of users making VI requests and β_{kBK} actions during a TTI. This work assumes $\alpha_{kBK} >> \beta_{kBK}$ equivalent with high traffic load and overloaded networks.

5.6.6 MDP ALGORITHM AND REWARD FUNCTION

P-ARMANS MDP-based algorithm establishes how to assign resources when a new user connection occurs within a heterogeneous wireless environment. P-ARMANS algorithm collects information about the geolocalised user and network conditions under its coverage as shown in Figure 22 and is explained in section 5.3.1 and 5.3.2.

Typically, when mobile users activate their radio interfaces (i.e., LTE and Wi-Fi in the proposed scheme), it is possible to select different options managed by several applications. For instance, it is possible the automatic switching from LTE to Wi-Fi when a Wi-Fi network is detected or automatic switching from Wi-Fi to LTE when a Wi-Fi connection has been lost or when is under a certain signal strength threshold.

P-ARMANS MDP-based is deployed at the level of a server/gateway entity placed between the mobile user and MNO. It dynamically selects the best network for a user according its type of traffic and ensures high QoS level through load balancing. Each not-yet-connected user is placed in state S_0 which represents the starting point of the MDP algorithm. The P-ARMANS MDP process is described next.

- n users are placed under the radio coverage of one or more transmitters. Connections and applications running determine BK requests thanks to $\alpha_{kBK,}$ actions. These actions determine a transition from the state S_0 to the state S_7. The β_{kBK} action represents the transition from the BK traffic state S_7, to S_0 and results in closing connectivity to that particular network. Finally, the action 1-β_{kBK} means that the user maintains an active connection. P-ARMANS evaluates the probability of these 1-β_{kBK} actions every TTI, when a new user connects to a network or when a user closes the session. P-ARMANS entity updates the counters of connected users every time any of the above mentioned three cases has been verified.
- a user could start a voice session or a BE session by means of α_{mVO} or α_{jBE} actions, respectively. These actions determine a transition from the state S_0 to state S_5 and S_6, respectively. The β_{mVO} and β_{jBE} actions represent the transition from the VO traffic state S_5 and BE traffic state S_6 to the initial state, S_0. These actions were verified when a user closes the VO or BE sessions. Similar to the BK actions, 1-β_{mVO} and 1-β_{jBE} represent the probability to maintain active a VO and a BE session, respectively. Both actions are analyzed by P-ARMANS entity every TTI, when a new user connects to the network or when it or other users closes the work session.
- a user could start the VI session with α_{iVI} actions where i=1,…,4. These actions cause a transition from the state S_0 to S_1, S_2 and S_3, respectively. Each user is able to make α_{iVI} transitions according to their device screen resolution. In this study the states S_1, S_2 and

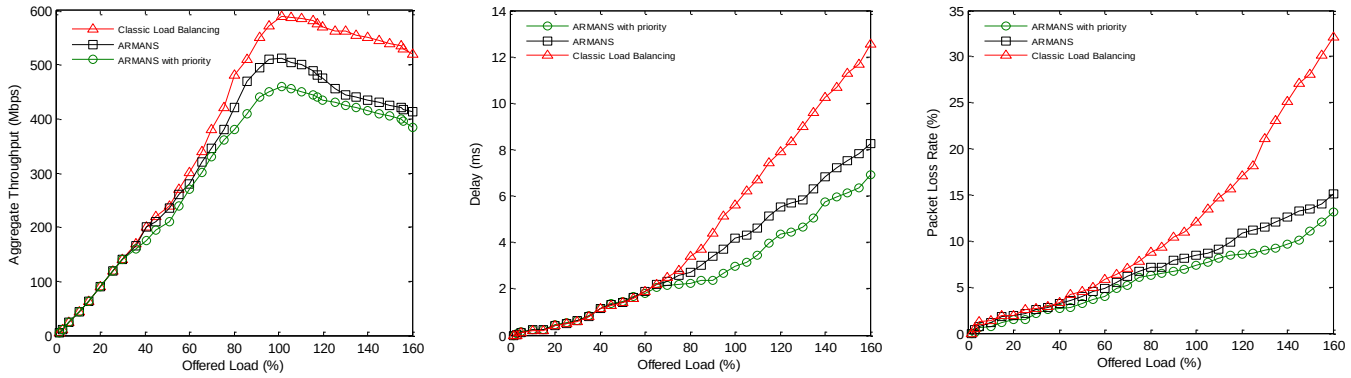
S_3, represent 4K, 2K and HD screen resolutions, respectively. Once the P-ARMANS entity recognizes a VI request by a 4K screen resolution user device, follows the α_{1VI} action and transitions from S_0 to S_1. Similarly it transitions to states S_2 and S_3, depending on the user device screen characteristics, respectively. β_{iVI} actions with i=1,…,4 are associated with closed VI sessions and the correspondent release of resources whereas 1-β_{iVI} actions maintain users in active mode for the VI traffic sessions.

- λ_i with i=1,…,4 represent the actions to stay at the same VI state decreasing the available VI data rate from maximum to a guaranteed minimum. Vice versa, 1-λ_i actions allow the transitions from de minimum to the maximum data rate. It is important to emphasize that the λ_i and 1-λ_i actions are available to TUs and BUs and are fully managed by P-ARMANS MDP algorithm.
- $\gamma_{i,i+1}$ actions represent the transition from the current state to the next worse state in terms of available VI data rate whereas $\gamma_{i+1,i}$ actions refer to transition from a current state to an immediately better one. The $\gamma_{i,i+1}$ actions make available the saved rate-related resources to other users. These actions are applied by P-ARMANS MDP algorithm to TU users in case of network overload, whereas the BU users are not performing these transitions. When the network load decreases, transitions are performed improving the state level Following this rule, a TU with a screen resolution such that it is originally placed in state S_1, S_2 and S_3 will be moved if necessary to the state S_2, S_3 and S_4, respectively.

As BUs have priority in comparison with TUs , P-ARMANS MDP-based will apply $\gamma_{i,i+1}$ actions to TUs before applying λ_i actions to BUs. Moreover, MDP will perform λ_i actions for BUs only when all TUs have been affected by $\gamma_{i,i+1}$ actions. Furthermore, we assume that VI content has priority with respect to any other type of traffic.

Parameters	Values
WiFi network	802.11n [78]
LTE network	LTE Cat. 6 [79]
Number of WiFi AP	1
Number of LTE cell	1
WiFi coverage radius	200 m
LTE cell radius	500 m
Transmit power of WiFi AP	100 mW (20 dBm)
Transmit power of LTE BS	39.8 W (46 dBm)
WiFi - LTE bandwidth	20 MHz
WiFi - LTE datarate (DL)	300 Mbps
Propagation model	(hybrid-indoor and outdoor)
User speed	< 3 Km/h

Table XVIII - Simulation Parameters



(a) (b) (c)

Figure 26 - Aggregate throughput for different load balancing schemes, with increasing amount of offered load (a), average delay for different load balancing schemes, with increasing the amount of offered load (b), and average packet loss rate for ARMANS and classic load balancing (c).

	Classic Load Balancing	ARMANS	Benefit (%)	ARMANS with priority	Benefit (%)
Average throughput (Mbps)	2.40	2.59	7.33%	3.25	26.15%
Standard deviation of throughput (Mbps)	4.18	2.96	29.18%	2.65	36.60%
Average delay (msec)	5.61	4.17	25.66%	2.95	47.42%
Standard deviation of delay (msec)	10.26	6.82	33.52%	5.75	43.95%
Average packet loss rate (%)	9.31	7.18	22.79%	6.52	30.10%
Standard deviation of packet loss rate	6.08	0.54	91.11%	0.46	92.33%
Average PSNR	18.62	26.58	29.94%	34.18	45.52%
Standard deviation of PSNR	16.99	1.83	89.21%	1.41	91.70%

Table XIX - Simulation Results of Scenario 1 and Scenario 2

5.7 ARMANS SIMULATION-BASED TESTING

The performance of the proposed ARMANS load-balancing algorithm is tested via in an OFDMA cellular system. The simulation parameters are given in Table XVIII. Next, the performance is analyzed when the number of UEs increases from 50 to 150. Results are evaluated in terms of average throughput [Mbps], average delay [sec], average packet loss rate (%), average PSNR and relative standard deviation when adding users to the networks. In Scenario 1, the ARMANS type of traffic load balancing has been analyzed compared to classic load balancing. In Scenario 2, ARMANS load balancing with resolution priority is evaluated. ARMANS with priority progressively decreases data rate from HQ to SQ starting from user with lower resolution to user with higher resolution thus according to Table XVII in which $P_1>P_2>P_3>P_4$. The aggregate throughput, average delay, and packet loss ratio for the simulated scenarios are illustrated in Figure 26, (a), (b) and (c), respectively. In Figure 27, the PSNR represents the end-user perceived quality level when receiving a 1080p HQ.

Results show how higher PSNR when using ARMANS and ARMANS with priority are achieved. Finally, average results for the simulation scenarios considered are listed in Table XIX. The results show that by using the proposed ARMANS with priority the selected users gain 26.15% in terms of throughput and records 47.42% and 30.10% decrease in average delay and average packet loss ratio, and 45.52% increase in PSNR.

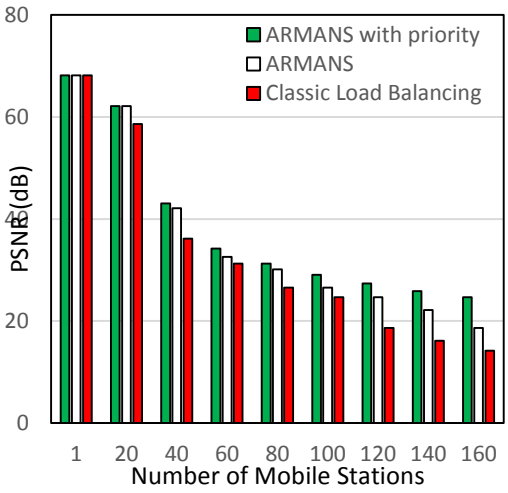


Figure 27 - PSNR achieved for a 1080p resolution using UDP traffic

5.8 P-ARMANS Simulation-based Testing

5.8.1 Multimedia Content

Four types of multimedia traffic were used for both IEEE 802.11n and LTE: VO, VI streaming BE, and BK contents. In Table XX the characteristics of the four traffic classes are detailed. The most common audio and video codecs ITU-T G.711 [87] and H.264/MPEG-4 AVC [88] and VP9 [89], respectively are considered. The encoding data-rates specified in standard (i.e., 64 kbps and 128 kbps), and typical packet sizes of 100, 1024 bytes, 1024 bytes, 512 bytes and 512 bytes are set for VO, VI, BE and BK traffic, respectively.

Several video data rate has been reported in Table XX. Each state S_i is characterized by a defined screen resolution. State S_1 corresponds to 2160p 4K with a screen resolution 3840x2160 pixels. Last generation of smartphone and tablet are essentially the same in terms of screen size and battery life, while increasing screen resolution up to 4K. Quad HD display, also known as 1440p has been the evolution of the 1080p, with a screen resolution of 2560x1440 pixels principally employed in a wide range of phablets. All devices able to elaborate such screen resolution are placed in state S_2. Nowadays, full HDi 1080p screen

resolution with 1920x1080 pixels is the most common format present in the mobile devices market. S_3 is the state that contains all devices capable to manage FHDi video resolution. Finally, the HD-ready state, also known as 720p and 1280x720 pixels of screen resolution is the state S_4. This state does not correspond to a particular set of devices but is referred to the capability to receive a data rate associated to a HD-ready resolution for devices belonging to state S_3 when overloaded networks occurred. For each state, Table XX reports a maximum and a minimum compressed data rate. These values are the recommended bitrates video codec H.264 AVC and VP9 for web browsing streaming with HTML5 Video and FLASH video file format.

The typical audio 128 Kbps quality, the most common format for audio streaming or storage, is used.

State S_i	Resolution	VP9 – H.264 AVC (Mbps)		Audio (Kbps)
		(max)	(min)	
2160p (4K)	3840x2160	13.0	10.4	128
1440p (Quad HD)	2560x1440	7.549	5.816	128
1080p (Full HD)	1920x1080	3.676	2.804	128
720p (HD Ready)	1280x720	2.674	1.229	128

Table XX - Encoding Video Data Rate

5.8.2 SIMULATION SCENARIOS

Table XXI details simulation parameters. The 802.11n [78] defines 32 Modulation and Coding Scheme (MCS) that consider several attributes such as spatial stream, modulation type (i.e., BPSK, QPSK, 16-QAM, 64-QAM), coding rate and maximum data rate according to 20 MHz or 40 MHz channels and 400 ns or 800 ns guard interval. In this work we consider a coverage radius about 200 meters and a transmitted power no higher than 100 mW. The expected bandwidth is 40 MHz and the maximum data rate in downlink according to the standard and MCS could reach 600 Mbps. Similarly, LTE [90] standard define 15 Channel Quality Indicator (CQI) carrying the information on how good/bad the communication channel quality is. The eNodeB has a higher radius coverage estimated about 500 meters, a transmitted power equal to 39.8 W and a bandwidth 20 MHz defined by the standard. The propagation model is a mixed configuration named hybrid-indoor and outdoor. This model allows the user performance evaluation when inside a building or outdoor, in free space and urban environment.

User speed is also considered. Static user (i.e., speed < 3 Km/h), slow moving users (i.e., 3 < speed < 15 Km/h), and high speed users (i.e., speed > 15 Km/h) are the three speed range simulated in the proposed scenarios.

P-ARMANS MDP algorithm has been evaluated by using the NS-3 network simulator [85]. The simulation topology shown in Figure 22 includes one P-ARMANS server communicating with a scalable number of n clients, LTE eNodeB and Wi-Fi AP WNs.

Parameters	Values
Wi-Fi network	802.11n [78]
LTE Network	LTE Cat. 6 [90]
Number of Wi-Fi AP	1
Number of LTE cell	1
Wi-Fi coverage radius	200 m
LTE cell radius	500 m
Transmit power of Wi-Fi AP	100 mW (20 dBm)
Transmit power of LTE BS	39.8 W (46 dBm)
Wi-Fi / LTE bandwidth	40 / 20 MHz
Wi-Fi / LTE data rate (DL)	600 / 300 Mbps
Propagation model	(hybrid-indoor and outdoor)
User speed	< 3 - <15 - >15 Km/h

Table XXI - Simulation Parameters

According to an MNO estimation on user agreements, Table XXII shows how TUs have a probability equal to 80% and BUs have a probability equal to 20% to request network resources.

Moreover, Table XXII includes assume predefined percentages of 4K, 2K and FHDi devices, both for TUs and BUs and conditional probabilities computed according to equation (35). For a TU we assumed a defined percentage of devices associated to a particular native screen resolution: 15%, 25% and 60% for 4K, 2K and FHDi devices, respectively. A BU has been characterized taking into account different percentages compared to TUs: 20%, 25% and 55% of users have a 4K, 2K and FHDi native screen resolution devices, respectively. According to equations (19) and (20), it is possible calculate the probability of the intersection of two dependent events (i.e., TU or BU and Si) given by the product of the probability of an event for the conditional probability of the other. A random access of a TU or BU with a particular screen resolution device shows that for the 48% is a TU with FHDi device requesting VI contents. The remaining of TUs VI requests are given by 2K and 4K devices with a percentage of 20% and 12%, respectively. Furthermore, there is only 4% of probability of BU random access with a 4K device VI request. A BU with a 2K native screen resolution device has a probability equal to 5% to request VI traffic whereas 11% is the percentage of VI contents requested for FHDi devices.

Network load considered a 5% of users generating BK, BE, and VO traffic only and the remaining 95% of users being involved in high-quality VI requests. The number of mobile devices has been progressively increased in steps of 1 every 2 seconds in order to collect

statistics in relatively stable conditions, and increasing the overall offered load. According to Table XXII, random TU and BU accesses after 200 seconds and without disconnection cases, consists of 100 users: 80 of these are TUs and 20 are BUs. First 100 users produce a total amount of traffic load around 613 Mbps. According to Table XX we assumed that both TUs and BUs are receiving the maximum data rate given its native screen resolution. This assumption is valid when the network load is between 1% and 95%. Considering 140 connected users at the same time, network load reaches a total amount of traffic load equal to 860 Mbps. Starting from random user number 141, P-ARMANS MDP algorithm starts a decision process based on load balancing and user priority.

Typical User (TU)			Business User (BU)		
$P_{TU} = 0.8$			$P_{BU} = 0.2$		
4K	2K	FHDi	4K	2K	FHDi
15%	25%	60%	20%	25%	55%
User with $P(TU \cap S_i)$			User with $P(BU \cap S_i)$		
12%	20%	48%	4%	5%	11%

Table XXII - Type of Users and Traffic Characteristics

The deployed P-ARMANS MDP can be summarized as follows:

1. λ_i are the first actions applied to TUs in the sequence λ_3-λ_2-λ_1. λ_i actions involve a reduction data rate from maximum to minimum, making available to new users resources previously occupied.
2. $\gamma_{i,i+1}$ actions in the sequence $\gamma_{3,4}$-$\gamma_{2,3}$-$\gamma_{1,2}$ are applied to TUs. TUs passed from the native state to the state S_{i+1} at the maximum data rate.
3. TUs suffer a further reduction of data rate with the actions λ_4-λ_3-λ_2.
4. the last step involves BUs. The actions λ_3-λ_2-λ_1 were applied to BUs to decrease the data rate from maximum to minimum, making it available for new users while preserving their native screen resolution state.

The proposed MDP ensures a gradual resource re-allocation and a growing number of users performs λ_3-λ_2-λ_1 actions as detailed in step 1. These actions allow an increase in the number of users connected to request VI contents, from 141 to 173. Step 2 forced TUs to decrease the original data rate via $\gamma_{3,4}$-$\gamma_{2,3}$-$\gamma_{1,2}$ actions. Thanks to these actions the maximum number of connected users increases to 208. Step 3 applies λ_4-λ_3-λ_2 actions and a further VI data rate reduction for TUs. New users requesting VI contents are served and the user count increases to 279 users. No other actions involve TUs as in step 4 MDP reduces the available BU's data rate via λ_3-λ_2-λ_1 actions. The number of BUs and TUs managed by MDP after 4 steps reaches 307 users. Given the two networks considered and their characteristics, no more users requesting VI contents were further accepted, whereas more VO, BE, BK requests could be accepted. Table XXIII details the VI data rate [Mbps] assigned to each user type in each step and the MDP proposed case. In particular the data rates assigned to TUs,

BUs and on average are indicated when the number of users increases. Following the actions in step 1, the TUs start from the average data rate of 6.04 Mbps and it has been limited to 4.70 Mbps. During step 4, there is only a data rate reduction for BUs. Figure 28 shows the trend of MDP process. The random access for user 145 needs unavailable resources. The MDP process acts in order to make available data rate to new users.

P-ARMANS MDP progressively decrease the data rate of connected user with the action sequences represented by Table XXIII. MDP is a 12-peaks curve where each peak represents the 95% traffic load supported. To each peak correspond 3 actions for step 1, 2 and 3, respectively. Step 4 involves BU users and 3 more actions.

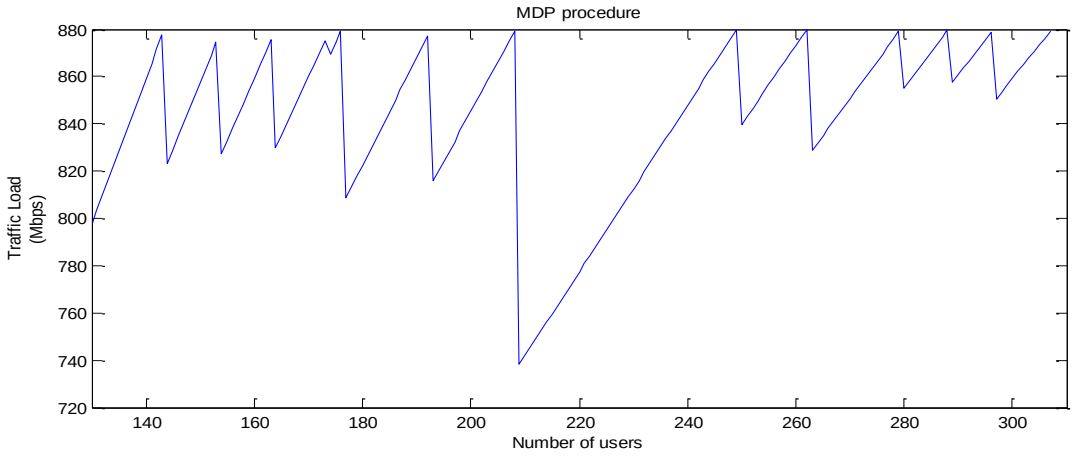


Figure 28 - MDP Process

Step	Step 1	Step 2	Step 3	Step 4
Actions	λ_3-λ_2-λ_1	$\gamma_{3,4}$-$\gamma_{2,3}$-$\gamma_{1,2}$	λ_4-λ_3-λ_2	λ_3-λ_2-λ_1
User Type	TUs	TUs	TUs	BUs
User No	141-173	174-208	209-279	280-307
TU Rate	6.04-4.70	4.69-3.66	3.65-2.31	2.31
BU Rate	6.51	6.51	6.51	6.50 -5.08
Avg. Rate	6.21-5.05	5.06-4.23	4.22-3.15	3.14-2.86

Table XXIII - P-ARMANS MDP Actions Sequence and Data Rate

5.9 SIMULATION-BASED TESTING RESULTS

Four scenarios have been analyzed and presented in Figure 29.

In Scenario 1, the mobile TU do not move, the losses are mainly due to the amount of traffic requested by the TUs and BUs connected to the AP and LTE. Each node generates traffic classified in four different type of traffic, device characteristics and screen resolution. The overall number of users and traffic gradually increases in order to overload both networks.

In Scenario 2 one TU user moves in a path towards and then away from the AP and LTE at a constant speed of 1 m/s. In this case, the losses are due to variable received power with increased or decreased distance from the AP and/or LTE, and the traffic load time-varying. All other settings are like in the previous scenario.

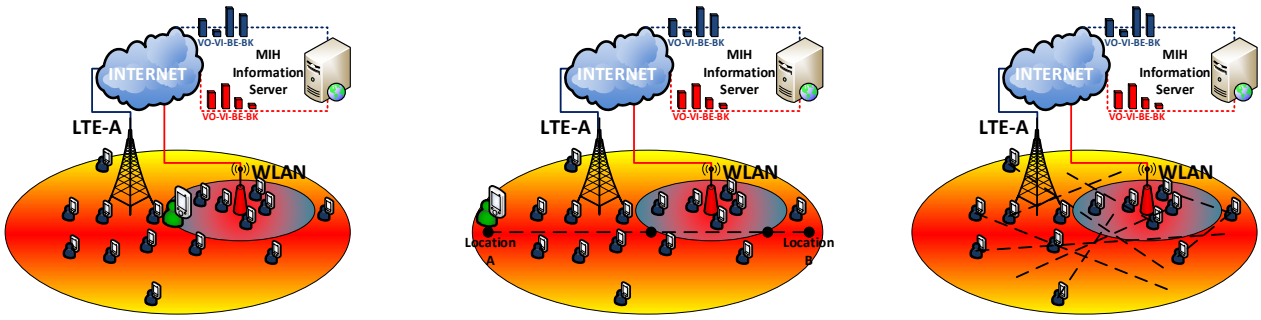
In Scenario 3 the total number of users is divided in mobile and non-mobile, and progressively the number of mobile users increases. Performance was evaluated for a static TU placed within LTE and AP area.

Scenario 4, similar to scenario 3, but the ARMANS load balancing with resolution priority is simulated. ARMANS with priority progressively decreases data rate from 4K to FHDi. Each scenario focuses on a TU (i.e., 4K, 2K, FHDi) and its performances evaluation for MDP in cases 1, 2 and 3, respectively.

Simulation-based testing analyzed the QoS in terms of packet delay, packet loss rate, average PSNR and aggregate throughput comparing P-ARMANS MDP with ARMANS load balancing for each of four proposed scenarios.

5.9.1 PACKET DELAY RESULTS

Figure 30 shows in detail the packet delay calculated for a TUs with a different screen resolution when the number of users is progressively increased. All the curves are characterized by a monotonically increasing trend with the increasing connected number of user. Scenario 1 is characterized by static users within a LTE and Wi-Fi area and the packet delay is evaluated for the users from 130 to 307. For all scenarios, the packet delay for a 2K screen resolution user is higher than FHDi and lower than 4K. Higher data rate requirements due to higher screen resolution implies higher packet delay when receiving a multimedia video content. Measured packet delay for the maximum amount of admitted user, falls within the range 10.55-15.8 ms for FHDi and 4K screen resolution, respectively. Scenario 2 similar to scenario 1 involves a TU at a constant speed of 1 m/s testing 3 different screen resolution devices. Speed increases the packet delay up to the range 20.32-22.25 ms at the maximum number of users and traffic load. Scenario 3 considers initially mobility and non-mobility users progressively increasing the number of mobility user. A generic TU is influenced by the growing number of mobility users and represent the worst study case to evaluate the QoS in terms of packet delay. In this case a 4K TU receives packet with a maximum delay of 29.89 ms whereas a FHDi and 2K TU receive video contents with a packet delay of 28.33 ms and 28.90 ms, respectively. Finally, in scenario 4 we evaluate the case in which no-priority action and MDP were taken. This case correspond to higher delay for all screen resolution. Simulated packet delay falls within the range 31.12-33.40 ms for FHDi and 4K screen resolution, respectively.



(a) (b) (c)

Figure 29 - Scenario 1 (a), 2 (b) and 3 (c)

5.9.2 AVERAGE PSNR DELAY RESULTS

Figure 31 deals with the average PSNR for different screen resolution and for all scenarios. All the curves are characterized by a monotonically decreasing trend with the increasing connected number of user. As evaluated for packet delay, PSNR has been evaluated in a relevant range of overloaded networks starting from user number 144. Before this threshold, the PSNR is predominantly around 37 dB. Figure 7 represents the evaluation of the PSNR when the number of user is increasing as well as their mobility. Scenario 1 is characterized by static and the growing number of users in overloaded case. The simulated PSNR for FHDi, 2K and 4K BU is equal to 31.84, 31.2 and 30.95 dB, respectively. From the user number 144 and the user number 307, the P-ARMANS MDP worsens the PSNR of 5.16-6.05 dB.

When BU user is moving according to scenario 2, the PSNR is influenced by a higher losses and power fluctuations. Whereby, the simulated PSNR loses from 2.24 dB up to 1.81 dB when the considered TU is moving at a constant speed of 1 m/s alongside and/or far away transmitters. Also in this scenario, the PSNR follows the rule "the higher screen resolution the lower PSNR": 29.14 dB, 29.24 dB and 29.6 dB are the simulated value for 4K, 2K and FHDi, screen resolution TU devices, respectively.

Increasing the number of mobile user in mobility, each user will be more affected from others. With respect to scenario 2, a further PSNR reduction has been estimated. FHDi screen resolution finds a reduction by up to 2.17 dB whereas 4K devices lose 2.36 dB. FHDi, 2K and 4K have a PNSR equal to 17.43, 27.2 and 26.78 dB, respectively. Finally, the scenario 4 considers the worst case of all users in mobility without MDP to optimize the traffic load. In this case, the higher simulated PSNR reduction equal to 1.8 dB belong to 2K screen resolution user (i.e., 27.2 dB to 25.4 dB) whereas 4K and FHDi PSNR reduction is equal to 1.58 dB (i.e., 26.78 dB to 25.20 dB) and 1.65 dB (i.e., 27.43 dB to 25.78 dB), respectively.

5.9.3 PACKET LOSS RATE RESULTS

In Figure 32, simulated PLR is characterized by curves with a monotonically increasing trend with the increasing connected number of user. In scenario 1 the static TU is slowly affected by PLR within the range 130-180 user where the trend is almost flat. In this range, PLR increased by up to 0.7%, whereas within the range 180-235 users we notice a further increment of 1.1% for a global PSNR of 5.8%. The range of users 235-280 have a PLR around 7.5% for both 2K, 4K and FHDi users. Last range (i.e., 280-307 users) shows different trends for 4K screen resolution devices compared to 2K and FHDi: 4K screen resolution suffer higher PLR (i.e., 10%) than 2K (i.e., 8.25%) and FHDi (i.e., 8.1%), respectively. The TU suffers high PLR adding mobility when requesting high data rate VI traffic. A significant PLR has been encountered in scenario 2 compared to scenario 1. After the first traffic overload and around 150 connected users, a growing PLR affect 4K, 2K, and FHDi users by up to 2.8%, 2.4%, and 1.8%, respectively. A second and significant PLR increase could be observed around the 240 connected users; the trend of the curves change from almost flat to a marked rise of the pendency. At the maximum number of supported users, PLR is equal to 22.45%, 21.4%,and 19.7% for FHDi, 2K and 4K screen resolution devices. Scenario 2 has a high impact on PLR when overload occurred at 307 users, with an increase of 12.46%, 13.15%, and 11.45% for FHDi, 2K and 4K screen resolution devices, respectively. In scenario 3 the PLR is influenced by the presence of many users in mobility and the difference between packets sent and packets received is reduced. PLR progressively increases up to the user 307 and the simulated values are equal to 34.9%, 33.1%, and 31.3%. Finally, the ARMANS algorithm without MDP optimization shows that the PLR is always higher and with a similar curve trend than scenario 3.

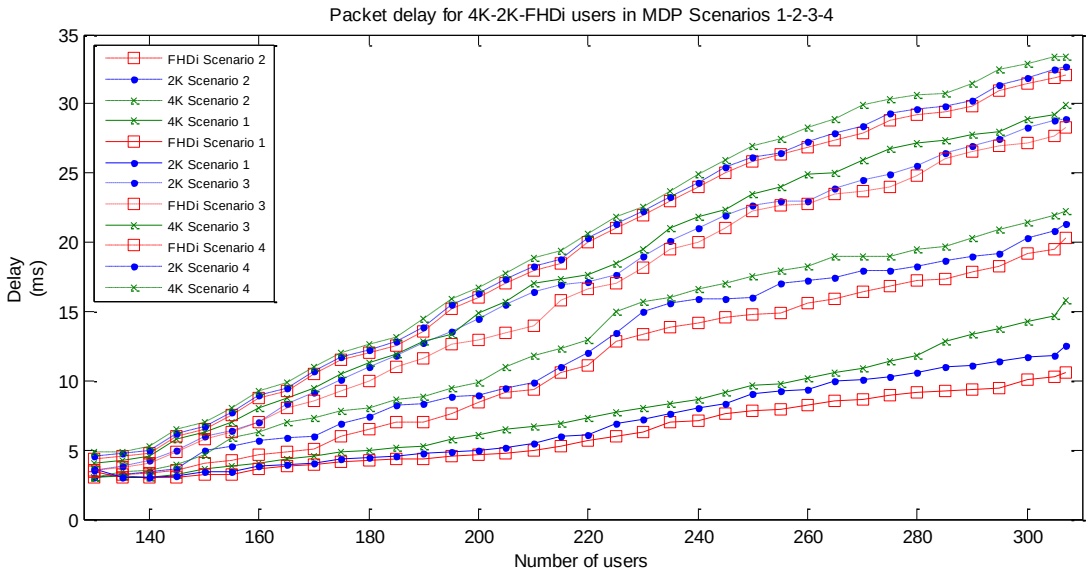


Figure 30 - Packet delay for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP

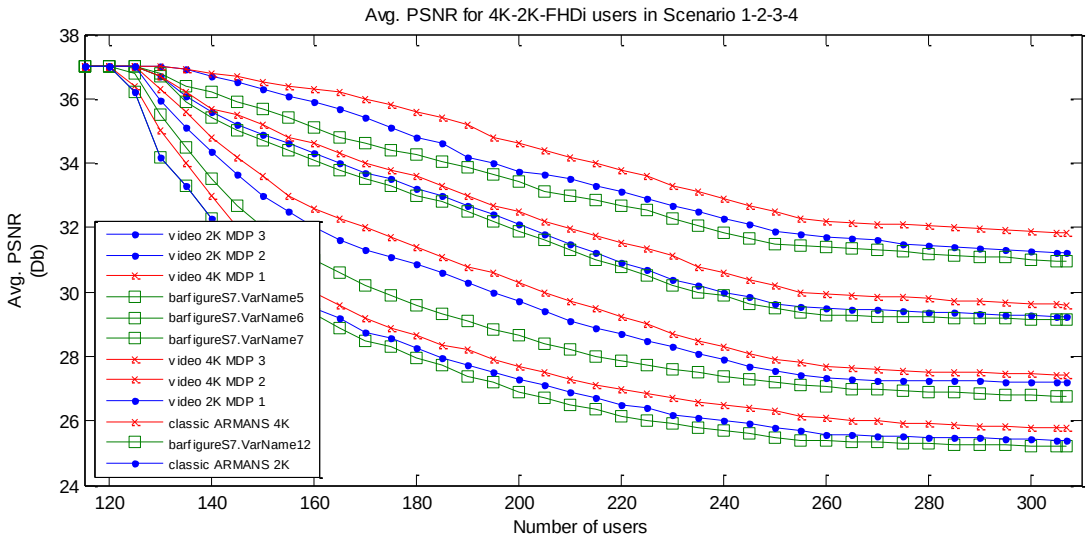


Figure 31 - Average PSNR for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP

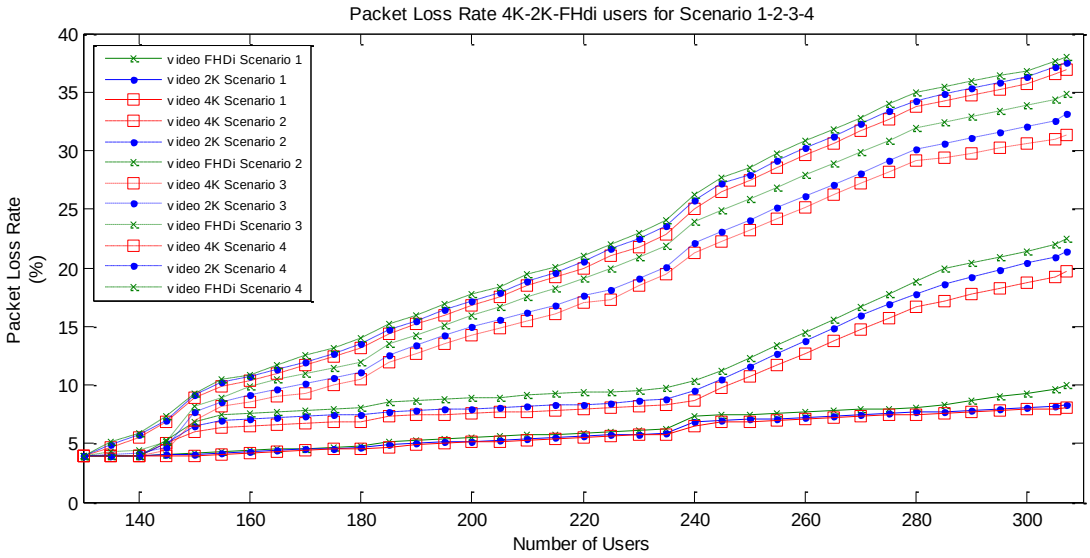


Figure 32 - Packet loss rate for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP

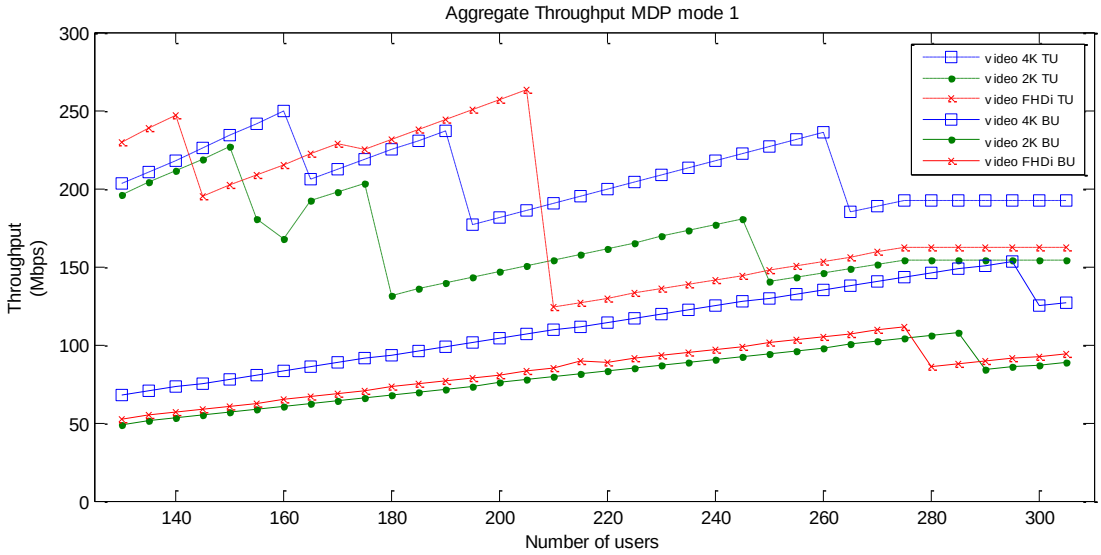


Figure 33 - Aggregate throughput for 4K, 2K and FHDi TUs and BUs evaluated with P-ARMANS MDP

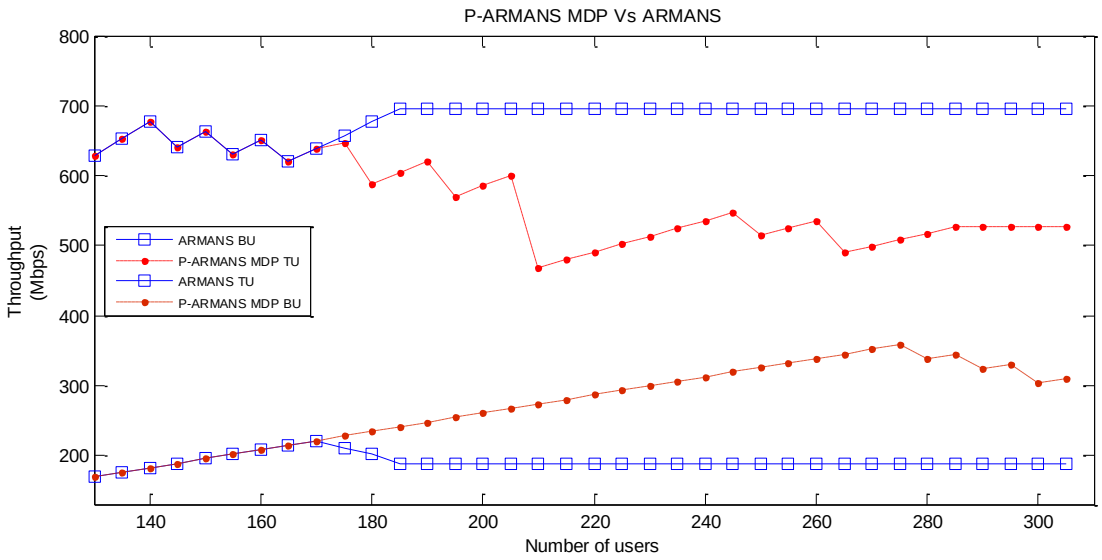


Figure 34 - Packet loss rate for 4K, 2K and FHDi TUs evaluated with P-ARMANS MDP and ARMANS

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Benefits (%)
Average Packet Delay (ms)					
FHDi	6.19946	11.1284	16.224	19.18432	15.295
2K	6.97405	12.14	16.9508	19.5	13.072
4K	7.92351	13.059	17.63	20.11892	12.37
Standard deviation	4.57	8.16	11.45	13.2	13.275
Average PSNR (dB)					
FHDi	34.0308	33.467	33.0332	27.925	15.46
2K	32.032	31.64	31.466	27.464	12.718
4K	29.547	29.996	28.8376	27.219	5.613
Standard deviation	2.95	2.98	2.93	3.05	3.93
Average PLR (%)					
FHDi	5.8135	10.0932	17.8189	21.103	15.562
2K	5.934	10.842	18.806	21.565	12.794
4K	6.3397	11.5905	19.889	22.029	9.714
Standard deviation	2.075	8.2	14.2	15.25	6.885

Table XXIV - Simulation Results Evaluation

5.9.4 AGGREGATE THROUGHPUT RESULTS

In Figure 33, a detailed trend of aggregate throughput for both TU and BU is drawn. MDP algorithm enforces the BUs priority rather than TUs, and applies to TUs the actions in step 1-2-3 according to Table XXIII. When no further actions can be applied to TU, MDP algorithm step 4 decreases the data rate of BU. According to the proposed percentage of BU and TU and the percentage of FHDi, 2K, and 4K, the user number 140 determines a traffic overload. MDP step 1 involves the actions λ_i with the sequence λ_3-λ_2-λ_1 to be applied to FHDi, 2K, and 4K screen resolution, respectively. Figure 33 shown that 130 users request are divided as follows: 229.38 Mbps, 196.3 Mbps and 202.8 Mbps for FHDi, 2K, and 4K screen resolution TU devices whereas 67.2 Mbps, 49.07 Mbps, and 52.57 Mbps for FHDi, 2K, and 4K screen resolution BU devices. The user 140 determines the first traffic overload and

the MDP algorithm action λ_3 reduces the available data rate to FHDi TU devices. Similarly, λ_2-λ_1 actions reduce the available data rate to 2K and 4K TUs, respectively. The ever-growing number of connected users drives P-ARMANS MDP to act with step 2 and step 3 in order to make available resources to the new users. The user number 178 characterizes the MDP step 4 and the actions λ_3-λ_2-λ_1 were applied to BUs to decrease the data rate from maximum to minimum according to the value in Table XXIII. The final resource allocation optimized by MDP is as follows: 191.93 Mbps, 154.2 Mbps, and 162.23 Mbps for FHDI, 2K, and 4K TUs, whereas 126.88 Mbps, 88.69 Mbps, and 94.07 Mbps for FHDI, 2K, and 4K BUs, respectively. The total amount of traffic is distributed in 309.65 Mbps for BUs and 526.86 Mbps for TUs.

5.9.5 P-ARMANS MDP AND ARMANS RESULTS

Finally, Figure 34 shows the difference between P-ARMANS MDP and ARMANS. The horizontal trend of BUs and TUs aggregate throughput is due to ARMANS. In ARMANS study case, the user number 185 requires not available resources. New users are admitted without increasing the overall radio resources. Whereby, the flat trends identify the maximum load to be shared among all users, both for BUs and TUs. On the other hand, P-ARMAN'S MDP progressively reduces and shares the wireless resources among users while ensuring higher number of connected devices and high QoS. MDP step 1, 2, and 3 act when the user number 141, 154, 164, 174, 177, 193, 209, 250, 263, and 280 require VI contents. Limited resources implicate actions α_{iVI}, λ_i, and $\gamma_{i,i+1}$ during step 1, 2, and 3 for TUs and when no further action could be applied to TUs, MDP acts with actions λ_i on BUs when a user 288 and 296 need VI multimedia contents.

Table XXIV summarizes the benefits introduced by P-ARMANS MDP analyzed in different scenarios and compared to ARMANS algorithm. Packet delay, PSNR and PLR are compared in terms of average values for each scenario and screen resolution devices. Average packet delay and average PLR, progressively increase when adding mobility to the users, starting from a scenario with static users and finally evaluating a scenario where all users are moving at a constant speed of 1 m/s. At the same time, average PSNR progressively decrease due to the increment of interferences produced by a growing number of user in mobility. Standard deviation represents a range in which fall all the simulated value of the employed metrics to evaluate the QoS. Benefits show the obtained improvements when comparing scenario 3 with scenario 4. P-ARMANS MDP has a average packet delay reduction of 12.37%, 13.07%, and 15.30% for 4K, 2K, and FHDi screen resolution, respectively. MDP algorithm gives advantages also in terms of PLR with percentages of 9.71%, 12.79%, and 15.56% for 4K, 2K, and FHDi screen resolution, respectively. Finally, the

PSNR shows an improvement of 15.46%, 12.72% and 5.61% when MDP algorithm manages the access network selection and load balancing over such HWN.

5.10 CONCLUSION

This work proposes the Adaptive Real-time Multi-user Access Network Selection algorithm (ARMANS) for improved HQ video delivery in a heterogeneous wireless network environment. ARMANS access index is introduced and indicates whether a user is better to access WiFi or LTE network, respectively. By making use of type of traffic management, the ARMANS framework is able to perform load balancing by selecting the most appropriate network for the user. The results show that the proposed ARMANS improves QoS in terms of available throughput, delay, packet loss and PSNR. ARMANS with priority can achieve up to 26.11% increase in throughput, and up to 45.52% increase in PSNR, when compared against a classic load balancing algorithm. In terms of future work, additional parameters and improvements could be integrated into the current solution in order to enhance also energy consumption and mobile user experience.

This work also proposes a novel P-ARMANS MDP algorithm to solve the access network selection and overloaded network problem. The authors proposed the P-ARMANS network selection algorithm to address the user to one network rather than another one, basing this selection on several user parameters such as screen resolution, type of traffic requested, user position compared to transmitters position, and user mobility. In this work we classified three screen resolution devices, the recent 4K, the 1080p 2K format and the FHDi screen resolution.

MDP plays a very important role in solving complex problems as load balancing breaking it in a set of sub problems. Each screen resolution devices is classified and treated like a state of the Markov chain. The decision process helps the real-time system to avoid network congestion while ensuring QoS and a high number of connected users.

The optimization is focused on multimedia high-quality video content that represents the biggest component of radio resources requested. MDP uses several actions and steps to dynamically manage the available resources and make new available resources to further users access network. A level of priority has been introduced among users: BUs and TUs. The first typology is represented by business users with a guaranteed data rate whereas the TU is a user without privileges. Compared to ARMANS algorithm, P-ARMANS MDP algorithm provides several benefits in terms of packet delay, PSNR, PLR and aggregate throughput. The average improvements are equal to 13.6%, 11.26%, and 12.69% for packet delay, PSNR, and PLR, respectively. The aggregate throughput results show that P-ARMANS ensures a bigger number of connected user through a slow and progressive

reduction of the data rate initially assigned to the user and the possibility to be available to new traffic requests.

REFERENCES

[1] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys.*, vol. 21, no. 6, pp. 1087–1092, Jun. 1953.

[2] S. Kirkpatrik, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science,*.

[3] B. Rosen, "Function Optimization Based On Advanced Simulated Annealing," in *Workshop on Physics and Computation*, 1992, pp. 289–293.

[4] "Wiley: Electromagnetic Optimization by Genetic Algorithms - Yahya Rahmat-Samii, Eric Michielssen." [Online]. Available: http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471295450.html.

[5] A. Gräf, "DAB Ensemble Planning – Problems and Techniques," *Telecommun. Syst.*, vol. 18, no. 1–3, pp. 137–154, Sep. 2001.

[6] "EN 302 304 - V1.1.1 - Digital Video Broadcasting (DVB); Transmission System for Handheld Terminals (DVB-H) - en_302304v010101p.pdf." [Online]. Available: http://www.etsi.org/deliver/etsi_en/302300_302399/302304/01.01.01_60/en_302304v010101p.pdf.

[7] "EN 300 744 - V1.6.1 - Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television - en_300744v010601p.pdf." [Online]. Available: http://www.etsi.org/deliver/etsi_en/300700_300799/300744/01.06.01_60/en_300744v010601p.pdf.

[8] "a122_DVB-T2_spec.pdf." [Online]. Available: https://telcogroup.ru/files/materials-pdf/DVB_standards/DVB-T/a122_DVB-T2_spec.pdf.

[9] A. Mattsson, "Single frequency networks in DTV," *IEEE Trans. Broadcast.*, vol. 51, no. 4, pp. 413–422, Dec. 2005.

[10] "Model and heuristic for a generalized access network design problem | SpringerLink." [Online]. Available: http://link.springer.com/article/10.1023/A:1019147029476.

[11] S. R. Tiourine, C. a. J. Hurkens, and J. K. Lenstra, "Local search algorithms for the radio link frequency assignment problem," *Telecommun. Syst.*, vol. 13, no. 2–4, pp. 293–314, Jul. 2000.

[12] M. Anedda *et al.*, "Heuristic optimization of DVB-T/H SFN coverage using PSO and SA algorithms," in *2011 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2011, pp. 1–5.

[13] J. R. Perez, J. Basterrechea, J. Morgade, A. Arrinda, and P. Angueira, "Optimization of the Coverage Area for DVB-T Single Frequency Networks Using a Particle Swarm Based Method," in *VTC Spring 2009 - IEEE 69th Vehicular Technology Conference*, 2009, pp. 1–5.

[14] L. Wang, "An improved cooperative particle swarm optimizer," *Telecommun. Syst.*, vol. 53, no. 1, pp. 147–154, May 2013.

[15] "ITU-R P.1546: Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3 000 MHz." [Online]. Available: https://www.itu.int/rec/R-REC-P.1546-5-201309-I/en.

[16] M. García-Lozano, S. Ruiz-Boqué, and F. Minerva, "Static delays optimization to reduce self-interference in DVB-T networks," in *2010 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2010, pp. 1–6.

[17] I. Eizmendi, G. Prieto, G. Berjon-Eriz, I. Landa, and M. Velez, "Empirical DVB-T2 Thresholds for Fixed Reception," *IEEE Trans. Broadcast.*, vol. 59, no. 2, pp. 306–316, Jun. 2013.

[18] J. Morgade, J. R. Perez, J. Basterrechea, M. T. Garcia, A. Arrinda, and P. Angueira, "Coverage optimization for DVB-T/H single frequency networks using a PSO algorithm," in *2009 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, 2009, pp. 1–6.

[19] S.-G. Chang and B. Gavish, "Telecommunications network topological design and capacity expansion: Formulations and algorithms," *Telecommun. Syst.*, vol. 1, no. 1, pp. 99–131, Dec. 1993.

[20] M. Rabinowitz and J. J. Spilker, "A new positioning system using television synchronization signals," *IEEE Trans. Broadcast.*, vol. 51, no. 1, pp. 51–61, Mar. 2005.

[21] "OFDM receivers - impact on coverage of inter-symbol interference and FFT window positioning - trev_295-brugger.pdf." [Online]. Available: https://tech.ebu.ch/docs/techreview/trev_295-brugger.pdf.

[22] A. Ligeti and J. Zander, "Minimal cost coverage planning for single frequency networks," *IEEE Trans. Broadcast.*, vol. 45, no. 1, pp. 78–87, Mar. 1999.

[23] "ITU-R Report BT.2254. Frequency and network planning aspects of DVB-T2." [Online]. Available: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-BT.2254-2012-PDF-E.pdf.
[24] "Sardegna DTV MUX list. List of broadcasters television." [Online]. Available: http://www.sardegnahertz.it/html/database/vista/elenco_tv.php?id=1&loc=1&pay=1.
[25] "ITU-R P.1812: A path-specific propagation prediction method for point-to-area terrestrial services in the VHF and UHF bands." [Online]. Available: http://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.1812-4-201507-I!!PDF-E.pdf.
[26] M. Lanza *et al.*, "Coverage optimization in single frequency Networks using simulated annealing," 2011, pp. 2789–2792.
[27] "EN 302 307 - V1.2.1 - Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications (DVB-S2) - en_302307v010201o.pdf." [Online]. Available: http://www.etsi.org/deliver/etsi_en/302300_302399/302307/01.02.01_40/en_302307v010201o.pdf. [Accessed: 27-Dec-2016].
[28] A. Meloni and M. Murroni, "Random Access in DVB-RCS2: Design and Dynamic Control for Congestion Avoidance," *IEEE Trans. Broadcast.*, vol. 60, no. 1, pp. 16–28, Mar. 2014.
[29] W. Sung, S. Kang, P. Kim, D.-I. Chang, and D.-J. Shin, "Performance analysis of APSK modulation for DVB-S2 transmission over nonlinear channels," *Int. J. Satell. Commun. Netw.*, vol. 27, no. 6, pp. 295–311, Nov. 2009.
[30] D. He *et al.*, "Improvements to APSK constellation with gray mapping," in *2014 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, 2014, pp. 1–4.
[31] H. Jiang and P. A. Wilford, "A hierarchical modulation for upgrading digital broadcast systems," *IEEE Trans. Broadcast.*, vol. 51, no. 2, pp. 223–229, Jun. 2005.
[32] H. Jiang, P. A. Wilford, and S. A. Wilkus, "Providing Local Content in a Hybrid Single Frequency Network Using Hierarchical Modulation," *IEEE Trans. Broadcast.*, vol. 56, no. 4, pp. 532–540, Dec. 2010.
[33] H. Jin, K. Peng, and J. Song, "Bit Division Multiplexing for Broadcasting," *IEEE Trans. Broadcast.*, vol. 59, no. 3, pp. 539–547, Sep. 2013.
[34] H. Méric, J. Lacan, F. Arnal, G. Lesthievent, and M. L. Boucheret, "Combining Adaptive Coding and Modulation With Hierarchical Modulation in Satcom Systems," *IEEE Trans. Broadcast.*, vol. 59, no. 4, pp. 627–637, Dicembre 2013.
[35] P. K. Vitthaladevuni and M. S. Alouini, "BER computation of 4/M-QAM hierarchical constellations," *IEEE Trans. Broadcast.*, vol. 47, no. 3, pp. 228–239, Settembre 2001.
[36] J. Zoellner and N. Loghin, "Optimization of high-order non-uniform QAM constellations," in *2013 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2013, pp. 1–6.
[37] F. Kayhan and G. Montorsi, "Constellation design for transmission over nonlinear satellite channels," in *2012 IEEE Global Communications Conference (GLOBECOM)*, 2012, pp. 3401–3406.
[38] E. D. Re *et al.*, "SALICE project: Satellite-Assisted Localization and Communication Systems for Emergency Services," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 28, no. 9, pp. 4–15, Sep. 2013.
[39] S. Morosi, S. Jayousi, E. Falletti, and G. Araniti, "Cooperative strategies in satellite assisted emergency services: COOPERATIVE STRATEGIES IN SATELLITE ASSISTED EMERGENCY SERVICES," *Int. J. Satell. Commun. Netw.*, vol. 31, no. 3, pp. 141–156, May 2013.
[40] T. Bruggen and P. Vary, "Analysis of digital modulation with unequal power allocation," in *2005 IEEE 61st Vehicular Technology Conference*, 2005, vol. 2, p. 1143–1147 Vol. 2.
[41] "Optimized APSK bit allocation for satellite communication - IEEE Xplore Document." [Online]. Available: http://ieeexplore.ieee.org/document/5586898/.
[42] "On the genetic optimization of APSK constellations for satellite broadcasting - IEEE Xplore Document." [Online]. Available: http://ieeexplore.ieee.org/document/6873465/.
[43] "A genetic algorithm tutorial - Whitley94.pdf." [Online]. Available: http://sci2s.ugr.es/sites/default/files/files/linksInterest/Tutorials/Whitley94.pdf.
[44] M. Lixia, M. Murroni, and V. Popescu, "PAPR reduction in multicarrier modulations using Genetic Algorithms," in *2010 12th International Conference on Optimization of Electrical and Electronic Equipment*, 2010, pp. 938–942.

[45] “Optimization of Non-Convex Multiband Cooperative Sensing With Genetic Algorithms - IEEE Xplore Document.” [Online]. Available: http://ieeexplore.ieee.org/document/5492140/.
[46] “Flexible Serially Concatenated Convolutional Turbo Codes with Near-Shannon Bound Performance for Telemetry Applications - CCSDS 131.2-O-1.pdf.” [Online]. Available: http://mtc-m16c.sid.inpe.br/col/sid.inpe.br/mtc-m18@80/2009/07.22.19.43/doc/CCSDS%20131.2-O-1.pdf?metadatarepository=&mirror=sid.inpe.br/mtc-m18@80/2008/03.17.15.17.24.
[47] “RECOMMENDATION ITU-R BS.1387-1 - Method for objective measurements of perceived audio quality - R-REC-BS.1387-1-200111-I!!PDF-E.pdf.” [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1387-1-200111-I!!PDF-E.pdf.
[48] M. Salovarda, I. Bolkovac, and H. Domitrovic, “Estimating Perceptual Audio System Quality Using PEAQ Algorithm,” in *2005 18th International Conference on Applied Electromagnetics and Communications*, 2005, pp. 1–4.
[49] A. A. M. Saleh, “Frequency-Independent and Frequency-Dependent Nonlinear Models of TWT Amplifiers,” *IEEE Trans. Commun.*, vol. 29, no. 11, pp. 1715–1720, Nov. 1981.
[50] “EL470 Satellite Modem - Product - Newtec.” [Online]. Available: http://www.newtec.eu/product/el470-satellite-modem.
[51] E. Casini, R. D. Gaudenzi, and A. Ginesi, “DVB-S2 modem algorithms design and performance over typical satellite channels,” *Int. J. Satell. Commun. Netw.*, vol. 22, no. 3, pp. 281–318, May 2004.
[52] R. De Gaudenzi, A. Guillen i Fabregas, and A. Martinez, “Performance analysis of turbo-coded APSK modulations over nonlinear satellite channels,” *IEEE Trans. Wirel. Commun.*, vol. 5, no. 9, pp. 2396–2407, Sep. 2006.
[53] “Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014–2019 White Paper,” *Cisco*. [Online]. Available: http://cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html. [Accessed: 08-Dec-2015].
[54] M. Murroni, “A power-based unequal error protection system for digital cinema broadcasting over wireless channels,” *Signal Process. Image Commun.*, vol. 22, no. 3, pp. 331–339, Mar. 2007.
[55] “ETSI - UMTS,” *ETSI*. [Online]. Available: http://www.etsi.org/technologies-clusters/technologies/mobile/umts. [Accessed: 12-Nov-2015].
[56] “LTE.” [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/98-lte. [Accessed: 12-Nov-2015].
[57] “IEEE SA - 802.11-2012 - IEEE Standard for Information technology--Telecommunications and information exchange between systems Local and metropolitan area networks--Specific requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.” [Online]. Available: https://standards.ieee.org/findstds/standard/802.11-2012.html. [Accessed: 12-Nov-2015].
[58] F. Zhou, L. Feng, P. Yu, and W. Li, “Energy-efficiency driven load balancing strategy in LTE-WiFi interworking heterogeneous networks,” 2015, pp. 276–281.
[59] M. Fadda, M. Murroni, C. Perra, and V. Popescu, “TV white spaces exploitation for multimedia signal distribution,” *Signal Process. Image Commun.*, vol. 27, no. 8, pp. 893–899, Sep. 2012.
[60] M. Sanna and M. Murroni, “Nonconvex Optimization of Collaborative Multiband Spectrum Sensing for Cognitive Radios with Genetic Algorithms,” *Int. J. Digit. Multimed. Broadcast.*, vol. 2010, pp. 1–12, 2010.
[61] J. D. Martínez-Morales, U. Pineda-Rico, and E. Stevens-Navarro, “Performance comparison between MADM algorithms for vertical handoff in 4G networks,” in *2010 7th International Conference on Electrical Engineering Computing Science and Automatic Control (CCE)*, 2010, pp. 309–314.
[62] R. Trestian, O. Ormond, and G.-M. Muntean, “Performance evaluation of MADM-based methods for network selection in a multimedia wireless environment,” *Wirel. Netw.*, vol. 21, no. 5, pp. 1745–1763, Jul. 2015.
[63] T. Bi, R. Trestian, and G.-M. Muntean, “RLoad: Reputation-based load-balancing network selection strategy for heterogeneous wireless environments,” 2013, pp. 1–3.
[64] T. Bi, Z. Yuan, R. Trestian, and G. M. Muntean, “URAN: Utility-based reputation-oriented access network selection strategy for HetNets,” in *2015 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, 2015, pp. 1–6.

[65] T. Bi, R. Trestian, and G. M. Muntean, "Reputation-based network selection solution for improved video delivery quality in heterogeneous wireless network environments," in *2013 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2013, pp. 1–8.

[66] Y. Luo, B. Qi, Y. Chen, and L. Tang, "Strengthen-Gray Relative Analysis access selection algorithm based on load balance in heterogeneous wireless networks," in *2012 Eighth International Conference on Natural Computation (ICNC)*, 2012, pp. 843–847.

[67] L. Militano, M. Condoluci, G. Araniti, A. Molinaro, A. Iera, and G. M. Muntean, "Single Frequency-Based Device-to-Device-Enhanced Video Delivery for Evolved Multimedia Broadcast and Multicast Services," *IEEE Trans. Broadcast.*, vol. 61, no. 2, pp. 263–278, Jun. 2015.

[68] Z. Yang, Q. Yang, F. Fu, and K. S. Kwak, "A novel load balancing scheme in LTE and WiFi coexisted network for OFDMA system," in *2013 International Conference on Wireless Communications Signal Processing (WCSP)*, 2013, pp. 1–5.

[69] R. K. Jain, D. W. Chiu, and W. R. Have, "A Quantitative Measure Of Fairness And Discrimination For Resource Allocation In Shared Computer Systems - fairness.pdf," 1984.

[70] K. Azarian, R. Patwardhan, C. Lott, D. Ghosh, R. Gowaikar, and R. Attar, "On Spatial Load Balancing in wide-area wireless networks," in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, 2012, pp. 2230–2235.

[71] L. Wang, W. Huang, Y. Fan, and X. Wang, "Priority-based cell selection for mobile equipments in heterogeneous cloud radio access networks," in *2015 International Conference on Connected Vehicles and Expo (ICCVE)*, 2015, pp. 62–67.

[72] M. Liu and P. Papantoni-Kazakos, "A random-access algorithm for data networks carrying high-priority traffic," *IEEE Trans. Commun.*, vol. 40, no. 1, pp. 84–96, Gennaio 1992.

[73] A. Burrell and P. Papantoni-Kazakos, "Priority Users in a Multi-Channel System," in *2006 40th Annual Conference on Information Sciences and Systems*, 2006, pp. 1218–1223.

[74] Z. Yuan and G.-M. Muntean, "A Prioritized Adaptive Scheme for Multimedia Services over IEEE 802.11 WLANs," *IEEE Trans. Netw. Serv. Manag.*, vol. 10, no. 4, pp. 340–355, Dec. 2013.

[75] Janghoon Yang, "A Semi-Markov Decision Process (SMDP) for Active State Control of A Heterogeneous Network," *KSII Trans. Internet Inf. Syst.*, vol. 10, no. 7, pp. 3171–3191, Luglio 2016.

[76] K. Jakimoski and T. Janevski, "Radio Access Technology Selection and Vertical Handover Decision Algorithms for Heterogeneous Mobile and Wireless Networks," *Adhoc Sens. Wirel. Netw.*, vol. 27, no. 1/2, pp. 95–109, Luglio 2015.

[77] E. Khloussy, X. Gelabert, and Y. Jiang, "Investigation on MDP-based radio access technology selection in heterogeneous wireless networks," *Comput. Netw.*, vol. 91, pp. 57–67, Nov. 2015.

[78] "IEEE SA - 802.11n-2009 - IEEE Standard for Information technology-- Local and metropolitan area networks-- Specific requirements-- Part 11: Wireless LAN Medium Access Control (MAC)and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput." [Online]. Available: https://standards.ieee.org/findstds/standard/802.11n-2009.html. [Accessed: 22-Mar-2016].

[79] "3GPP Long Term Evolution (LTE): Quality of Service (QoS) in LTE," 18-Apr-2016. [Online]. Available: http://4g-lte-world.blogspot.it/2013/01/quality-of-service-qos-in-lte.html. [Accessed: 18-Apr-2016].

[80] R. K. Jain, D. W. Chiu, and W. R. Have, "A Quantitative Measure Of Fairness And Discrimination For Resource Allocation In Shared Computer Systems - fairness.pdf," 1984.

[81] "LTE ue-Category." [Online]. Available: http://www.3gpp.org/keywords-acronyms/1612-ue-category. [Accessed: 22-Mar-2016].

[82] "G.711 : Pulse code modulation (PCM) of voice frequencies," 22-Mar-2016. [Online]. Available: https://www.itu.int/rec/T-REC-G.711. [Accessed: 22-Mar-2016].

[83] "RFC 3984 - RTP Payload Format for H.264 Video." [Online]. Available: http://tools.ietf.org/html/rfc3984#page-2. [Accessed: 22-Mar-2016].

[84] "Video Encoding Settings for H.264 Excellence," 22-Mar-2016. [Online]. Available: http://www.lighterra.com/papers/videoencodingh264/. [Accessed: 22-Mar-2016].

[85] T. R. Henderson, M. Lacage, G. F. Riley, C. Dowell, and J. Kopena, "Network simulations with the ns 3 simulator," presented at the SIGCOMM'08, Seattle, Washington, USA, 2008.

[86] H. S. Chang, J. Hu, M. C. Fu, and S. I. Marcus, "Markov Decision Processes," in *Simulation-Based Algorithms for Markov Decision Processes*, Springer London, 2013, pp. 1–17.
[87] "G.711 : Pulse code modulation (PCM) of voice frequencies." [Online]. Available: https://www.itu.int/rec/T-REC-G.711. [Accessed: 22-Mar-2016].
[88] "Video Encoding Settings for H.264 Excellence." [Online]. Available: http://www.lighterra.com/papers/videoencodingh264/. [Accessed: 22-Mar-2016].
[89] "The WebM Project | VP9 Video Codec Summary." [Online]. Available: https://www.webmproject.org/vp9/. [Accessed: 04-Oct-2016].
[90] "3GPP Long Term Evolution (LTE): Quality of Service (QoS) in LTE." [Online]. Available: http://4g-lte-world.blogspot.it/2013/01/quality-of-service-qos-in-lte.html. [Accessed: 18-Apr-2016].