UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

RAFAEL KUNST

A QoS-Aware Resources Sharing Architecture for Homogeneous and Heterogeneous Wireless Networks

Thesis presented in partial fulfillment
of the requirements for the degree of
Doctor of Computer Science

Advisor: Prof. Dr. Juergen Rochol

Porto Alegre
July 2017

CIP – CATALOGING-IN-PUBLICATION

Kunst, Rafael

A QoS-Aware Resources Sharing Architecture for Homogeneous and Heterogeneous Wireless Networks / Rafael Kunst. – Porto Alegre: PPGC da UFRGS, 2017.

139 f.: il.

Thesis (Ph.D.) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2017. Advisor: Juergen Rochol.

1. Resources sharing. 2. Wireless networks. 3. Resources broker. I. Rochol, Juergen. II. Título.

ACKNOWLEDGMENTS

Addis Ababa, Ethiopia, November 13th, 2016.

The ideas for this acknowledgments text are in my mind for a long time. I am not sure why, but now, during an extended layover in Ethiopia, while returning from IEEE LCN conference, it seems to be the right moment to turn the ideas into words. I hope I can find the right words to express all the emotions involved in the process that leaded to the production of this thesis. Consequently, I expect to properly thank those persons who were important or indispensable during the whole process.

The epigraph that you will read within this volume is extracted from the so called Arts of War, written by Sun Tzu and was a constant source of motivation. Every time I felt unmotivated, I used to read the part of the book which says:

> "If you know the enemy and know yourself, you need not fear the result of a hundred battles. If you know yourself but not the enemy, for every victory gained you will also suffer a defeat. If you know neither the enemy nor yourself, you will succumb in every battle."

This quote was a constant reminder that I knew myself but not always all the "enemies" which took my attention off the production of this thesis. Many people helped keeping me focused and deserve to be acknowledged.

First of all, I would like to thank my family, especially Carol, my future wife, who always believed in my potential and was present during the hardest moments to keep me motivated. I also want to acknowledge those people who worked with me during my entire academic life. All my professors, colleagues at UFRGS and Unilasalle, and friends who are also into the academic life are included in this group. It is indispensable, however, to name some people: my advisor, Juergen Rochol, professors Bampi and Edison, thank you for sharing your experience with me. Leandro and Jean, thanks for working hard on the production of papers derived from this thesis. Alexandra and Mariana, thank you both for all the discussions regarding the PhD student life.

AGRADECIMENTOS

Addis Abeba, Etiópia, 13 de novembro de 2016.

As ideias para este texto de agradecimento estão em minha mente há algum tempo. Não sei explicar exatamente por que, mas este momento, durante uma longa escala na Etiópia, ao retornar do LCN, parece ser o momento certo para transformar ideias em palavras. Espero poder encontrar as palavras certas para expressar todas as emoções envolvidas no processo que conduziu à produção desta tese. Consequentemente, espero agradecer adequadamente às pessoas que foram importantes ou imprescindíveis durante todo o processo.

A epígrafe que você vai ler neste volume foi extraída do livro Arte da Guerra, escrito por Sun Tzu e foi uma fonte constante de motivação. Toda vez que me sentia desmotivado, costumava ler a parte do livro que diz:

> "Se você conhece o inimigo e se conhece, não precisa temer o resultado de cem batalhas. Se você conhece a si mesmo, mas não o inimigo, para cada vitória obtida, você também sofrerá uma derrota. Se você não conhece nem o inimigo nem você mesmo, você sucumbirá em todas as batalhas ".

Esta citação foi um lembrete constante de que eu conhecia a mim mesmo, mas nem sempre conhecia a todos os "inimigos" que desviaram minha atenção da produção desta tese. Muitas pessoas ajudaram a manter-me focado e merecem ser reconhecidas por isso.

Em primeiro lugar, gostaria de agradecer a minha família, especialmente à Carol, minha futura esposa, que sempre acreditou no meu potencial e estava presente nos momentos mais difíceis para me manter motivado. Eu também quero citar aquelas pessoas que trabalharam comigo durante toda a minha vida acadêmica. Todos os meus professores, colegas na UFRGS e no Unilasalle, e amigos que também estão na vida acadêmica estão incluídos neste grupo.

No entanto, é indispensável citar algumas pessoas: meu orientador, Juergen Rochol, professores Bampi e Edison, obrigado por compartilhar sua experiência comigo. Leandro e Jean, obrigado por trabalharem muito na produção dos artigos derivados desta tese. Alexandra e Mariana, obrigado por todas as discussões sobre a vida do estudante de doutorado.

"If you now your enemy and yourself,
you will succeed in every battle".
— SUN TZU.

ABSTRACT

The static model currently applied by governmental authorities for allocating the spectrum of frequencies and the increasing demand for network resources imposed by modern applications and services may lead to a resources scarcity problem in the near future. Dealing with this problem demands improvements on resources allocation. One of the ways of providing such improvements is by allowing resources sharing among network operators in both homogeneous and heterogeneous network scenarios. These network operators may implement different technologies, such as collective use of spectrum and licensed shared access to the spectrum of frequencies.

Many related works have been proposed in the same context of the presented research, however these related works generally identify the need for additional resources and search for available resources without taking into account the QoS requirements of the resources renter and the costs involved in the resources sharing initiative. Therefore, in this thesis, a novel architecture is proposed to facilitate the implementation of resources sharing and consequently encourage network operators to lease their underutilized resources taking into account both the cost and the QoS requirements. This approach allows the network operator which is serving the resources to improve its profits at the same time that allows quality of service improvements to the resources renter.

The main contributions of the proposed architecture include but are not limited to the design of a multilevel resources broker to control the resources sharing process. This broker is concerned on dynamically establishing a service level agreement that takes into account the quality of service requirements of resources renter. This process focuses on exchanging a small amount of control information to prevent the overhead from interfering with the legitimate traffic of the network operators. Another important contribution of the proposed approach is to improve the resources allocation in comparison with related work. Furthermore, the proposed solution is capable of taking fast decisions regarding resources allocation, what leads to the implementation of fast handover, allowing the traffic steering without interfering with incumbent users.

The proposed architecture is modeled analytically and simulated using Matlab to evaluate its behavior in three different scenarios, considering both homogeneous and heterogeneous networks. The overhead in practical operation scenarios is kept under 1% of the total network traffic, what is considered not to interfere with the transmissions of the network operators. The fast decisions taken by the resources sharing architecture are based on accurate traffic load forecasting, what leads to fast handover, attaining times up to 46% lower than the maximum allowed handover duration. Results also show that both delay and jitter metrics are controlled to be maintained below their specific thresholds of the analyzed applications and therefore, the QoS is guaranteed for the resources renter, considering the coexistence of up to 500 devices.

Keywords: Resources sharing. wireless networks. resources broker.

Arquitetura para Compartilhamento de Recursos com QoS em Redes Sem Fio Homogêneas e Heterogêneas

RESUMO

O atual modelo de alocação espectral implementado pelas autoridades governamentais somado à crescente demanda por recursos imposta pela implementação de modernas aplicações e serviços de rede irá resultar em um problema relacionado à escassez de recursos em um futuro próximo. Lidar com este problema demanda esforços no sentido de melhorar a alocação de recursos. Uma das maneiras de atingir este tipo de melhoria é permitir o compartilhamento de recursos entre operadores em redes homogêneas e heterogêneas que podem implementar diferentes tecnologias, como a utilização coletiva do espectro e de recursos licenciados.

Diversos trabalhos relacionados à esta pesquisa foram propostos. Entretanto, estes trabalhos geralmente identificam a necessidade de obter recursos adicionais, porém buscam por esses recursos sem levar em conta os requisitos de qualidade de serviço e o custo envolvido no compartilhamento desses recursos. Considerando esse contexto, nesta tese, uma nova arquitetura é proposta para permitir a implementação do compartilhamento de recursos e para encorajar operadores a alugarem recursos sobressalentes levando em conta o custo e a qualidade de serviço oferecida. Esta abordagem permite que operadores tenha ganhos com o aluguel dos recursos, ao mesmo tempo em que o cliente recebe serviços com maior qualidade.

As principais contribuições da arquitetura proposta incluem o projeto de um controlador de recursos para coordenar o processo de compartilhamento. Esse controlador busca estabelecer contratos de serviço dinâmicos levando em conta a qualidade de serviço requerida. Para tanto, é necessária a troca de informações que, no caso da arquitetura proposta, é mantida baixa para evitar que a rede seja sobrecarregada e acabe interferindo com o tráfego de dados. Além disso, a solução proposta é capaz de tomar decisões rápidas sobre a alocação de recursos, o que permite o redirecionamento do tráfego sem que ocorram interferências com os demais usuários.

A arquitetura proposta é modela analiticamente e simulada com o auxílio da ferramenta Matlab. O desempenho da proposta é medido em três diferentes cenários, considerando tanto redes homogêneas, quanto heterogêneas. A sobrecarga gerada pela troca de informações de controle corresponde a menos de 1% do tráfego total da rede, o que é desprezível do ponto de vista da interferência com o tráfego de dados. As decisões rápidas tomadas pela arquitetura são baseadas na previsão acurada do tráfego futuro da rede e permitem o redirecionamento do tráfego para outras redes em um tempo até 46% abaixo do limite máximo especificado na literatura para este tipo de redirecionamento. Os resultados mostram ainda que as métricas de atraso e variação do atraso também são mantidas abaixo dos limites especificados, o que indica que a qualidade de serviço é garantida nos cenários avaliados.

Palavras-chave: compartilhamento de recursos, redes sem fio heterogêneas, *broker*.

LIST OF FIGURES

2.1 Resources Sharing Taxonomy . . . 21
2.2 Inter-network Signaling . . . 27
2.3 Cognitive Functions . . . 29
2.4 Spectrum Sensing Dimensions . . . 29
2.5 Centralized Spectrum Sharing . . . 31
2.6 Distributed Spectrum Sharing . . . 32
2.7 Neural Network . . . 37

4.1 Architecture Design . . . 49
4.2 Simulation Tool Use Case Diagram . . . 64

5.1 ON/OFF Behavior Markov Chain . . . 67
5.2 GQM model of the proposed solution . . . 72
5.3 Underutilized resources considering a total frame capacity of 113Mbps . . . 74
5.4 Frame capacity . . . 75
5.5 Throughput behavior with variable number of connections . . . 77
5.6 Delay Comparison . . . 78
5.7 Percentage of collisions with variable number of connections . . . 79
5.8 Update Overhead . . . 80
5.9 Request Overhead . . . 80
5.10 Actual vs. Predicted Traffic Load for each Wi-Fi Network . . . 83
5.11 Actual vs. Predicted Traffic Load for each class of service in LTE-LSA Network . . . 83
5.12 Video and VoIP Traffic Steering from LTE-LSA to Wi-Fi 2 . . . 84
5.13 HTTP Traffic Steering from LTE-LSA to Wi-Fi 3 . . . 84
5.14 Average delay for differente Wi-Fi networks . . . 85
5.15 Video Surveillance IoT Architecture . . . 87
5.16 Traffic load in each MNO - IoT scenario . . . 88
5.17 Average Delay in each MNO - IoT scenario . . . 89
5.18 Average Jitter in each MNO - IoT scenario . . . 89

B.1 Current vs. Predicted Traffic Load for each Wi-Fi Network with the Multiple Linear Regression Model . . . 105
B.2 Current vs. Predicted Traffic Load for each CoS in the LTE-LSA Network with the Multiple Linear Regression Model . . . 106

B.3 Current vs. Predicted Traffic Load for each Wi-Fi Network using the Neural Network Model 107
B.4 Current vs. Predicted Traffic Load for each CoS in the LTE-LSA Network using the Neural Network Model 108
B.5 Current vs. Predicted Traffic Load for each Wi-Fi Network using the Regression Tree Model 109
B.6 Current vs. Predicted Traffic Load for each CoS in the LTE-LSA Network using the Regression Tree Model 110
B.7 Current vs. Predicted Traffic Load for each Wi-Fi Network using the Fourier Model 110
B.8 Current vs. Predicted Traffic Load for each CoS in the LTE-LSA Network using the Fourier Model 111

LIST OF TABLES

3.1 Set of Considered Publications 40
3.2 Selected papers per publication 42
3.3 10 most relevant papers 42
3.4 Related Works Analysis 46
3.5 Comparison between this proposed approach and related work 47

4.1 Raw Update Data Structure 50
4.2 Resources Request Structure 53
4.3 Classes of Services of the Proposed Architecture 55
4.4 Tickets Database Table Structure 56
4.5 Parameters used to calculate the cost 59
4.6 Scenario Configuration File Structure 63

5.1 VoIP traffic parameters 68
5.2 Video traffic parameters 68
5.3 HTTP traffic parameters 69
5.4 Simulation Setup Parameters 71
5.5 Evacuation Duration 86

A.1 SLR papers set 103

B.1 Average Time for Traffic Load Forecasting in Wi-Fi Networks 112
B.2 Mean Absolute Percentage Error of Wi-Fi Traffic Load Forecasting Models 112
B.3 Average Time for Traffic Load Forecasting for Each CoS in the LTE-LSA Network 112
B.4 Mean Absolute Percentage Error of Traffic Load Forecasting Models for each CoS in the LTE-LSA Network 113

LIST OF ABBREVIATIONS AND ACRONYMS

3GPP	3rd Partnership Project
AMR	Adaptive Multi Rate
BS	Base Station
BES	Best Effort Service
CoS	Class of Service
CRN	Cognitive Radio Networks
CORE	Cognitive Radio Trial Environment
CUS	Collective Use of Spectrum
CAC	Connection Admission Control
CPE	Customer Premises Equipment
DSA	Dynamic Spectrum Access
ETSI	European Telecommunication Standards Institute
ESA	Exclusive Spectrum Access
FIFO	First In First Out
FDD	Frequency Division Duplexing
GQM	Goals/Questions/Metrics
HSS	Home Subscriber Service
ISM	Industrial, Scientific, and Medical
IEEE	Institute of Electrical and Electronics Engineers
IMT	International Mobile Telecommunication
ITU	International Telecommunication Union
IoT	Internet of Things
LSA	Licensed Shared Access
M2M	Machine-to-Machine
MAC	Medium Access Control
MNO	Mobile Network Operator

MME	Mobility Management Entity
MS	Multimedia Service
MLRM	Multiple Linear Regression Model
NRA	Network Resources Administrator
NN	Neural Network
PDN	Packet Data Network
P-GW	Packet Data Network Gateway
PHY	Physical
PCRF	Policy Control and Charging Rules Function
QoS	Quality of Service
RAN	Radio Access Network
RTS	Real Time Service
RTM	Regression Tree Model
SLA	Service Level Agreement
S-GW	Serving Gateway
SINR	Signal to Interference Plus Noise Ratio
SDR	Software Defined Radi
SLR	Systematic Literature Review
TDD	Time Division Duplexing
UMTS	Universal Mobile Telecommunication System
VoIP	Voice over IP
WFQ	Weighted Fair Queuing
CFBP	Cascade Forward Back Propagation

CONTENTS

1 INTRODUCTION 15

1.1 Research Problem and Motivation 15

1.2 Objectives and Contributions 16

1.3 Results Summary 17

1.4 Outline of the Thesis 18

2 RESOURCES SHARING IN HETEROGENEOUS WIRELESS CELLULAR NETWORKS 20

2.1 Background on Resources Sharing 20

2.1.1 Mobile Network Operators Resources Sharing 20

2.1.2 Collective Use of Spectrum 22

2.1.3 Exclusive Spectrum Access 23

2.1.4 Licensed Shared Access 23

2.2 Background on Heterogeneous Wireless Networks 25

2.2.1 Inter-Network Communication and Signaling 26

2.2.2 Cognitive Wireless Networks 28

2.2.3 Traffic Load Forecasting 35

3 RELATED WORK 39

3.1 Definition of Research Questions 39

3.2 Search for Suitable Papers 39

3.3 Applicability Assessment 40

3.4 Classification of Selected Papers 41

3.5 Related Work Analysis 43

3.6 Related Work Summary 46

4 AN ARCHITECTURE FOR RESOURCES SHARING IN WIRELESS CELLULAR NETWORKS 48

4.1 Design of the Resources Sharing Architecture 48

4.1.1 Update Level 49

4.1.2 Resources Level 51

4.1.3 Decision Level 52

4.2 Simulation Tool 63

5 PERFORMANCE EVALUATION . . . 66
5.1 Simulation Scenario . . . 66
5.1.1 Traffic Models . . . 66
5.1.2 Simulation Parameters Setup . . . 70
5.2 Case Studies . . . 71
5.2.1 LTE-Advanced Operators Coexisting in the same Geographical Area . . . 73
5.2.2 Fast Handover in heterogeneous network scenarios . . . 81
5.2.3 Video Surveillance in Smart Cities . . . 86

6 CONCLUSION AND FUTURE WORK . . . 91
6.1 Summary of Contributions . . . 91
6.2 Final Remarks and Suggestions for Future Work . . . 92

APPENDIX A ORIGINAL SLR PAPERS SET . . . 93

APPENDIX B TRAFFIC LOAD FORECASTING MODELS EVALUATION . 104
B.1 Evaluation of Traffic Load Forecasting Models . . . 104
B.1.1 Traffic Load Forecasting with the Multiple Linear Regression Model . . . 104
B.1.2 Traffic Load Forecasting using the Neural Network Model . . . 105
B.1.3 Traffic Load Forecasting using the Regression Tree Model . . . 108
B.1.4 Traffic Load Forecasting using the Fourier Model . . . 109
B.2 Comparison Table of Performance Metrics for each Traffic Load Forecasting Model . . . 111

APPENDIX C PUBLISHED PAPER – IEEE WIMOB 2016 . . . 114

APPENDIX D PUBLISHED PAPER – IEEE LCN 2016 . . . 123

APPENDIX E PUBLISHED PAPER – IEEE ICC 2017 . . . 128

REFERENCES . . . 135

1 INTRODUCTION

The traffic generated by mobile network operators is constantly growing and by 2020 it is expected to overload the existing licensed spectrum (CISCO, 2016). This scenario in conjunction with the current spectrum allocation model, which prioritizes main network operators who purchase exclusive rights over portions of the spectrum of frequencies, may lead to a network resources scarcity problem. Dealing with this resources scarcity problem is especially challenging in situations in which Mobile Network Operators (MNOs) aim to guarantee quality of service (QoS) to their customers. Based on this context, in Section 1.1, the research problem and the motivations of this thesis are discussed.

1.1 Research Problem and Motivation

An emerging approach to increase the amount of resources available to wireless MNOs is to encourage resources sharing. In this thesis, the concept of resources sharing is considered in two different scenarios, regarding to the technologies involved in the sharing initiative. The first one is called homogeneous and comprises the collaboration among operators which implement the same technology, *e.g.* only LTE-Advanced or IEEE 802.11 devices participate in the resources sharing initiative. The second scenario is called heterogeneous, since it allows the resources sharing in network environments composed of multiple technologies, *e.g.* LTE-Advanced and IEEE 802.11 devices coexisting in the same geographical area.

The main motivation of the novel solution proposed in this thesis is to deal with the spectrum scarcity problem by implementing QoS-aware resources sharing in both homogeneous and heterogeneous network scenarios. This motivation is based on the premise that both commercial and non-commercial wireless network operators have a certain amount of unused resources, especially during off-peak hours. By sharing unused resources with coexisting operators in the same geographical area, a MNO can improve its profit, while the resources renter is able to improve the QoS offered to its customers (KUNST et al., 2016a).

The hypothesis tested in this thesis is that it is possible to implement a resources sharing mechanism which deals with the resources scarcity problem and guarantees the QoS offered to the wireless cellular networks customers in both homogeneous and heterogeneous scenarios. The following research questions are defined in order to confirm or refute the hypothesis.

1. How to gather the information needed to manage the resources without causing high impact on the network control traffic?
2. How to take time sensitive resources allocation decisions without interfering with the incumbent users traffic?
3. What is the best approach to allocate on demand traffic in heterogeneous network scenario?

Many related works have been proposed in the same context of the presented research questions. These works typically consider the resources availability by analyzing the spectrum of frequencies in two different ways. In situations in which licenses are not required for accessing a given frequency range, the concept of Collective Use of Spectrum (CUS) is applied. Otherwise, when licenses are necessary, Exclusive Spectrum Access (ESA) or Licensed Shared Access (LSA) are considered to implement resources sharing (KUNST et al., 2016b)(FERNANDEZ; KUNST; ROCHOL, 2017). Most of the current research, although very relevant, generally identify the need for additional resources and search for available resources without taking into account the QoS requirements of the resources renter and the costs involved in the resources sharing initiative. This common approach of related works may lead to the allocation of network resources which are not suitable for the MNOs.

Considering the previously defined research questions and the aspects not fully covered by the related works, in this thesis, a novel resources sharing architecture is presented to deal with the spectrum scarcity problem by implemented a QoS-aware approach. The main component of the proposed architecture is a multilevel broker, which is implemented to allow resources sharing among MNOs which coexist in a given geographical area. This broker is able to operate considering both homogeneous and heterogeneous network environments. QoS requirements of the resources renter are considered via the establishment of a dynamic Service Level Agreement (SLA). This multilevel broker provides interfaces, a control protocol, and cognitive decision algorithms to allow communication and resources sharing involving different MNOs. The control protocol has been designed to exchange a small amount of information to prevent the overhead from significantly interfering with the network traffic.

The proposed architecture is modeled analytically and simulated using Matlab to evaluate its behavior in both homogeneous and heterogeneous scenarios. The results, when applicable, are compared to related works in order to evaluate the relevance of the proposed approach. Considering the discussed context, the specific objectives and contributions of this thesis are presented in section 1.2.

1.2 Objectives and Contributions

Two classes of objectives are defined in this thesis. The first one refers to those objectives related to the design principles of the proposed architecture. This first kind of objectives are defined as follows.

- Design and implementation of a resources sharing architecture which allows resources sharing involving network operators which coexist in the same geographical area;
- Design of a solution that allows the sharing of different kinds of network resources, for example, spectrum and capacity to allow heterogeneous network operators to share resources;

- To allow resources sharing in different types of allocation regimes, for example, collective use of spectrum and licensed shared spectrum access;
- Proposal of QoS-aware decision algorithms which allow fast and reliable selection of network operators with available network resources;
- Proposal of a model to estimate the cost of resources sharing considering the priority and the current load of the resources renter;

The second class of objectives is related to the expected performance of the proposed architecture. These objectives are:

- Contribute with the state of the art on resources sharing by improving the results obtained by related research;
- Keep the overhead generated by the proposed architecture low to avoid interfering with the legitimate traffic of the MNOs;
- To provide a solution which allows time sensitive resources allocation;
- Analyze the performance of the proposed architecture in both homogeneous and heterogeneous scenarios with variable amount of MNOs, variable number of connections, and variable wireless channel conditions.

After attaining the aforementioned objectives, the main expected contributions of this thesis are:

- A novel resources sharing architecture which improves resources sharing efficiency in comparison with related work;
- To offer QoS guarantees to wireless MNO customers in homogeneous and heterogeneous network resources sharing scenarios;
- To propose algorithms to implement fast resource allocation decisions to avoid interference;
- A low overhead resources sharing control mechanism.

A brief summary of the results obtained in this thesis is presented in Section 1.3.

1.3 Results Summary

Three simulation scenarios have been defined to evaluate the performance of the proposed resources sharing architecture, each one to address one research question. To deal with research question number 1, a homogeneous network scenario composed of LTE MNOs is proposed. In this scenario, the mechanism used for gathering information from the MNOs which participate in the resources sharing initiative is evaluated. The outcomes of the simulations show that the overhead is sufficiently low to avoid interfering with the network traffic.

Research question 2 is answered in the second case study. In this case, a fast handover heterogeneous network scenario is considered. In order to implement the fast handover, network traffic is predicted to allow the provisioning algorithm to take in advance decisions regarding resources sharing. Multiple Linear Regression Model (MLRM) is implemented due to the level of accuracy in traffic forecasting. The implementation of such model leads to fast handover, attaining times up to 46% lower than the maximum allowed handover duration.

The third question is addressed in the last evaluation case, in which Internet of Things (IoT) sensors are used to implement a video surveillance solution destined to monitor countries borders. This solution works in real time and thus demands strict delay and jitter QoS requirements. In order to evaluate the performance of the proposed resources broker in this situation, five MNOs are available in a heterogeneous network environment. Results show that both delay and jitter metrics are kept below their specific thresholds for video applications and therefore, the QoS is guaranteed for the resources renter, considering the coexistence of up to 500 sensors.

The organization of the remainder of this thesis is presented in section 1.4.

1.4 Outline of the Thesis

Chapter 2 focuses on studying different approaches of network resources sharing found in the literature. The first part of the chapter is dedicated to present background aspects on network resource sharing. A resources sharing taxonomy is compiled based on the classifications found in the literature. In this taxonomy, three classes of resources sharing are highlighted: CUS, LSA, and ESA. Details and ramifications of these classes are discussed. The second part of the chapter presents background aspects on heterogeneous wireless networks. In this part, internetworking communication and signaling is discussed. Moreover, cognitive network functions and traffic loading forecasting solutions are depicted in this chapter.

Related works are discussed in Chapter 3, based on a Systematic Literature Review (SLR) conducted in order to identify related research relevant to this thesis. The methodology used to gather and analyze the data is presented and the results of the SLR are discussed. The first step towards the search of relevant papers is the definition of research questions. Based on three research questions, a search for suitable publications resulted in an initial set of 103 papers. These papers passed through an applicability assessment, what led to 31 relevant papers, which were classified according to the SLR methodology and the 10 considered more relevant were analyzed and compared with the solution proposed in this thesis.

In Chapter 4, the design of the network resources sharing architecture is discussed. The multilevel broker receives special attention, since it is the main component of the architecture. The explanation of each level of the broker is provided, as well as the specification of the control structures demanded by the architecture to operate. Resources controller and resources provisioning algorithms used by the broker to take resource allocation decisions are also presented and discussed in this chapter. Another important aspect covered in this chapter is the estimation of the resources sharing cost. Towards this estimation, a formula is proposed based on the priority of the request and the current load of resources server. Finally, the design of the simulation tool used for obtaining results is presented.

The aim of Chapter 5 is to evaluate the performance of the proposed architecture. Towards that aim, the simulation scenarios are presented and discussed. Special focus is given to the traffic models and traces implemented to specify the traffic generated to be classified into the three classes of service supported by the proposed approach. After detailing the simulation scenarios and parameters, the results are presented considering three network scenarios: (I) a homogeneous scenario composed of four LTE MNOs, to reflect the scenario typically found in Brazil, (II) a heterogeneous network scenario composed of LTE costumers operating in LSA frequencies and IEEE 802.11af network operators, and (III) a scenario composed of five different network operators which coexist in the same geographical area. The results show that the proposed approach is able to improve the resources allocation in comparison with related work, while the overhead is kept low. Moreover, the results obtained via simulations show that, considering the analyzed scenarios, the QoS requirements of the resource renters are guaranteed.

Chapter 6 finished this thesis with the presentation of conclusions and contributions of the proposed approach. The aspects not fully covered in this thesis which will constitute future work are also discussed. After this chapter, Appendix A presents details on the SLR presented in Chapter 3. Moreover, Appendix B provides further details on the reasons for the selection of MLRM as the artificial intelligence approach to implement traffic load forecasting. Finally, Appendixes C, D, and E discuss the publications derived from this thesis.

2 RESOURCES SHARING IN HETEROGENEOUS WIRELESS CELLULAR NETWORKS

The aim of this chapter is to provide background on two fundamentals concepts of this thesis. The first one is the concept of resources sharing, which is discussed in Section 2.1. The approach used to explain resources sharing is to divide the subject into three groups according to the type of sharing. First, CUS is discussed in 2.1.2. The concepts related to ESA are detailed in subsection 2.1.3. The last aspect related to resources sharing background is LSA, which is depicted in 2.1.4. The second main background concept discussed in this thesis regards to heterogeneous wireless networks. In this case, three main aspects are taken into account: (I) inter-network communication and signaling, (II) cognitive wireless networks, and (III) traffic load forecasting. These aspects are depicted in subsections 2.2.1, 2.2.2, and 2.2.3, respectively.

2.1 Background on Resources Sharing

Traffic demand is increasing due to the characteristics of wireless services and applications. As a consequence, the demand for network resources is also increasing. A report published by the Universal Mobile Telecommunication System (UMTS) Forum showed that such demand has been forecast to increase over 30 times before 2020 (UMTS Forum, 2011a). Considering the future scenario, the availability of exclusive network resources tends to become scarce. A potential solution to deal with the consequent network resources scarcity is to allow opportunistic resources sharing among different MNOs. Background aspects regarding MNOs resources sharing are provided in subsection 2.1.1.

Network resources can be shared among different MNOs according to three main approaches. The first one is called CUS because specific licenses are not demanded to allow devices access to network resources. On the other hand, there are two scenarios in which licenses are necessary to provide network resources access: (I) ESA, which is discussed in 2.1.3, and (II) LSA, depicted in 2.1.4. These approaches and their ramifications are compiled into a taxonomy show in Figure 2.1.

2.1.1 Mobile Network Operators Resources Sharing

A MNO can be considered as a mobile broadband services provider in the context of wireless communications. Typically, this kind of operator aims at providing high QoS Internet access to its customers. Considering the current overload of spectrum resources, to achieve the aim of providing QoS-enabled services, the MNOs may deploy heterogeneous wireless networks that geographically coexist and operate in different frequency bands such as licensed, LSA, and CUS frequencies in a harmonized and coordinated manner.

The first choice of a MNO is to provide services using the traditional licensed spectrum

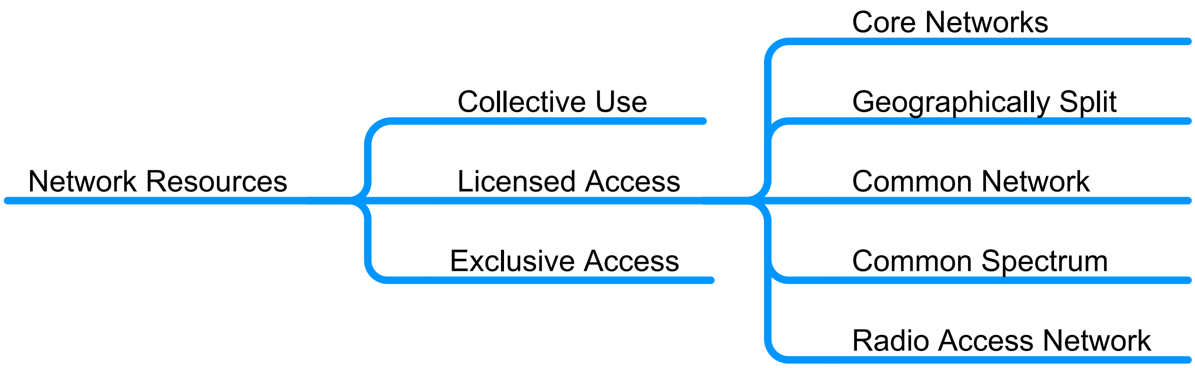


Figure 2.1: Resources Sharing Taxonomy

frequencies. However, research papers and reports in both academic and industrial fields - (4G AMERICAS, 2014) (CISCO, 2016) (J.; W., 2011) (HOSSAIN et al., 2014) (UMTS Forum, 2011a) (WiMAX Forum, 2008) - argue that the amount of traffic expected by 2020 exceeds the capacity of the traditional spectrum resources used by the MNOs. According to Hossain *et al.* (HOSSAIN et al., 2014), the existing wireless network of MNOs will not be able to deal with the substantial increase in total mobile broadband data. By 2020, the wireless communication technologies are expected to have 1000 times higher volume of mobile data and up to 100 times greater number of connected devices (OSSEIRAN et al., 2014) (METIS, 2013).

This current context urges the existence of alternative paths to implement high QoS solutions. In this sense, resources sharing among MNOs is an emerging solution which allows better resources usage efficiency by permitting temporary unoccupied resources (spectrum of frequencies or channel capacity) to be opportunistically leased to secondary users (ECC, 2011). The availability of implementing resources sharing is limited to a certain geographical location and time, however when the implementation is feasible, the MNO can increase its spectrum bandwidth and consequently the available amount of resources.

Different types of resources may be available to the MNOs, covering a range of both micro and macro cells and both spectrum and network capacity resources. These types of resources are available due to the existence of heterogeneous wireless network technologies, which operate in more than one kind of spectrum frequency. The MNO relies on technologies like Frequency Division Duplexing (FDD) and Time Division Duplexing (TDD) to ensure harmonious coexistence between CUS, ESA, and LSA spectrum frequencies.

Practical examples of coexistence among heterogeneous networks are already implemented or under development and testing. In Europe, the regulatory agencies are planning to provide access at a frequency range of 2.3 - 2.4 GHz through the LSA technology. This band is designed for mobile broadband services according to International Mobile Telecommunication (IMT) and International Telecommunication Union (ITU) radio regulations. Moreover, this frequency range is suitable for the application of TDD into the LSA technology in which the Base

Station (BS) implements time division to split downstream and upstream signals. In contrast, FDD is used by the BS on licensed bands where the transmitter and receiver operate at different frequencies. Another example of regulatory effort towards resources sharing is the one recently published by the 3rd Partnership Project (3GPP), which standardized the 2.3 - 2.4 GHz spectrum as LTE band for TDD-LTE spectrum sharing. A third example of successful implementation of resources sharing is described in an European Telecommunication Standards Institute (ETSI) document (ETSI, 2013), which recommends the use of the small cells LSA leverage the bandwidth expansion of MNOs. The small cells approach makes it easier to carry out the scheduled and unscheduled evacuations of users on LSA band.

2.1.2 Collective Use of Spectrum

In CUS regime, devices operate under a general authorization, which means that no license is demanded to access network resources. In this regime, a limited amount of network resources, *e.g.* spectrum frequencies, can be accessed by an unlimited amount of independent devices (MUSTONEN et al., 2014). The access can occur at anytime and anywhere in a limited geographical area with strict configuration conditions, such as a limited transmission power (UMTS Forum, 2011b).

One of the main challenges faced by CUS regime is the unlimited and typically unpredictable amount of devices competing for the available network resources. This situation is obvious for unlicensed technologies, such as those which operate in Industrial, Scientific, and Medical (ISM) frequencies, especially in the 2.4GHz range of frequencies, since many devices are sharing the access to the frequencies, what may lead to mutual interference. However, even in scenarios where a MNO is providing and consequently controlling resources access, for example, when IEEE 802.22 is used, it is not guaranteed that the transmissions will be interference-free, what leads to an uncertainty related to the quality of QoS provided to each device.

Recent research has deeply explored CUS, especially correlating it to the concepts of Software Defined Radio (SDR) and Cognitive Radio Networks (CRN). One of the more important works in this area is proposed by Lee and Akyildiz (AKYILDIZ et al., 2008). The approach of the authors is to conduct a survey considering the developments and open research issues in spectrum management in CRN. The architecture and the cognitive functions defined in the IEEE 802.22 standard, *i.e.* Spectrum Sensing, Spectrum Sharing, Spectrum Decision, and Spectrum Mobility, are explained in detail. The main contribution of the article, however, consists in the proposal of an architecture for mobility management, which allows intercell resource allocation in a CRN. Specifically, the work is focused on how to deal with handoffs to implement spectrum mobility.

Gardellin *et al.* (GARDELLIN; DAS; LENZINI, 2013) dealt with coexistence of different CR cellular networks composed of a TV transmitter and diverse CR cells which provided access to microphones as primary user and IEEE 802.22 BSs and Customer Premises Equipments (CPE) as secondary users. The coexistence problem is stated in terms of channel assignment between the cells considering both a cooperative and a non-cooperative scheme for coexistence. The outcomes of the simulations considering each scheme are compared using a proposed fairness index, which is based on the throughput of each CPE. The channels are assigned based on their quality, measured considering the Signal to Interference Plus Noise Ratio (SINR). Using the mentioned parameter, the authors attempted to find an appropriate set of channels without harmful interference, making the coexistence possible. The results showed that the cooperative operation achieved a better fairness index when compared to non-cooperative and random methods.

2.1.3 Exclusive Spectrum Access

Exclusive spectrum access is typically implemented by commercial MNOs. With this model, a given MNO obtains exclusive access to a certain frequency range to provide both broadband services and cellular communications. The spectrum access is controlled by regulatory authorities which lease portions of the spectrum for exclusive usage, in concessions that last for long periods (*e.g.* 15 or 30 years). The concession of each frequency range is provided to an unique MNO for the duration of the leasing. Therefore, the exclusive spectrum access provides the licensee with an interference free area of the spectrum. Generally, the concession is based on a market-centric approach (*e.g.* auctions) to allocate spectrum frequencies to the highest bidder. This approach is leading to an exponential increase of prices based on exclusive spectrum usage rights (PONOMARENKO-TIMOFEEV et al., 2016).

The auction model demands high investments by the MNO to operate in a particular spectrum band. The operation of this band is guaranteed to be long-term and is exempt from harmful interference from other radio communication services or other MNOs. The MNO can make an internal decision about the network deployments within the rules of the licensing agreement, which generally allows operation in large geographical coverage areas. The internal decision may include the possibility of leasing underutilized resources for secondary resource renters.

In the event of resources leasing, the resource renter is demanded to operate in more than one frequency and even in more than one duplexing mode (MUSTONEN et al., 2014) (PALOLA et al., 2014). An example of this kind of implementation can be found in the literature by analyzing the scenario proposed by Palola *et al.* (PALOLA et al., 2014), which shows the deployment of four BSs simultaneous accessing LSA (TDD) and exclusive (FDD) spectrum bands.

2.1.4 Licensed Shared Access

LSA is a controlled sharing approach in which a limited amount of devices receive individual licenses to access network resources that are already assigned to one or more incumbent users (UMTS Forum, 2013). Incumbent users, in this context, are those network operators which own the network resources. LSA allows resources to be shared during a limited time period in a limited geographical area that is not currently being used by the incumbent. In contrast with CUS, the implementation of LSA considers sharing rules that guarantee a certain level of QoS to all the authorized devices. This guarantee is possible due to the celebration of a SLA to predefine the access conditions and the amount of resources that will be guaranteed to the LSA user.

Two models of SLA can be implemented to allow resources sharing: static or dynamic. The static model is based on a predefined SLA and considers conditions as specific exclusion zones, a predefined duration of operation, and specific spectrum frequencies that can be used, among others. On the other hand, the dynamic model typically takes advantage of cognitive functions, allowing spectrum sharing on a frequency, location, and time basis. In this model, SLA is celebrated dynamically, considering the restrictions applied by the incumbent user. The dynamic model is more complex since it demands a management system for providing, updating, and maintaining the access conditions (KHUN-JUSH et al., 2012).

Both static and dynamic SLA can be celebrated to allow resources sharing among different network operators. Different kinds of resources can be shared. Based on the definitions of Costa-Perez *et al.* (COSTA-PEREZ et al., 2013) and on a 3GPP report (3GPP, 2012), a classification of these kinds of resources is presented. Five classes of resources are defined: (I) core networks, (II) geographically split, (III) common network, (IV) common spectrum, and (V) radio access network.

- **Core Networks:** refers to multiple MNOs sharing a common network infrastructure. For operators that have multiple frequency allocations, it is possible to share infrastructural elements, however it is not possible to share the radio frequencies. In this case, the operators connect directly to their own dedicated carrier.
- **Geographically Split Networks:** is the situation where various licensed MNOs cover different geographical areas, *e.g.* parts of a country, but cooperate to provide joint coverage. Therefore, a larger geographical area will be covered.
- **Common Network:** in this case, an operator which covers a specific geographical area allows other operators to use this coverage for their subscribers. Outside this geographical area, coverage is provided by each of the operators independently.
- **Common Spectrum:** corresponds to common spectrum network sharing when one operator has a frequency license and shares the allocated spectrum with other operators or a number of operators decide to pool their allocated spectrum.
- **Radio Access Network:** in this case, multiple MNOs share a common core network. The operators define deployment details, therefore different parts of the network's radio access infrastructure can be shared.

Dixit *et al.* (DIXIT; PERIYALWAR; YANIKOMEROGLU, 2013) proposed a framework to implement a cooperative coexistence between primary (licensed) and secondary (unlicensed) users on LTE networks. The main goal of the research is to optimize spectrum utilization, bringing as an advantage to the primary user the possibility of earning profits by leasing spectrum white spaces. A pricing model is introduced as a way to allow temporary access for secondary users when the network resources are underutilized.

A survey involving Radio Access Network (RAN) and business models for network virtualization is presented by Costa-Perez *et al.* (COSTA-PEREZ et al., 2013). The authors also proposed physical infrastructure sharing among different wireless service providers. The proposed approach allows on-demand resources negotiation for providing specific services, like Voice over IP (VoIP), live streaming, and even the emerging machine-to-machine (M2M) communication services.

The efforts towards a novel approach for spectrum sharing in the United States are discussed by Sohul *et al.* (SOHUL et al., 2015). Although the article is a theoretical survey, the authors present and discuss important scenarios of spectrum sharing based on LSA regime. The approach proposed by the authors allows the coexistence of heterogeneous networks, but is not focused on assuring QoS for the users of the spectrum access system nor on the cost of spectrum sharing.

Chatzikokolakis *et al.* (CHATZIKOKOLAKIS et al., 2015) analyze the requisites and technical enablers of spectrum sharing in the context of heterogeneous networks and different frequency allocation regimes. Beyond the theoretical analysis, the authors also propose a simple spectrum sharing mechanism based on fuzzy logic. The proposed algorithm is used to match the needs of the spectrum renter by selecting the more suitable spectrum frequencies to serve its demands. A functional architecture is used to allow communication among primary and secondary users. The results obtained via simulation show that the proposed architecture and the artificial intelligence algorithm together provide an increase on the spectrum allocation efficiency.

Spectrum sharing among co-primary 5G small cell networks is investigated by Singh *et al.* (SINGH et al., 2015). A non-cooperative protocol is proposed to keep the overhead low. Such a protocol is based on minimizing the cost involved for a given network operator to rent resources from another 5G network operator. The model used to describe the costs is based on spectrum favors. Two approaches are proposed to meet various network operation scenarios. An instantaneous reciprocity model is applied in situations where the operators are considered impatient. On the other hand, a long-term reciprocity is proposed to be used when operators have persistent and publicly known identity, so the operators can learn from each other behavior. In both approaches, the cost of spectrum favors is calculated based on a repeated non-cooperative game.

2.2 Background on Heterogeneous Wireless Networks

Many technological challenges are faced by the MNO which requests resources during a resources sharing initiative. Some of the most challenging are the different technologies, the variable topologies implemented by the resources provider, and the different kinds of resources *i.e.* spectrum of frequencies and channel capacity. In order to deal with these challenges, the resources sharing architecture must be aware of the technologies involved in the resources sharing initiative. Especially, an inter-network signaling protocol must be implemented to permit the communication and traffic steering among heterogeneous networks.

According to an ETSI Report (ETSI, 2013), the resources sharing regime in heterogeneous wireless networks considers that the MNO which is providing resources determines the frequency ranges or the channel capacity available for sharing. Then, the physical implementation of the resources sharing follows one of the following models. The first one is called macro cells and relies on the implementation of high power BSs. In this case, a high area of coverage is available, what leads to a reduced probability of interference. Generally, different spectrum frequencies are deployed to allow heterogeneous networks to coexist. One example of the macro cell resources sharing is the 3GPP approach which allows resources sharing among commercial service providers.

The second approach is the deployment of small cells. In this situation, micro cells, pico cells, and femto cells can be considered. Resources sharing is implemented through the deployment of low power BSs, what limits the geographical area for resources sharing and consequently increases the probability of interference. The low transmission power of small cells, as well as their typical small coverage area brings the advantage of allowing a small cell deployment to cover a limited geographic area. This characteristic creates the opportunity for resources sharing in areas where multiple technologies coexist in the same spectrum frequency.

It is important to emphasize that the mentioned models are not mutually exclusive and a MNO may deploy more than one model in different parts of the resources sharing region. The definition of the target cell takes into account the number of users in the deployment area, the network load, and the QoS requirements, among other factors. In this thesis, three case studies are considered and all involve the coexistence of heterogeneous MNOs in both small and macro cells.

2.2.1 Inter-Network Communication and Signaling

Inter-network resources sharing in the approach proposed in this thesis only happens when the resources renter is unable to find enough resources to access the Internet on its own network. In this case, an inter-network communication protocol is started, as detailed in Figure 2.2. The specification of such a protocol is based on the approach of the LTE standard published by the 3GPP (3GPP...,). The choice of keeping the compatibility with the LTE-Advanced protocol is justified because LTE is considered the most important network in the proposed architecture, since they play the role of primary users and typically serve a higher number of users in a comparison with other technologies. Although the proposed signaling protocol is compatible with LTE, the illustration is an interpretation of the protocol adapted to the scenarios to be

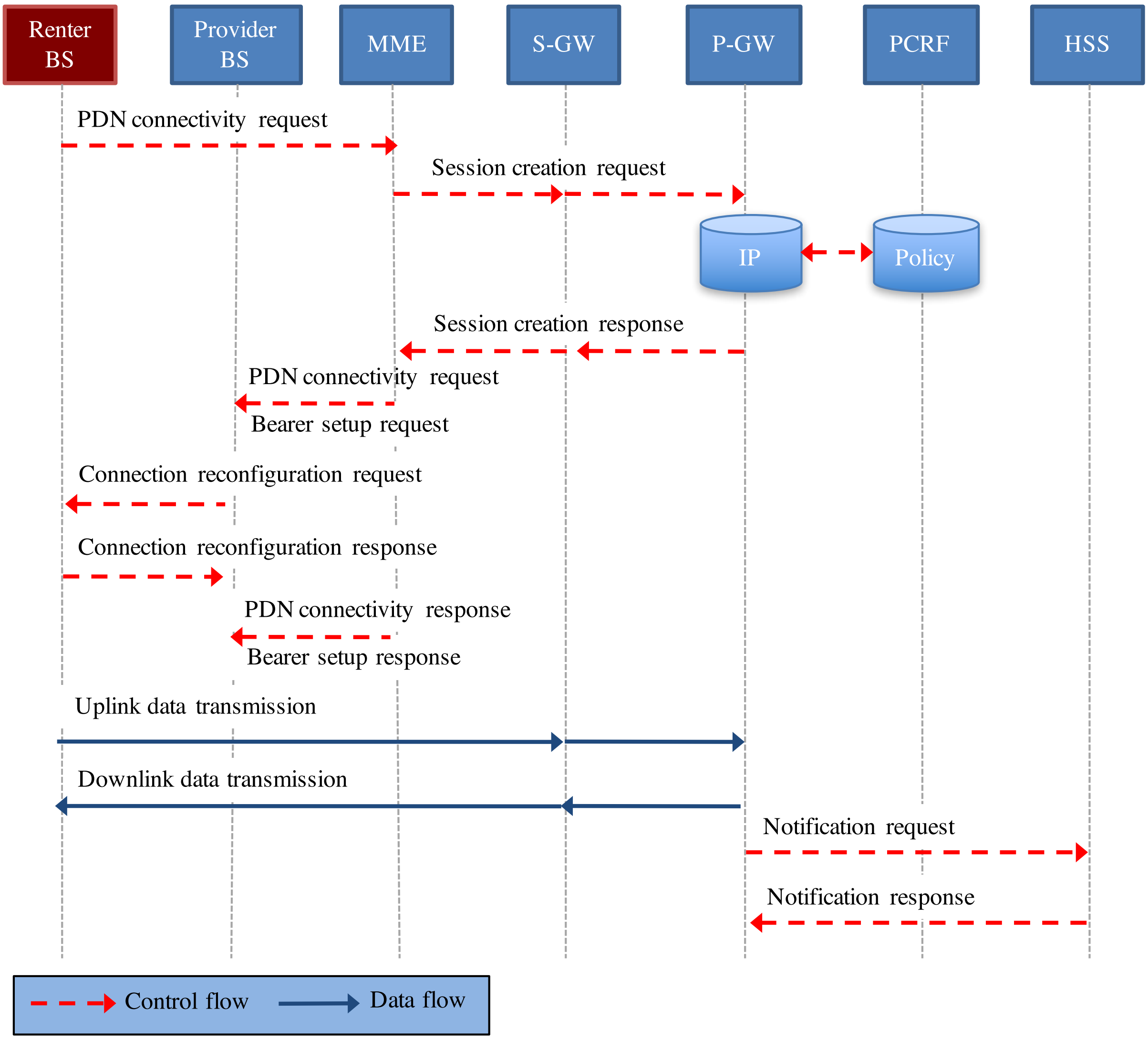


Figure 2.2: Inter-network Signaling

covered in this thesis without demanding modifications on the original protocol.

Two paths are represented in Figure 2.2 (I) control information (dashed line) and (II) data flow (continuous line). The communication flow is always initiated by the network operator which wants to rent resources. The technology used by the renter may be LTE or another wireless technology, such as IEEE 802.11, IEEE 802.16, or IEEE 802.22, among others. The only modification required to the renter, if it is not a LTE device, is the feature of formatting inter-networking messages according to the presented protocol. In the case of a LTE renter, no modification is required. The renter starts the process by communicating with a BS to dynamically establish a SLA to access shared resources made available by the LTE MNO.

As shown in the figure, the renter begins this communication by sending a Packet Data Network (PDN) connectivity request to the resources provider. This message is automatically forwarded to the Mobility Management Entity (MME), which controls all signaling between the devices within the resources provider and its core network. In order to provide such control, the

MME receives information from a Home Subscriber Service (HSS), which holds information about authorized network users, as QoS profiles, roaming restrictions, and PDNs which can be accessed by a given device. In the sequence, the MME demands to the Serving Gateway (S-GW) the creation of a transmission session.

This session is only created after the Policy Control and Charging Rules Function (PCRF) verifies network, SLA, and QoS policies, and the PDN Gateway (P-GW) provides IP connectivity. After receiving a successful response, the MME sets the bearer up, allowing the renter to access the shared network resources. Finally, the MME informs the HSS about the new communication. The data transmissions pass through a S-GW, which is responsible for controlling the mobility of devices between different BSs as well as for administrative tasks such as collecting information for charging purposes, SLA compliance verification, and lawful inspections.

2.2.2 Cognitive Wireless Networks

A possible approach to deal with resources sharing is the implementation of Dynamic Spectrum Access (DSA) techniques using the concepts of CRN, which is an enabling technology to DSA by allowing unlicensed network users to opportunistically access licensed spectrum frequencies, through resources sharing techniques. However, resources sharing is not the only challenge of CRNs, since characteristics as the variability on the spectrum frequencies usage, and guaranteeing QoS to the network applications also must be taken into account. According to Akyldiz *et. al.* (AKYILDIZ et al., 2006), CR users should implement four techniques to deal with these challenges:

1. Determine which portions of the spectrum are available;
2. Select the best available channel;
3. Coordinate access to this channel with other users;
4. Vacate the channel when a licensed user is detected.

These four techniques are important to enable the implementation of CRNs. However, cognitive functions are demanded to allow the correct operation of such techniques by turning available updated information about the CRN status.

To be part of a CRN, a device must be able to perform four basic functions. These functions, which are defined to work in Physical (PHY) and Medium Access Control (MAC) layers are called Spectrum Sensing, Spectrum Sharing, Spectrum Mobility, and Spectrum Decision. The aim of such functions is to efficiently manage the access of CR devices to the spectrum of frequencies. Towards this aim, control messages are exchanged to optimize the sharing of spectrum resources by the CRN users. Figure 2.3 shows the relationship among the aforementioned functions, which are detailed in the next sections.

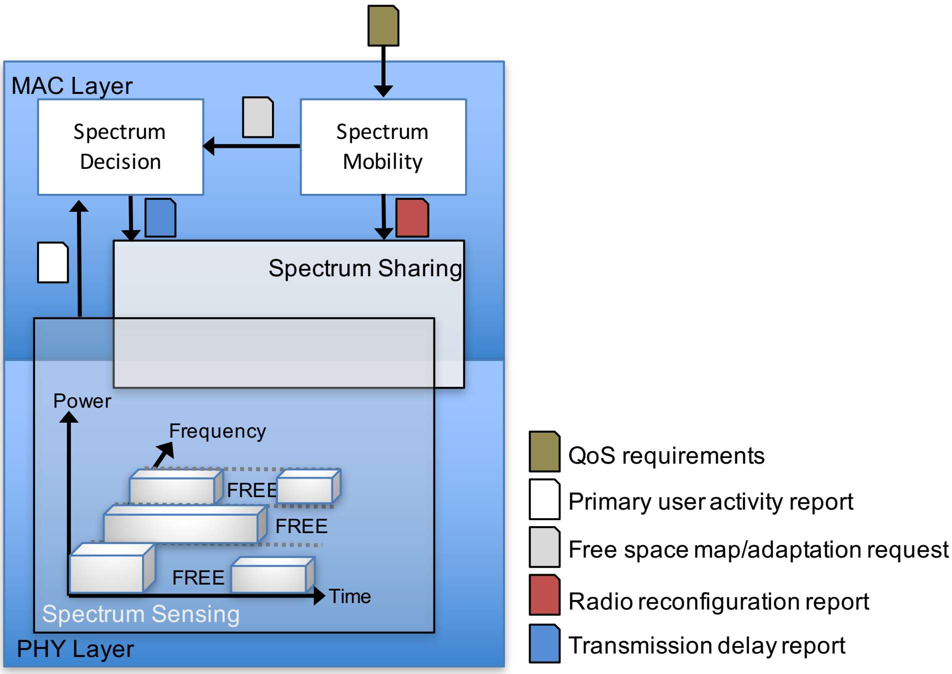


Figure 2.3: Cognitive Functions

2.2.2.1 *Spectrum Sensing*

Spectrum sensing is the most important function to allow the correct implementation of CRNs. This function analyzes the spectrum of frequencies to determine its usage characteristics. According to Yucek *et.al* (YUCEK; ARSLAN, 2009), these characteristics are typically classified into five different dimensions, called frequency, time, space, code, and angle, as illustrated in Figure 2.4.

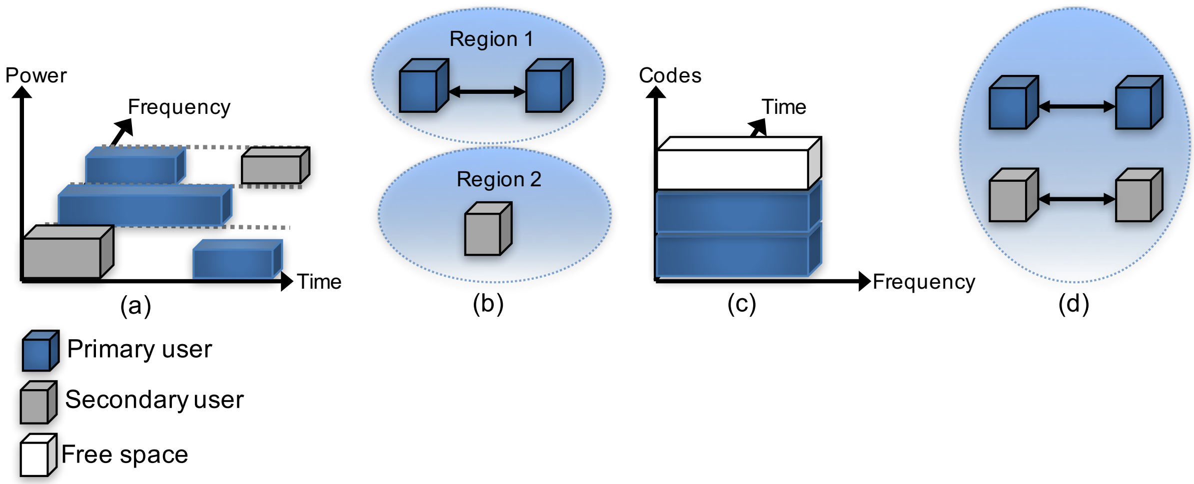


Figure 2.4: Spectrum Sensing Dimensions

Time and frequency dimensions are considered together in spectrum sensing, as can be seen in Figure 2.4 (a). In this case, parts of the spectrum of frequencies are divided into narrower chunks of frequency bands. Spectrum opportunity in this dimension is based on the principle that all the bands are not used simultaneously, *i.e.*, some bands might be available for opportunistic usage. This principle involves also the availability of a specific part of the spectrum in time. Thus, there will be times when bands of frequencies will be available for opportunistic usage.

Space dimension (Figure 2.4 (b)) bases the spectrum sensing in the Location (latitude, longitude, and elevation) and distance of primary users. This dimension considers that the spectrum of frequencies can be available in some parts of a geographical area, while occupied in other parts of the same area at a given instant of time. This approach takes advantage of the propagation loss (path loss) in space. Therefore, the recognition of a primary user transmission can be conducted by analyzing the interference level. No interference typically means that no primary user is transmitting in the analyzed area. However, the probability of false negatives, *i.e.*, the probability of not detecting the primary user, exists due to the so called hidden terminal problem.

The spectrum over a wideband might be used at a given time through spread spectrum or frequency hopping, what does not mean that there is no availability over this band. Simultaneous transmission without interfering with primary users would be possible in code domain using an orthogonal code with respect to codes that primary users are using, as illustrated in Figure 2.4 (c). This requires the opportunity in code domain, *i.e.* not only detecting the usage of the spectrum of frequencies, but also determining the used codes. To enable code dimension spectrum sensing, the spreading code, and the timing information are needed to allow secondary users to synchronize their transmissions with the primary user.

The last spectrum sensing dimension defined in the literature is angle, which is shown in Figure 2.4 (d). In this case, knowledge about transmission direction of primary users' beam, *i.e.* azimuth and elevation angle, as well as the location of the primary user are considered. This approach permits that even if a primary user is transmitting in a specific direction, the secondary user can transmit in the same spectrum band. However, the secondary user must transmit in other direction to avoid creating interference on the primary user's transmission. Considering the aforementioned dimensions, spectrum sensing can be divided into five techniques that represent different approaches to detect the transmission of primary users. These techniques are called Energy-based detection, Waveform-based detection, Cyclostationarity-Based Sensing, Radio Identification Based Sensing, and Matched-Filtering (YUCEK; ARSLAN, 2009).

2.2.2.2 Spectrum Sharing

The concept of spectrum sharing in CRNs is introduced to enhance the efficiency of spectrum usage by allowing primary users to share spectrum with secondary ones (MITOLA; MAGUIRE, 1999). The shared nature of the radio frequency channel, however demands the coordination of transmission attempts between primary and secondary users. Indeed, spectrum sharing should include much of the functionality of a MAC protocol (AKYILDIZ et al., 2008). Spectrum sharing is classified by the literature according to four aspects: architecture, spectrum allocation behavior, spectrum access technique, and scope.

The architecture of a spectrum sharing protocol is typically classified as centralized or decentralized. In the centralized approach (HAKIM et al., 2010) (SRINIVASA; JAFAR, 2008), the spectrum allocation and access procedures are controlled by a central entity. In this case, the central entity can be responsible for both sensing and creating a spectrum allocation map. However, other possible approach is the application of a distributed sensing procedure, where secondary users perform measurements that are forwarded to the central entity, that constructs the spectrum allocation map. These procedures are presented in Figure 2.5.

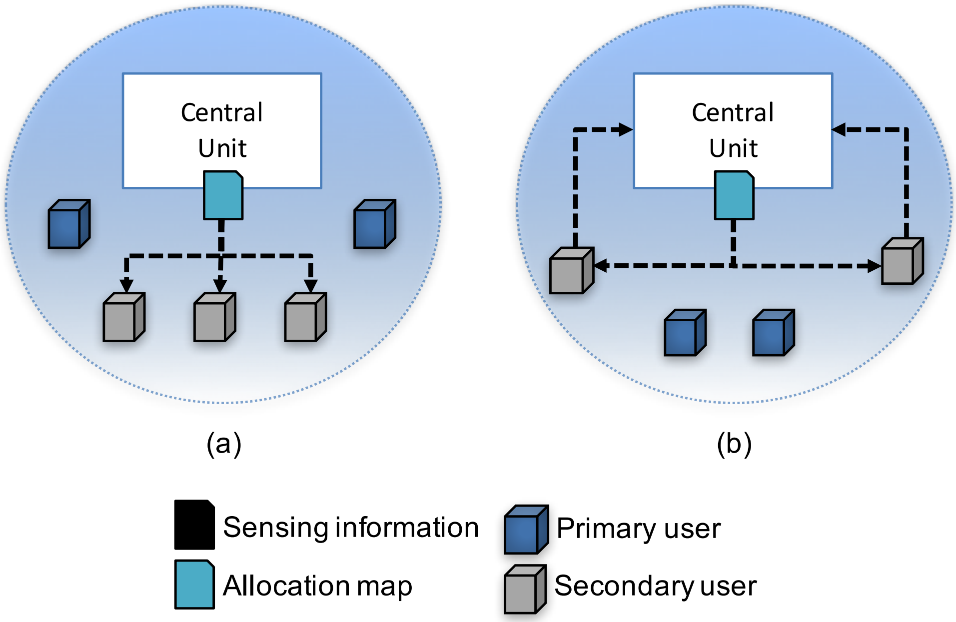


Figure 2.5: Centralized Spectrum Sharing

Figure 2.5 (a) illustrates the organization of a truly centralized architecture. The central unity is responsible for both sensing the spectrum and for the allocation map generation. This map is broadcasted to the secondary users. Figure 2.5 (b), in turn, shows the second approach, where secondary users perform spectrum sensing and inform the measurement results to the central entity that processes the received information to assemble the spectrum allocation map. In the sequence, this map is transmitted to the secondary users which will be able to access the spectrum of frequencies in a limited geographical region for a specific amount of time.

Distributed spectrum sharing relies on a set of policies that must be followed by the network users (KESHAVARZ et al., 2010). These policies are typically defined in a local network scope. However, the policies may also be considered for the global scope of a given MNO. The case of a local scope is exemplified in Figure 2.6 (a), in which the network users access the policies database in order to decide whether to access the spectrum of frequencies in a given instant of time or not.

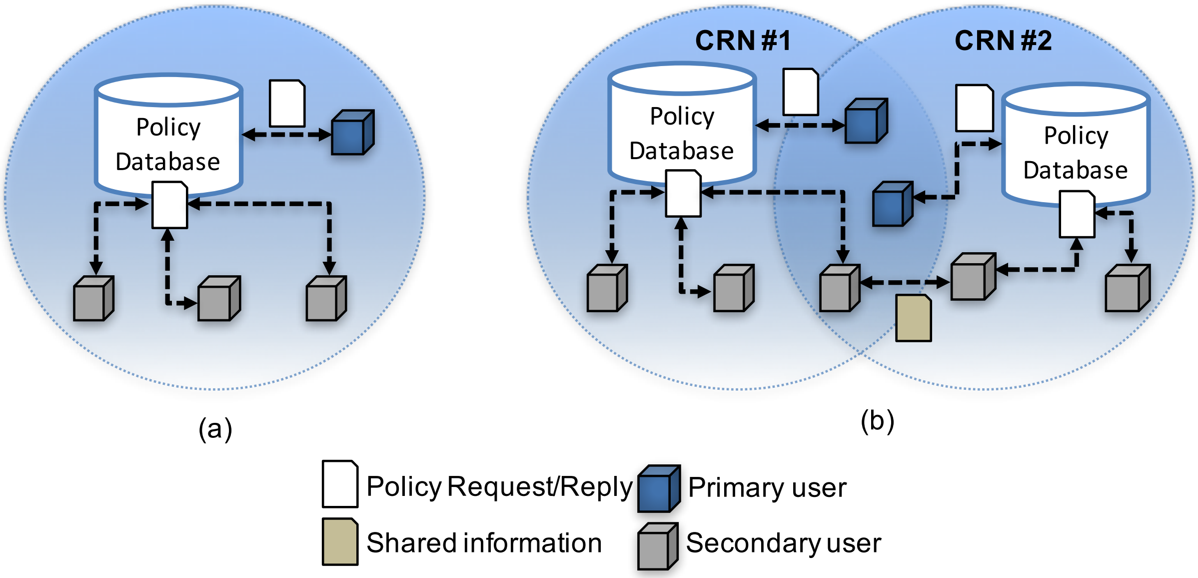


Figure 2.6: Distributed Spectrum Sharing

The second possible approach for implementing distributed spectrum sharing consists on the collaboration between at least two different CRNs, as can be seen in Figure 2.6 (b). In this case, network users first perform the already discussed spectrum sharing procedures. In the sequence, a given secondary user is able to share its measurement results with other secondary user which is part of another CRN. An example of this second approach is shown in the figure, where a secondary user belonging to CRN #1 shares information with a secondary user of CRN #2.

The behavior of the network users is also considered to provide spectrum sharing. Network users are classified into two groups: (I) non-cooperative or non-collaborative users, or (II) cooperative or collaborative users. Users belonging to the first group present a selfish behavior, while users in the second group share information to improve their spectrum sharing measurements and decisions (AKYILDIZ et al., 2008).

Non-cooperative spectrum sharing is a simple method where a network user solely takes the decision about accessing the spectrum of frequency or not. An advantage of this method is that it does not require frequent message exchange between neighbors. Moreover, due to the simplicity, this is a cheap, and easy to implement method. However, there are shortcoming regarding mainly the performance of the method, because interference with other network users is not considered. Furthermore, non-cooperative solutions may result in spectrum underutilization. An example of non-cooperative spectrum sharing can be seen in Figure 2.6 (a).

Cooperative spectrum sharing, in turn, exploits the interference measurements of each user such that the effect of the communication of one user on others is considered. A common technique used in these schemes is forming clusters to share interference information locally. This localized operation provides an effective balance between a fully centralized and a distributed scheme. Figure 2.6 (b) shows an example of this cooperative method, where CR users are able to share information with their neighbors. Cooperative approaches generally outperform non-cooperative ones. Moreover, cooperative techniques result in a certain degree of fairness, as well as improved throughput. On the other hand, the performance degradation of non-cooperative approaches is generally offset by the significantly low information exchange and hence, energy consumption (AKYILDIZ et al., 2008).

The third classification for spectrum sharing in CRNs is based on the access technology. The first approach to allow this kind of spectrum sharing is called overlay spectrum sharing. In this case, network users are allowed to use a portion of the spectrum which has not been used by licensed users. This minimizes interference with primary users. The second approach found in the literature is underlay spectrum sharing. The spread spectrum techniques are exploited in this approach, such that the transmission of a secondary user is regarded as noise by licensed users. Underlay techniques can utilize higher bandwidth at the cost of a slight increase in complexity. Considering this trade-off, hybrid techniques can be considered for the spectrum access technology in CRNs (SRINIVASA; JAFAR, 2008).

Finally, spectrum sharing can be implemented inside a CRN (*i.e.* intra-network spectrum sharing), or among multiple coexisting CRNs (*i.e.* inter-network spectrum sharing). Intra-network spectrum sharing focus on spectrum allocation between the entities of a CRN, as shown in Figure 2.6 (a). In this case, secondary users must access the available frequencies of the spectrum without causing interference to the primary users. On the other hand, inter-network spectrum sharing poses unique challenges that have not been considered previously in wireless communication systems (AKYILDIZ et al., 2008). The CR architecture enables multiple systems to be deployed in overlapping locations and spectrum, as in the example presented in Figure 2.6 (b).

2.2.2.3 Spectrum Decision

Spectrum decision is a very important function to allow the implementation of CRNs, because it provides to secondary users the capability to decide which is the best spectrum band among the available ones. This decision must be taken considering the QoS requirements of the applications and consequently will impact in the overall network performance. Besides the QoS requirements of the application, spectrum decision must take into account other aspects, such as the activities of other users in the CRN and the channel conditions. The implementation of this function usually consists of two steps regarding the activities of the network users: first, each spectrum frequency is characterized, based on not only local observations of secondary users but also on statistical information of primary networks. In the sequence, based on the results obtained from such a characterization, the most appropriate spectrum band can be chosen.

Another important aspect that affects the decision function is the time varying radio frequency channel propagation conditions. In this sense, four main characteristics must be considered:

- **Interference:** the power of a secondary user can be derived from the amount of interference it generates to the primary user. This power measurement can then be used for the assessment of channel capacity.
- **Path Loss:** this characteristic is closely related to distance and frequency. As the operating frequency increases, the path loss increases, which results in a decrease in the transmission range. If transmission power is increased to compensate for the increased path loss, interference at other users may increase.
- **Errors:** depending on the modulation scheme and the interference level of the spectrum band, the error rate of the channel changes. This characteristics may result in information loss, what impacts on the network performance and consequently on the QoS level of the applications.
- **Transmission Delay:** to address varying levels of interference, path loss, and errors, different types of link layer protocols are required what results in different link layer delays. It is desirable to identify the spectrum bands that combine all the characterization parameters previously described for accurate spectrum decision.

These four characteristics must be considered through probabilistic models. A typical approach found in the literature to address this question is the definition of an error generation model that takes into account the variability of the radio frequency channel conditions (KUNST et al., 2011). After the decision process, the spectrum mobility function must reconfigure the network devices to reflect the decision taken.

2.2.2.4 Spectrum Mobility

The spectrum mobility function has the goal of changing the CR devices configurations to access the spectrum frequency defined by the spectrum decision function. Spectrum mobility rises a new type of handoff in CRNs, which is called spectrum handoff (AKYILDIZ et al., 2008). Protocols for both PHY and MAC layers of the network stack must adapt to the channel parameters of the operating frequency. This adaptation is necessary because the radio frequency channel conditions typically chance when the transmission frequency is altered (KUNST et al., 2011).

Spectrum handoff should be transparent to avoid affecting the QoS provided to secondary users' applications. Another goal of the spectrum handoff is to implement fast transitions between two frequencies to avoid increasing the network latency. To reduce the performance degradation, it is fundamental for a secondary user to have information about the duration of a spectrum handoff. This information can be provided by the sensing function. After the latency information is available, the secondary user can assess whether the QoS of ongoing communications can be preserved or not.

The implementation of spectrum handoff demands changes on the operating frequency of the secondary user's device. Such a change must be dynamic to permit moving among frequencies when a primary user starts a transmission.

2.2.3 Traffic Load Forecasting

One of the key features demanded for the accurate resources sharing is correctly assessing the future traffic demands of the MNOs involved in the resources sharing initiative. The purpose of traffic load forecasting models is to discover the future trend of the traffic to allow the estimation of the channel capacity resource occupation. In this thesis, to carry out the traffic load forecasting, machine learning algorithms are considered. Initially, four algorithms are considered and compared in order to decide which one is more suitable to estimate the future traffic load. Details on these algorithms are presented in the upcoming subsections. Details on the comparison among these models are available in Appendix B.

2.2.3.1 Multiple Linear Regression Model

MLRM is a statistic model largely implemented for predictive analysis. This model is used to explain the relationship between one continuous dependent variable and two or more independent variables. The MRLM is broadly applied in areas such as trend line, telecommunications, finance, economy, environmental science, and epidemiology, among others. MLRM has many practical uses, among the most important applications, it can be used for forecasting through fitting the predictive model to an observed data set. The purpose of the MLRM is to establish a relationship among the group of predictors (*e.g.* historical mean traffic per second, minute or hour). MLRM allows to understand which predictors have the greatest impact, and it aims to calculate the best fitting curve by minimizing the least squares errors.

Many researches have applied this model in the context of data communications. Papadoupli, Raftopoulus, and Shen (PAPADOPOULI; RAFTOPOULOS; SHEN, 2006) propose a short term traffic load forecasting in wireless networks. The authors evaluate several traffic forecasting algorithms which consider the recent traffic history and information related to the current traffic flow. According to Liu and Lee (LIU; LEE, 2015), the MLR and other six algorithms are able to carry out the throughput prediction in mobile data networks. Further, they develop an information theoretic lower bound to define the prediction error. Niami *et al.* (NIAMI et al., 2014) apply the prediction over metrics such as the number of retransmissions needed and time expected to transmit a data packet to adjust the routing metrics in ad-hoc wireless networks. In fact, the proposed solution anticipates the signal strength using linear regression over the historical measurements of the link quality. Noulas *et al.* (NOULAS et al., 2012) leverage the historical information on coarse granularity from the Foursquare social platform to predict the user mobility. The user, global, and temporal features sets are analyzed, then such features are trained in the supervised classification problem to predict the next check-in state.

The presented examples show that MLRM can be applied to perform the prediction of scenario variables based on historical information. In this thesis, the MLRM is used to perform the traffic load forecasting in different scenarios, involving heterogeneous networks. The output of this model is used by the proposed resources sharing architecture, especially by the provisioning algorithm to take decisions on which is the best resources provider in a given scenario.

2.2.3.2 Neural Network Model

Neural Networks (NNs) are extensively used in computer science. This model is based on a wide collection of simple neural units. Each neural unit is connected with several others, and such connections can improve or impede the activation status of adjacent neural units. Each neural unit computes using a summation function, considering a threshold such that the signal must exceed the limit before propagating to other neurons. These systems are trained and self-learn, in situations where the solution or feature detection is hard to express in a traditional computer program.

NNs can estimate almost any function in an efficient and stable manner when the underlying data relationships are unknown (RODRIGUES; NOGUEIRA; SALVADOR, 2010). The NN model is a nonlinear, nonparametric, adaptive modeling approach which relies on the observed or historical data rather than on an analytical model (FENG; SHU, 2005). The architecture and parameters of the NN are determined by the dataset.

NNs typically consist of multiple layers, and the signal path traverses from front to back. Back propagation is the use of forwarding stimulation to reset weights on the front neural units. This is sometimes done in combination with training when the correct result is known. Modern NNs allow stimulation and inhibition with connections interacting in a much more chaotic and complex fashion. Dynamic NNs are the most advanced and can form new connections and neural units dynamically, based on predefined rules.

A NN is composed of nodes interconnected according to weights to create different layers of neurons. A NN comprises at least one input layer, one or more hidden layers, and an output layer. The most traditional NN architecture is called feed-forward because, in this kind of NNs, the information travels through the network only in the forward direction, *i.e.* from the input layer towards the output layer, as illustrated in Figure 2.7.

Using a NN as a predictor involves two phases: (I) the training phase and (II) the prediction phase (BARABAS et al., 2011). In the training phase, the training set is presented to the input layer, and the parameters of the NN are dynamically adjusted to achieve the desired output value for the input set. The most common learning algorithm is the back propagation algorithm,

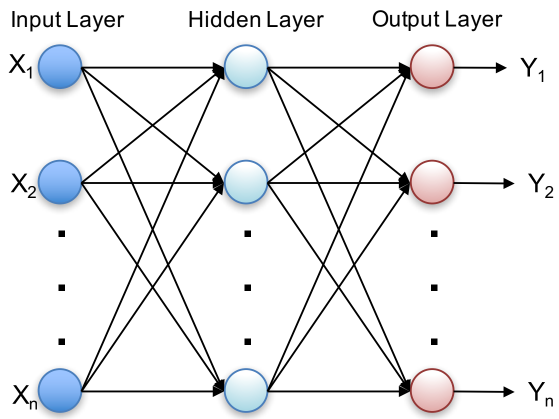


Figure 2.7: Neural Network

which is based on the backward propagation of the error. In this algorithm, the weights change continuously until the output error falls below a predetermined value. In this way, the NN can recognize patterns between input sets and the corresponding target values. The prediction phase represents the testing of the NN. A new input (not included in the training set) is then presented to the NN, and the output is calculated, thereby predicting the outcome of a new input data.

The number of hidden layers and the number of nodes in each layer is usually chosen empirically. NN must have at least one hidden layer to predict nonlinear values. Too many hidden layers slow down the training process and increase the complexity of the network. To improve the nonlinearity of the solution, the activation functions of neurons in the hidden layer are sigmoid functions, while the output nodes have linear transfer functions.

2.2.3.3 Fourier Model

The curve fitting using the Fourier series model can be applied in various fields, such as engineering, seismology and economics, among others (STRANG, 1994). Fourier applied this technique to find the solution of the heat equation. Furthermore, the Fourier analysis in forecasting overcomes certain limitations that other models have in capturing the seasonality phenomena (LYE; YUAN; CAI, 2009).

Lye *et al.* (LYE; YUAN; CAI, 2009) decomposes the given time series (*i.e.* historical data) into a linear combination of sinusoids (*i.e.* frequency components) via an orthogonal transform method, developing a new method for forecasting based on the analysis of the frequency domain. In this sense, the Fourier approach has been used for forecasting changes in electricity load, transportation, network traffic, and prices, which are variables that are clearly related to cyclic and recursive variations. The Fourier series can be applied in the context of network traffic forecasting by identifying traffic patterns based on previous measurements of the traffic.

2.2.3.4 Regression Tree Model

Regression Tree Model (RTM) uses a decision tree as a predictive model which observes an item to draw conclusions about its target value. This model is applied in the fields of machine learning, statistics, and data mining, among others. The RTM is often used to model the correlation of load consumption and other factors (*e.g.* number and priority of users) to apply the load forecasting regression method. In this approach, a binary tree is constructed based on a data set. The goal is to produce subsets of the data which are as homogeneous as possible in accordance with the response variables. According to Wilkinson (WILKINSON, 2004), the classification of the RTM can be built by growing and pruning the tree.

According to Strobl *et al.* (STROBL; MALLEY; TUTZ, 2009), a regression tree is simple nonparametric regression approach, which has as main characteristic the featured space, *i.e.* the space spanned by all predictor variables is recursively split into a set of rectangular areas. They advocate that the splitting is created such that observations with similar response values are grouped. Thus, after the splitting is completed, a constant value is predicted within each resulting area. Qiang *et al.* (XU et al., 2013) proposed the network performance forecast for real-time, iterative mobile applications. This solution is based on a machine learning framework which implements the concept of regression trees to identify the trend of the network performance over short, fine-grained time windows using previously available observations.

3 RELATED WORK

The related works were surveyed based on the concepts of SLR (KITCHENHAM et al., 2009). Such a methodology was previously used, for example, by the software engineering community to map the research on that area. In this thesis, an adapted methodology is used, which is divided in four main steps, namely, (I) definition of research questions, (II) search for suitable papers, (III) applicability assessment, and (IV) classification of the selected articles. These steps are presented and discussed in the following subsections.

3.1 Definition of Research Questions

The first step of the methodology consists in defining a set of research questions. The answers to these questions allow depicting the current status of research and applications of resources sharing in wireless networks. Three research questions have been defined as the base for the application of the methodology:

1. Which is the most relevant publication media related to resources sharing in wireless networks?
2. What is, in quantitative terms, the research activity related to resources sharing in wireless networks in the past three years?
3. Which are the 10 most relevant papers in the area of resources sharing in wireless networks?

These questions have been considered during the process of searching for suitable papers.

3.2 Search for Suitable Papers

The second step of the methodology is the selection of papers that contributed for the evolution of wireless network resources sharing research. The search process was focused on papers published in worldwide relevant journals, magazines, and conferences in the fields of data communications and networks. The selection of publication medias considered the h5-index of the publication, computed based on the Google Scholar Metrics. Papers published between 2014 and 2016 are considered. It is important to emphasize that the year 2017 is not considered since it is not possible to collect papers published during the whole year, what could lead to imprecise results. In order for the publication to be included in the search process, the minimum h5-index was set to 50.

Table 3.1 lists the publications from where the initial set of articles has been collected. Analyzing the table it is possible to answer the first research question of the SLR. Considering the h5-index, the most relevant publication media in the field of resources sharing in wireless networks is the IEEE Communications Magazine, with a value of 102.

Table 3.1: Set of Considered Publications

Publication	**h5-index**
IEEE Communications Magazine	102
IEEE Transactions on Wireless Communications	82
IEEE INFOCOM	80
IEEE Journal on Selected Areas in Communications	77
Communications of the ACM	73
IEEE Transactions on Communications	60
IEEE/ACM Transactions on Networking	57
Journal of Network and Computer Applications	56
IEEE Wireless Communications	56
Elsevier Computer networks	50
IEEE Communications Letters	50

Papers were retrieved from digital libraries (*e.g.*, IEEE Explorer, ACM Digital Library, and Elsevier). The definition of the initial set of articles was based on a search on the metadata of the publications, considering the following string: **("RESOURCES" OR "SHARING") AND ("WIRELESS" AND "NETWORK")**. Exceptions regarding the date of publication were conceded to three papers notably relevant in the area, which were published in 2013 and are closely related to the solution proposed in this thesis.

3.3 Applicability Assessment

In the third step of the methodology, the initial set of papers passes through a screening process aiming to apply an exclusion criteria. The goal of this step is to individually observe each paper to decide whether it should be kept or excluded from analysis. As a result of this step, a smaller set of papers, which is more closely related to resources sharing in wireless networks research is obtained.

The applicability assessment methodology consists in the inspection of the papers to attribute a grade to each one based on the three applicability assessment questions, defined as follows:

1. Is the paper related to wireless networks resources sharing?
2. Does the paper mention heterogeneous network technologies or multioperator scenarios?
3. Do the authors of the paper propose and evaluate a resources sharing architecture?

This phase allows the exclusion from the initial set of those papers which are not the main focus of the proposed work, *e.g.* guest editorials and other documents which are relevant, but not suitable for this SLR. First, the title, keywords, and abstract information are analyzed to answer each of the predefined research questions. If the title, keywords, and abstract are considered not sufficient to answer all the questions, the introduction, methodology, and conclusion sections are also read. Based on Kitchenham's methodology (KITCHENHAM et al., 2009), three possible answers are defined, each one having a correlated grade: Yes=1, Partial or not clear=0.5, and No=0.

After the articles are graded, a decision is taken on whether each study must be kept or excluded from the original set, based on exclusion criteria. More formally, such criteria are:

- Papers mainly organized as comments or personal opinions are excluded from the set, since they usually do not present a validation methodology;
- Papers which are graded below 2, since at least 2 of the questions received 'no' or 'partial' as answer, indicating a not very relevant publication for this SLR;

After applying these criteria onto a original set of 103 papers, a total of 31 had a grade of at least 2.0 and therefore were considered relevant by the SLR methodology. The complete list of analyzed papers and their grades is available in Appendix A.

3.4 Classification of Selected Papers

In this section, the resulting classification of the papers after applying the SLR methodology is presented. Two approaches are used to show and explain the collected data. First, a correlation between the publications and the amount of selected papers is shown. Then, the 10 most relevant papers according to the methodology are presented.

The analysis methodology considered the search for papers in the previously selected publication medias using the aforementioned search string. The amount of papers found in each publication is presented in Table 3.2. In this table, the Original Set field represents the total amount of papers returned by the search string and the Final Set field is the number of papers considered relevant in that given publication media. The analysis of this table allows to answer the second research question presented in section 3.3. A total of 103 papers were considered, of which, 100 have been published in the last three years.

In order to answer the final research question of this SLR, the 10 most relevant papers are presented in Table 3.3. The order considered for selection of these papers follows the criteria of the highest grade. In cases in which two or more papers are assigned the same grade, the second criteria used for ordering is the year of publication. In this situation, the newer papers are going to be presented first.

The most relevant papers are going to be analyzed in the next section, where they will be compared with the solution proposed in this thesis.

Table 3.2: Selected papers per publication

Publication	**Original Set**	**Final Set**
IEEE Communications Magazine	23	11
IEEE Transactions on Wireless Communications	26	8
IEEE INFOCOM	0	0
IEEE Journal on Selected Areas in Communications	7	1
Communications of the ACM	1	0
IEEE Transactions on Communications	6	1
IEEE/ACM Transactions on Networking	4	1
Journal of Network and Computer Applications	8	0
IEEE Wireless Communications	11	3
Elsevier Computer networks	8	2
IEEE Communications Letters	9	4
Total	**103**	**31**

Table 3.3: 10 most relevant papers

Paper Title	**Q1**	**Q2**	**Q3**	**Grade**
Toward spectrum sharing: opportunities and technical enablers	1.0	1.0	1.0	3.0
Coordination protocol for inter-operator spectrum sharing in co-primary5G small cell networks	1.0	1.0	1.0	3.0
Synergistic spectrum sharing in 5G HetNets: A harmonized SDN-enabled approach	1.0	1.0	1.0	3.0
Beyond Coexistence: Traffic Steering in LTE Networks with Unlicensed Bands	1.0	1.0	1.0	3.0
Spectrum access system for the citizen broadband radio service	1.0	1.0	1.0	3.0
Self-coexistence in cellular cognitive radio networks based on the IEEE802.22 standard	1.0	1.0	1.0	3.0
Radio access network virtualization for future mobile carrier networks	1.0	1.0	1.0	3.0
Secondary User Access in LTE Architecture Based on a Base-Station-Centric Framework With Dynamic Pricing	1.0	1.0	1.0	3.0
Wireless resource sharing for multiple operators: Generalization, fairness, and the value of prediction	1.0	1.0	0.5	2.5
Advanced spectrum sharing in 5G cognitive heterogeneous networks	1.0	1.0	0.5	2.5

3.5 Related Work Analysis

The selection of related works to be considered for further analysis in this thesis follows the SLR methodology. Especially, the 10 papers considered more relevant in the resources sharing in wireless networks field of research. The analysis of each paper is presented in Table 3.4, considering the following key points:

- Paper title, authors, and publication media;
- Proposal and contributions of the paper;
- Network technologies considered in the proposed solution;
- Resources sharing regime;

#1 (CHATZIKOKOLAKIS et al., 2015)	
Paper Title	Toward Spectrum Sharing: Opportunities and Technical Enablers
Authors	Konstantinos Chatzikokolakis, Panagiotis Spapis, Alexandros Kaloxylos, and Nancy Alonistioti
Publication media	IEEE Communications Magazine, July 2015
Proposal	A framework is proposed to enable the MNOs and other spectrum license holders to exchange information about spectrum availability. A spectrum sharing mechanism based on fuzzy logic is also presented to facilitate the selection the most suitable spectrum to cover the MNOs needs.
Network Technologies	Multiple technologies using LSA frequencies
Sharing Regime	Licensed Shared Access
#2 (SINGH et al., 2015)	
Paper Title	Coordination Protocol for Inter-Operator Spectrum Sharing in Co-Primary 5G Small Cell Networks
Authors	Bikramjit Singh, Sofonias Hailu, Konstantinos Koufos, Alexis A. Dowhuszko, Olav Tirkkonen, Riku Jantti, and Randall Berry
Publication media	IEEE Communications Magazine, July 2015
Proposal	A coordination protocol is proposed to allow spectrum sharing among LTE MNOs. The protocol is non-cooperative, but assumes an agreement to a set of negotiation rules. The signaling overhead is low, and knowledge of a competitor's channel state information is not considered.
Network Technologies	LTE
Sharing Regime	Exclusive spectrum access
#3 (AKHTAR; WANG; HANZO, 2016)	

Paper Title	Synergistic Spectrum Sharing in 5G HetNets: A Harmonized SDN-Enabled Approach
Authors	Auon Muhammad Akhtar, Xianbin Wang, and Lajos Hanzo
Publication media	IEEE Communications Magazine, January 2016
Proposal	A Software Defined Networks based synergistic spectrum sharing technique is proposed. The proposal relies on the availability of distributed reports regarding to spectrum usage of 5G heterogeneous networks.
Network Technologies	Multiple in the context of Software Defined Networks
Sharing Regime	Collective Use
#4 (ZHANG et al., 2016)	
Paper Title	Beyond Coexistence: Traffic Steering in LTE Networks with Unlicensed Bands
Authors	Ning Zhang, Shan Zhang, Shaohua Wu, Ju Ren, Jon W. Mark, and Xuemin Shen
Publication media	IEEE Wireless Communications, December 2016
Proposal	The authors consider existing architectures to study how to efficiently utilize heterogeneous network resources for various service provisioning in LTE networks with unlicensed bands. The main focus of the study relies on traffic steering to distribute traffic among heterogenous LTE cells, radio access technologies, and spectrum bands based on the desires of the network or users.
Network Technologies	LTE
Sharing Regime	Collective Use
#5 (SOHUL et al., 2015)	
Paper Title	Spectrum Access System for the Citizen Broadband Radio Service
Authors	Munawwar M. Sohul, Miao Yao, Taeyoung Yang, and Jeffrey H. Reed
Publication media	IEEE Communications Magazine, July 2015
Proposal	The efforts toward a spectrum sharing system in the U.S. are presented by summarizing different interest groups' standpoint on the FCC proposed framework. A spectrum access system architecture is also proposed to accommodate the tiered access to shared spectrum.
Network Technologies	Multiple
Sharing Regime	Exclusive spectrum access and licensed shared access
#6 (GARDELLIN; DAS; LENZINI, 2013)	

Paper Title	Self-Coexistence in Cellular Cognitive Radio Networks Based on IEEE 802.22 Standard
Authors	Vanessa Gardellin, Sajar K. Das, and Luciano Lenzini
Publication media	IEEE Wireless Communications, April 2013
Proposal	Considering the IEEE 802.22 as the standard reference for cellular network mechanisms, the authors addresses coexistence issues and propose two channel assignment schemes for cooperative and non-cooperative CR devices.
Network Technologies	IEEE 802.22
Sharing Regime	Exclusive spectrum access
#7 (COSTA-PEREZ et al., 2013)	
Paper Title	Radio Access Network Virtualization for Future Mobile Carrier Networks
Authors	Xavier Costa-Peres and Joerg Swetina, Tao Guo, Rajesh Mahindra, and Sampath Rangarajan
Publication media	IEEE Communications Magazine, July 2013
Proposal	A solution based on spectrum sharing is presented. This solution is called network virtualization substrate, which can be natively implemented in BSs. The performance of the proposed solution is evaluated in a LTE network by means of simulation, showing that it can meet the needs of future virtualized mobile carrier networks in terms of isolation, utilization, and customization.
Network Technologies	Multiple
Sharing Regime	Collective use and licensed shared access
#8 (DIXIT; PERIYALWAR; YANIKOMEROGLU, 2013)	
Paper Title	Secondary User Access in LTE Architecture Based on a Base-Station-Centric Framework With Dynamic Pricing
Authors	Soumitra Dixit, Shalini Periyalwar, and Halim Yanikomeroglu
Publication media	IEEE Transactions on Vehicular Technology, January 2013
Proposal	A dynamic incentive-based pricing model is proposed to allow temporary wireless for secondary users during periods of low primary user demand, thus improving spectrum usage in the temporal domain. The implementation of the proposed framework to LTE infrastructure requires minimal enhancements and can be potentially attractive to wireless service providers.
Network Technologies	LTE and IEEE 802.16
Sharing Regime	Licensed shared access

#9 (MALANCHINI; VALENTIN; AYDIN, 2016)	
Paper Title	Wireless Resource Sharing for Multiple Operators: Generalization, Fairness, and the Value of Prediction
Authors	Ilaria Malanchini, Stefan Valentin, and Osman Aydin
Publication media	Computer Networks, February 2016
Proposal	A theoretical framework for multi-operator Scheduling is formulated. This formulation allows to analyze sharing guarantees and spectral efficiency for a large number of parameters and covers various fixed and dynamic resource sharing policies as special cases.
Network Technologies	Multiple
Sharing Regime	Exclusive spectrum access
#10 (YANG et al., 2016)	
Paper Title	Advanced Spectrum Sharing in 5G Cognitive Heterogeneous Networks
Authors	Chungang Yang, Jiandong Li, Mohsen Guizani, Alagan Anpalagan, and Maged Elkashlan
Publication media	IEEE Wireless Communications, April 2016
Proposal	A spectrum flowing scheme is proposed for 5G cognitive heterogeneous cellular networks, which improves both spectral and energy efficiency.
Network Technologies	Multiple device to device technologies
Sharing Regime	collective use

Table 3.4: Related Works Analysis

3.6 Related Work Summary

In this section, the related works are summarized considering their main features. These features are then compared to the solution presented in this thesis. Different approaches have been considered to implement resources sharing both in CUS and LSA regimes. Most of these proposals involve the implementation of complex algorithms aiming to mitigate interference or to find spectrum access opportunities using artificial intelligence techniques. Although very relevant, these solutions generally identify the need for additional resources and search for available resources without taking into account the QoS requirements of the resources renter. This common approach of related works may lead to the allocation of network resources that are not suitable for the users.

Table 3.5 summarizes the relevant previous works considering whether or not five different aspects are covered by each proposal. The first aspects refer to the type of network resources that can be shared, *i.e.* the support for sharing CUS and/or LSA resources. The support for heterogeneous networks is also considered to compare the proposals. Finally the cost and the support to QoS are also analyzed.

Table 3.5: Comparison between this proposed approach and related work

Related Work	**CUS**	**LSA**	**Heterogeneous**	**Cost**	**QoS**
#1 - Chatzikokolakis *et al.*	X	X	X		X
#2 - Singh *et al.*	X	X	X	X	
#3 - Akhtar *et al.*	X		X		
#4 - Zhang *et al.*	X		X		
#5 - Sohul *et al.*		X	X		
#6 - Gardellin *et al.*	X				X
#7 - Costa-Perez *et al.*	X	X	X		
#8 - Dixit *et al.*		X		X	
#9 - Malanchini *et al.*			X		
#10 - Yang *et al.*	X		X		
Proposed Architecture	X	X	X	X	X

Taking into account the limitations of the currently proposed solutions, in this thesis, an architecture is proposed to support the sharing of CUS and LSA resources in heterogeneous network environments considering both the QoS requirements and the cost of renting the resources. Another aim of the proposed solution is to deal with the trade-off between meeting the QoS requirements and reducing the cost of the resources rental. Details on the proposed solution are provided in the next chapter.

4 AN ARCHITECTURE FOR RESOURCES SHARING IN WIRELESS CELLULAR NETWORKS

In this chapter, the concepts of the proposed architecture are discussed. The designed solution is presented in section 4.1. Details on the three levels that compose the resources broker are provided. In section 4.2, the simulation tool used to obtain results to evaluate the behavior of the proposed solution is presented and discussed.

4.1 Design of the Resources Sharing Architecture

The design of the proposed architecture is presented in Figure 4.1. The illustration is divided in two parts which communicate through a polling and reply mechanism. The the left side of the image represents spectrum users which coexist in a geographical area considering a scenario that allows one network to communicate with all neighboring network operators. In the right side of the figure, the structure of a broker is represented. This broker is responsible for coordinating resources sharing among the spectrum users.

In the illustration, an example scenario is presented to reflect the behavior of typical spectrum users. In this scenario, four LTE-Advanced network operators are represented to illustrate the reality of LTE-Advanced frequencies allocation in Brazil (Anatel, 2015). An IEEE 802.22 CRN and an IEEE 802.11 (Wi-Fi) MNO were also represented to indicate that the proposed approach allows the coexistence and resources sharing between primary users and secondary users belonging to both homogeneous and heterogeneous network technologies with the restriction that they are located in the same geographical area.

Another important aspect to emphasize is related to the direction of the resources sharing. The proposed architecture allows resources sharing in two ways, *i.e.* each MNO can dynamically assume the role of a resources provider or the role of a resources renter. In Figure 4.1, the direction of resources provider is represented by a straight connector, while the possibility of resources renting is represented by a dashed connector. Although, for the sake of simplicity in the representation each secondary user is communicating with only one LTE, the architecture, indeed, allows the secondary users to communicate with any other MNO which implements the signaling protocol within the same geographical area. Details on this signaling protocol have been previously described in chapter 2.

Different types of resources can be shared in the proposed approach. Further than allowing the cooperation between primary and secondary users, the architecture permits the cooperation between wireless network technologies which operate using diverse kinds of resources, *e.g.* spectrum of frequencies and channel capacity. In Figure 4.1 LTE is an example of technology based on channel capacity, while IEEE 802.22 conducts spectrum sensing to directly transmit over the spectrum of frequencies. In order to turn feasible the translation between two kind of resources, a centralized entity is necessary. Moreover, a centralized approach is indicated in

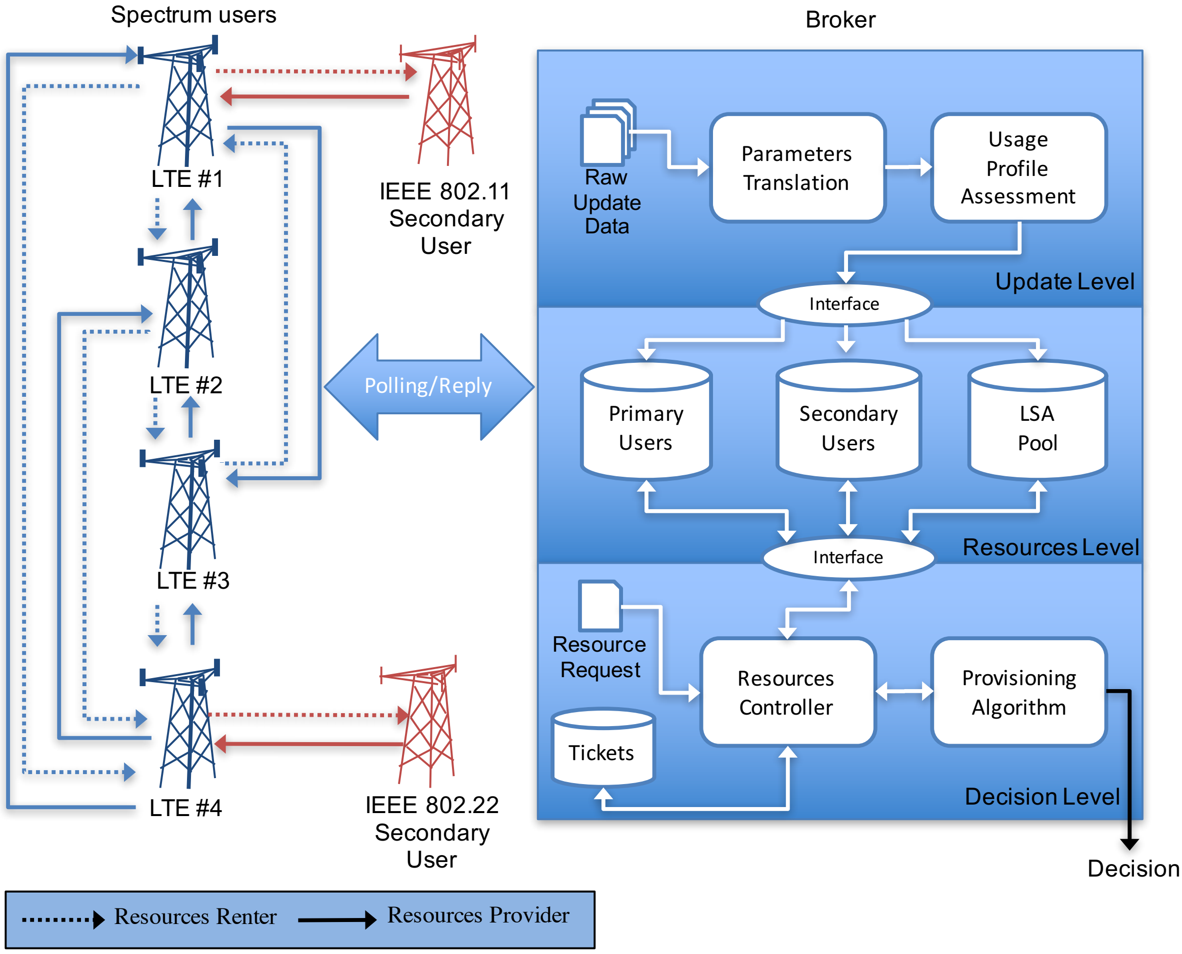


Figure 4.1: Architecture Design

situations where the primary user activity does not chance constantly. The application scenario of the proposed approach fits this criteria in a long-term observation, *e.g* 24 hours, since wireless cellular MNOs deal with predictable traffic in most situations (GUIZANI et al., 2015).

A novel multilevel broker is proposed to play the role of the centralized entity in the resources sharing architecture. Three levels areß defined to provide independent and simultaneous control of different tasks of spectrum sharing management. These levels are named accordingly to the function executed by each one: (I) Update Level, (II) Resources Level, and (III) Decision Level. These levels are interconnected by interfaces which implement the flow of information that allows information exchange among the different levels of the broker.

4.1.1 Update Level

The update level is responsible for collecting operation parameters from the network operators which participate in the spectrum sharing initiative. The updating mechanism is based on the implementation of a polling-based technique, which is controlled by the Parameters Translation block of the broker. This block is discussed in section 4.1.1.1. After the parameters are translated, a cognitive function is applied to assess the usage profile of each MNO. Details on this function are presented in section 4.1.1.2.

4.1.1.1 Parameters Translation

This block allows the configuration of the interval between polls. The precise definition of such interval is crucial to deal with the trade-off between having accurate information about the current resources usage profile of each MNO and the overhead generated by the control information transmitted to update the broker.

Another important function performed by the Parameters Translation block is the translation of the raw update data into useful information to allow the architecture to take proper decisions regarding resources sharing. Therefore, the definition of the structure used by the MNOs to update the broker is very important. The definition this structure is presented in Table 4.1.

Table 4.1: Raw Update Data Structure

Field	**Size**	**Description**
Network Operator ID	1 byte	Uniquely identifies the network operator in the spectrum sharing architecture
Average Delay (expressed in ms)	8 bytes	Updated assessment of the average delay. Performed by the network operator
Average Jitter (expressed in ms)	8 bytes	Updated assessment of the average jitter. Performed by the network operator
Average Throughput (expressed in Mbps)	8 bytes	Updated assessment of the average throughput. Performed by the network operator

Upon receiving the raw update data structured according to the presented organization, the Parameters Translation block performs a SINR estimation in the radio frequency channel and collects the timestamp of the instant when the raw update information was received. These two parameters complement the ones informed by the network operator and are used respectively to estimate the overall network load, considering the system model presented in **??**, and to provide information for historical assessment of the load of each QoS parameter.

4.1.1.2 Usage Profile Assessment

The pre-processed raw data is received by the Usage Profile Assessment block, which applies the concepts of cognition to keep track of the historical information provided by the network operators. This historical information is taken into account to define the current usage profile of the network in order to minimize the effect of abnormal behaviors of the traffic that may occur in realistic operation scenarios. The weight given to the historical information (α) and the weight considered to the most recent update ($1 - \alpha$) can be set and modified as parameters of this block. Equation (4.1) is applied to calculate the weighted load (ℓ) of each considered QoS parameter of the network operators.

$$\ell = \alpha.\sum_{i=1}^{n} \ell_{(t-i)} + (1 - \alpha).\ell_{(t)} \tag{4.1}$$

This equation considers a pre-defined number (n) of historical evaluations of ℓ and performs an exponential smoothing to obtain the weighted load of a given QoS parameter. The same equation is applied to the remaining QoS parameters to obtain the complete assessment of the usage profile of each network operator. The processed usage profile information is then sent to the Resources Level using the proper interface. The value of α was empirically defined based on a set of simulations. The value which provided better results was 0.3, meaning that a weight of 0.7 is given to the most updated information and a weight of 0.3 is given to the historical information. The amount of historical evaluations considered to calculate the load of a giver parameter (n) is set to 5, based on simulation, which showed that this value provides the best results.

4.1.2 Resources Level

For the sake of simplicity in dealing with resources allocation, two classes of users are considered to coexist in the same geographical area in the proposed approach. All the users within the specified geographical area are classified by the resources broker to fit into one of the classes. The first class is called primary users. This class is composed of users who hold a license issued by a regulatory agency given rights to occupy a pre-determined range of the spectrum of frequencies. On the other hand, the second class is called secondary users and comprehend all those network users who aim to opportunistically access the available network resources, playing the role of resources renters.

In order to help accommodating these two classes of users in the same geographical area, the broker must have knowledge about three types of frequencies (UMTS Forum, 2013):

1. **Exclusive Use:** this kind of frequencies relies on licenses granted by regulatory bodies and is controlled by network operators who hold individual usage rights for a specified range of frequencies for a defined period of time in a given geographical area. In such cases, the regulator takes a responsibility to protect the licensed user against interference and provides a legal basis for ensuring a certain level of QoS. The exclusive use is implemented, for example, in the frequency plan of LTE networks.
2. **Shared Use:** refers to the range of spectrum frequencies which are license-exempt. In this case, the right to use the spectrum is afforded to devices that meet certain technical

conditions to share the spectrum and which have a low probability of causing interference to other services. The regulator takes no responsibility for protecting individual users of license-exempt devices against interference and does not provide a legal guarantee for ensuring a certain QoS level. An example of license-exempt application is the 2.4GHz spectrum for the provision of Wi-Fi access service based on the IEEE 802.11 standard.

3. **Exclusive Shared Use:** is the most recent model of spectrum access an is the basis for the so called LSA regime. This kind of frequency range is licensed and works as a complementary source of network resources to MNOs when they face resources shortage. The access to such frequencies is similar to the exclusive use, with the difference that the duration of such access is reduced. The transmissions in the exclusive shared access frequencies are controlled by a regional Network Resources Administrator (NRA). The role of NRA is performed by the broker in the proposed architecture.

The Resources level of the broker is responsible for providing information regarding the users currently operating in the geographical area as well as about the available ranges of frequencies of each type. Therefore, this level implements three databases which are often updated by the Update level and provide the decision level with information about the current resource allocation status in an on-demand basis.

A Primary Users database is specified to store regulatory information regarding the exclusive usage rights afforded to license holders. It is important to note that the proposed architecture allows primary users to share these resources with opportunistic network operators in exchange for profits that may involve financial gains or credit for future resources renting.

The Secondary Users database allows the resources broker to register opportunistic and license-exempt MNOs. Such a registration provides important information for the decision level that allow the control of shared resources access. This database does not store information about shared use ranges of frequencies, since the access to these frequencies is by definition not controlled by outside entities, which is the case of the proposed multilevel broker.

Finally, a LSA Pool database is defined to store information about ESA frequencies. This database is accessed mainly in situations where primary users are in need of complementing their network resources. The available frequencies in the LSA Pool and the conditions for access are established by the regulatory bodies.

4.1.3 Decision Level

Request for resources renting are received and processed by the Decision Level of the multilevel broker. Details on the format of the requests are provided in section 4.1.3.1. The pricing mechanism proposed for the resources sharing architecture is presented in section 4.1.3.2. A model used to assess the resources availability is presented in 4.1.3.3. Specific details of the request are processed by a Resources Controller which is responsible for finding candidate network operators for resources sharing. The features of this controller are discussed in section 4.1.3.4. The Resources Controller takes into account the price of the resources.

4.1.3.1 Resource Request Structure

The Resource Request contains all information demanded by the multilevel broker in order to decide which resources will be designated for sharing taken into account the QoS requirements and the cost. A structure is defined to format such requests as shown in Table 4.2.

Table 4.2: Resources Request Structure

Field	Size	Description
Network Operator ID	1 byte	Uniquely identifies the network operator in the spectrum sharing architecture
Class of Service (CoS)	3 bits	Identifies the class of service demanded by the resources renter. Used for QoS purposes
Maximum Delay (expressed in ms)	8 bytes	The largest delay supported by the application for which the rented resources will be allocated
Maximum Jitter (expressed in ms)	8 bytes	The largest jitter supported by the application for which the rented resources will be allocated
Minimum Throughput (expressed in Mbps)	8 bytes	The lowest throughput supported by the application for which the rented resources will be allocated
Duration (expressed in Hours)	4 bytes	The expected duration of the resources loan
Priority	3 bits	The priority level of the request. The informed value influences on the kind of resources offered by the broker

Two fields defined in the resources request structure deserve further explanation. The Class of Service (CoS) field is 3 bits long to support the three classes defined in the proposed architecture, plus one bit reserved for future use. The classes of service defined in the broker architecture are designed to accommodate traffic from different classes of services defined in different network operators technologies. These classes are specified as follows:

- **001 - Real Time Services (RTS):** is the configuration that provides highest level of QoS guarantees. This class is designed for delay and jitter sensitive real time transmissions, for example VoIP and Video conferences. In this scenario, the Decision Level of the broker is going to consider all resources providers to decide which one is able to provide the QoS level desired by the resource renter. This kind of selection may lead to higher costs to obtain the shared resources.
- **010 - Multimedia Services (MS):** comprehends non real time multimedia services which typically demand high throughput but not strict delay and jitter requirements. Since this class is considered a medium QoS service, the Decision Level is going to prioritize

cheaper network resources comparing to RTS, such as those provided by the LSA pool of frequencies or even those provided by secondary users, such as IEEE 802.22 network operators or IEEE 802.11 networks available in the geographical area.

- **011 - Best Effort Services (BES):** provides the lowest QoS level in the proposed multilevel broker. BES is designed to support best effort transmissions without strict QoS requirements. Therefore, only free or very cheap shared resources will be considered by the provisioning algorithm. For example, preference will be given to obtain resources from shared use frequency ranges.

The second field of the resources request structure that deserves special attention is Priority. The priority of a request is defined by the resources renter and is related to the amount of investment that such network operator is willing to make in order to rent resources from the resources provider. A high priority indicates that the network operator is able to rent more expensive resources than in a low priority situation. This field was defined to be 3 bits long to allow the setup of three values of priority currently defined in the architecture, but also to support future enhancements on the proposed architecture. The currently defined levels of priority are the following:

- **001 - High Priority:** when high priority is set in the resources request structure, the resources controller and the provisioning algorithm will search for resources using all the available network operators and the LSA pool of frequencies. In other words, this means that all the three types of frequencies will be taken into account in the decision process. In this case, the price of the resources will be placed in second plan when deciding which is the best resources renting option for the desired QoS.
- **010 - Medium Priority:** is designed to be used by applications that demand QoS guarantees which are not very strict. In this situation, the decision process will not take into account the more expensive network resources, for example, those belonging to network operators which hold licenses to access exclusive use frequency ranges. Since the price of the resources is taken into account, the preference will be given to shared use and exclusive shared use frequencies.
- **011 - Low Priority:** focuses on finding cheap resources options for renting. In this case, the QoS level will not be the main concern of the decision process, meaning that the network provider which offers the best cost-benefit considering the trade-off between price and QoS will be selected.

Table 4.3 summarizes the features of each CoS and the corresponding priorities.

4.1.3.2 Resources Pricing

The resources pricing mechanism is designed to serve as an incentive to MNOs to share resources. Many proposals on pricing algorithms have been published recently. The majority

Table 4.3: Classes of Services of the Proposed Architecture

Class of Service	**Supported Priorities**	**Exclusive Use Frequencies**	**Shared Use Frequencies**	**Exclusive Shared Use Frequencies**
RTS	High/Medium	X	X	X
MS	all		X	X
BES	Low/Medium		X	

of such proposals can be classified into three groups:

1. **Pricing:** in this approach, the profit gained by resource provider is specified as a price that must be paid in currency by the resources renter in order to access the shared resources (CAO; CHEN; LIU, 2015).
2. **Auction:** in this case, many resources providers advertise information about their available resources and the corresponding price. A resource renter then chooses among the available options (YI; CAI, 2015).
3. **Favors:** this model offers no financial profit to the resources provider. Instead, it is based on favors traded among resources providers which expect to receive similar favors in the future. Generally, the control of such favor exchange is done by implementing the concept of tickets (SINGH et al., 2015).

The proposed architecture relies on the third group of pricing. Such an assumption is justified because the aim of the proposal is to allow sharing among MNOs which will access the spectrum of frequencies for a long time (*e.g.* cellular LTE operators) in the same geographical area. This kind of scenario fits perfectly to the reciprocity demanded by the favors exchange mechanism. Moreover, no payment control is necessary, what simplifies the process of celebrating a dynamic SLA.

In practical terms, the Decision Level of the broker implements a Tickets database, which is updated by the Resources Controller when a new resources sharing transaction is completed. This database logs every transactions and keeps track of the amount of resources shared and received by each MNO. The Tickets database implements a table with the structure shown in Table 4.4 to allow the correlation between two MNOs to decide whether a resource request can be served based on the current tickets balance of the operators.

4.1.3.3 Resources Assessment Model

The approach defined in this thesis demands an accurate assessment of the amount of resources controlled by each operator. For the sake of simplicity, it is assumed that all operators, including the secondary users, are able to access predefined spectrum bands as their main re-

Table 4.4: Tickets Database Table Structure

Field	**Description**
transaction_id	Unique identification of a transaction within the architecture
server_id	Unique identification of the operator which is the resources server
renter_id	Unique identification of the operator which is the resources renter
tickets	Amount of tickets invested in the specific transaction
start_time	Time stamp indicating the expected start of the transaction
duration	Expected duration of the transaction
finish_time	Time stamp indicating the expected end of a transaction

source. This assumption is close to reality, since this kind of allocation is standard for LTE and LSA regime. In this situation, the spectrum frequency is always available to the MNO which is responsible for managing the access of the clients to the spectrum of frequency. The amount of resources is then correlated with the transmission capacity of each operator.

The capacity is modeled considering the Shannon's model, based on an adaptation of the solution presented by Simona et al. (SIOMINA; YUAN, 2012), as defined in equation (4.2). The channel bandwidth (B) is considered to calculate the theoretical channel capacity (C), which is the one of the resources shared in the proposed architecture.

$$C = B \log_2 \left(1 + \frac{P.g}{\sigma^2}\right) * \psi \tag{4.2}$$

In (4.2), P represents the transmission power, g is the gain provided by the transmitting antenna, and σ^2 is the noise power. Besides the SINR, the link efficiency (ψ) is considered to model a more realistic scenario.

The resources demand in a given instant of time ($d(t)$) takes into account the individual demand ($d_i(t)$) of the *ith* active connection of each network operator. The total number of active connections is represented by n. Moreover, the overhead, caused by both cyclic prefix insertion (ϑ_{CP}) and pilot subcarriers used for synchronization (ϑ_{PS}) is considered. Therefore, $d(t)$ is calculated as defined in (4.3).

$$d(t) = \left(\sum_{i=1}^{n} d_i(t)\right) + \vartheta_{CP} + \vartheta_{PS} \tag{4.3}$$

In order to simplify the decision process implemented by the Decision Level of the broker, the resources occupation factor ($\delta(t)$) in a given instant of time is calculated using equation (4.4). It is important to highlight that this equation correlates the current demand ($d(t)$) with the capacity of a network operator (C). The demand is originally calculated in unit of Mb, while the capacity is obtained in terms of Mbps. Therefore, to guarantee the consistency of $\delta(t)$ factor, the demand must be observed during the period of one second, to transform its unit into Mbps before applying the equation.

$$\delta(t) = \frac{(\sum_{i=1}^{n} d_i(t)) + \vartheta_{CP} + \vartheta_{PS}}{B \log_2 \left(1 + \frac{P.g}{\sigma^2}\right) * \psi} \tag{4.4}$$

Situations where δ is close to zero represent that resources are underutilized. On the opposite, a value of δ near to 1 indicates that the resources are compromised, what may lead to resources scarcity. This values does not surpass 1, because a connection admission control is used to avoid overloading the resource providers. The proposed system model allows to simulate the amount of resources which are available to each MNO. These resources must be managed to first accommodate the MNO clients, since they play the role of primary users. After serving these primary users, the resources may not be completely compromised. In this situation, the network operator may play the role of a resources server by allowing opportunistic users belonging to other network operator to take advantage of the underutilized resources. These opportunistic users are provided with access to the network resources through a leasing process that might involve a profit to be earned by the resources server. To allow the communication among the serving network operator and the opportunistic users, all network operators involved in the transaction are required to implement a signaling protocol, which is detailed in Section 2.2.1.

Another important aspect that must be considered in the system model is the overhead generated by the polling/reply based update process used by the Update Level. This approach imposes a trade-off between the interval between updates and the resulting overhead. It is common sense that the more updates are received, the higher the generated overhead. Beyond the interval, the overhead is also affected by the number of network operators managed by the broker and by the size of the update message. The aim of the proposed architecture is to keep the overhead low in order to avoid interfering with the legitimate traffic of the MNOs.

4.1.3.4 Resources Controller

Every time a resources request is received, it is processed by a Resources Controller. This entity of the broker has direct access to the Tickets database. Through the proper interface it is also able to retrieve information from the databases in the Resources Level of the broker. The aim of the resources level is to have updated knowledge about the network resources status and feed the Provisioning algorithm with possible resources servers for a given request. Towards this aim, the execution of the Resources Controller follows the specification of algorithm 1.

The inputs of the Resources Controller algorithm is a resource request. Such algorithm interfaces with the Resources Level and therefore is able to access the Primary User, Secondary User, and LSA databases. In the first stage, the algorithm classifies the resource request according to the priority informed by the requesting operator considering the CoS (as defined in

Algorithm 1 Resources Controller

Require: r ▷ A struct containing a resource request as specified in Table 3.2
Require: get_mno([databases], [QoS Requirements]) ▷ A Procedure that returns operators which match the QoS requirements

$p \leftarrow r.Priority$ ▷ Gets the priority of the request
$d \leftarrow r.Delay$ ▷ Gets the largest delay supported by the application
$j \leftarrow r.Jitter$ ▷ Gets the largest jitter supported by the application
$t \leftarrow r.Throughput$ ▷ Gets the slowest throughput supported by the application
switch p **do** ▷ Tries the standard types of operators, as defined in Table 3.3
    case High
        $mno \leftarrow get\_mno([Primary, Secondary, LSA], [d, j, t])$
    end case
    case Medium
        $mno \leftarrow get\_mno([Secondary, LSA], [d, j, t])$
    end case
    case Low
        $mno \leftarrow get\_mno([Secondary], [d, j, t])$
    end case
end switch
for all mno **do**
    $cost(i) \leftarrow [mno.Id, mno.Tickets]$ ▷ Gets the amount of tickets to be paid
end for
if $cost = \varnothing$ & $p =$ High **then**
    return 0
else if $cost = \varnothing$ & $p =$ Medium **then**
    $mno \leftarrow get\_mno([Primary], [d, j, t])$
    for all mno **do**
        $cost(i) \leftarrow [mno.Id, mno.Tickets]$ ▷ Gets the amount of tickets to be paid
    end for
    if $cost = \varnothing$ **then return** 0
    end if
else if $cost = \varnothing$ & $p =$ Low **then**
    $mno \leftarrow get\_mno([Primary, Secondary], [d, j, t])$
    for all mno **do**
        $cost(i) \leftarrow [mno.Id, mno.Tickets]$
    end for
    if $cost = \varnothing$ **then return** 0
    end if
end if
return provisioning($cost$) ▷ Calls the Resources Provisioning Algorithm

Table 4.3). The function called $get\_mno(< Types\ of\ Resource >, < QoS\ Parameters >)$ is responsible for searching the databases of Resources Level to retrieve candidate resource providers which have enough resources to guarantee QoS. This retrieval of information takes into account the restrictions imposed by QoS parameters specified in the resources request, *i.e.* maximum delay, maximum jitter, and minimum throughput.

As stated in Table 4.3, requests from an application using RTS class of service typically have high priority which will indicate to the algorithm that it should try to obtain resources from all the available service providers databases. MS requests, in turn, can be described as a medium priority class of service, what leads the algorithm to prioritize borrowing resources from Shared Use Frequencies (Secondary Users) and Exclusive Shared Use Frequencies (LSA). Finally, BES is used for low priority services and, therefore, the algorithm will try to obtain resources only from secondary users in the first attempt. In cases where the first attempt to find resource providers in medium and low priority requests returns no result, the algorithm will expand the selection ranges to consider more expensive service.

After accessing the Resources Level databases, the algorithm calculates the cost of each resource available. The cost (ζ), in the proposed architecture, follows the model of favors exchanged among resources providers. The cost of each favor is influenced by three main factors: (I) the type of service provider (ρ), (II) the amount of resources currently compromised by the selected resources provider (ℓ) at a given instant of time, and (III) the priority of the request (κ). ζ is calculated using (4.5).

$$\zeta = \rho.\kappa.\left(\frac{\ell_{RTS} + \ell_{MS} + \ell_{BES}}{L}\right) \tag{4.5}$$

In this equation, L represents the total amount of available resources in a given resources provider. The values related to the priorities and types of service providers are summarized in Table 4.5.

Table 4.5: Parameters used to calculate the cost

ρ		κ	
Type of Provider	**Value**	**Priority**	**Value**
Shared Use	1	Low	1
Exclusive Shared Use	2	Medium	2
Exclusive Use	3	High	3

It is important to highlight that the resources broker estimates the initial cost without considering the duration of the loan, since this information is not accurate at this first stage of analysis, because the expected duration may differ from the real duration of a transaction in a realistic scenario. Considering that the instantaneous load of a network operator may vary from 0% to 100% of the available resources, the cost will vary between 0 and 9 (applying equation

4.5). After the transaction if finished, the initial cost is multiplied by the duration of the loan. Since the duration of a sharing transaction is computed in unit of hours by the broker, the final price of the favor, as a consequence, will be computed in a unit of tickets per hour. In order to guarantee fairness in the resources sharing transactions among network operators, the favor will be registered by the broker considering its final cost.

The Resources Controller algorithm generates an array of candidate resources providers. Each entry of the array is composed of the unique identification of the service provider and the cost of this transaction. The resulting array is used as the input to the Resources Provisioning algorithm, which is called once the candidate resources provider array is ready. The aim of the Resources Provisioning algorithm is to take a decision on which resource providers is the best to serve a specific request.

4.1.3.5 Resources Provisioning

The resources provisioning algorithm receives from the resources controller a list of candidate resource providers which in a first analysis have enough resources to guarantee the QoS demanded by a resources renter. This list is composed of arrays containing the unique identification of the operator within the proposed architecture and the cost of each transmission. Since the analysis conducted by the resources controller takes into account only the current capacity of each operator, in the provisioning level, further analysis is conducted in order to analyze aspects related to the expected traffic load of the candidate resources providers during the duration of the resources rental.

This function of the broker is constantly running with the goal of taking in advance decisions to allow the architecture to work in network environments where fast evacuation of frequencies may be required. This kind of evacuation is expected especially in situations where exclusive or exclusive shared use frequencies are being rented. This in advance decision demands this level of the broker to forecast the traffic of the network operators in order to identify possible evacuation routes. Such forecast requires knowledge about the historical traffic load, which is stored in the resources level. The resulting values also serve as inputs to the Resources Provisioning Algorithm.

Such algorithm supports both the usage of traffic models or traces to describe expected behavior of the network operators. In order to forecast the traffic behavior, a MLRM is implemented using Matlab. This model is based on a traffic measurement Y, which is related to a single predictor X for each observation. Therefore, the conditional mean function can be described as in (4.6), where α is the intercept and β is the coefficient.

$$\mathrm{E}[Y \mid X] = \alpha + \beta X \tag{4.6}$$

Considering that multiple predictors (n) are available from the traffic models or from the traces, a multiple linear regression model is considered, according to (4.7).

$$\mathrm{E}[Y \mid X] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \tag{4.7}$$

The variability of the ith measurement Y around its mean value is specified in (4.8).

$$\mathrm{E}[Y \mid X_i] = \alpha + \beta_1 X_{i,1} + \beta_2 X_2 + \cdots + \beta_n X_{n,i} + \epsilon_i \tag{4.8}$$

In this case, the error assumptions for ϵ_i are that $\mathrm{E}[\epsilon_i] = 0$ and $\mathrm{var}(\epsilon_i) = \sigma^2$. The accuracy of the forecast can be measured by the mean absolute percent error (η), which is given by (4.9). In this equation, e_t represents the actual network occupation based on network traces or traffic models and y_t is the forecast occupation of the same network in a given instant of time.

$$\eta = \frac{1}{n}\left(\sum_{t=1}^{n}\left|\frac{e_{(t)}}{y_{(t)}}\right|\right) \tag{4.9}$$

The resulting forecast points compose a continuous traffic function, $f(x)$, which describes the occupied area of each analyzed network. In this context, let $f : D \to R$ be a function defined on a subset D of R and let $I = [a, b]$ be a close interval contained in D. This closed interval represents the start and the end time of the forecast. Finally, let $P = \{[x_0, x_1], [x_1, x_2], \cdots, [x_{n-1}, x_n]\}$ be a partition of I such as $P = \{a = x_0, x_1, \cdots, x_n = b\}$. Thus, a Riemann sum (S) of f over I with partition P is defined in (4.10).

$$S = \sum_{i=1}^{n} f(x_i^*)(x_i - x_{i-1}) \tag{4.10}$$

When the number of points in P increase indefinitely, it is possible to apply (4.11) to calculate the expected occupied area of each network, which can be related to the occupied network capacity.

$$A_{occupied} = \int_a^b f(x)dx = \lim_{x\to\infty}[s^*(P, f)] \tag{4.11}$$

This value is normalized considering the total capacity (A_{total}) area of each network operator. Its complement therefore represents the percentage of available resources of a given network. Let $\Theta = \{o_0, o_1, \cdots, o_{n-1}, o_n\}$ be a set of network operators. Thus, the free capacity percentage of the network operators is given by (4.12).

$$\forall o \in \Theta, A_{free}(o) = 1 - \left(\frac{A_{occupied}(o)}{A_{total}(o)}\right) \tag{4.12}$$

As previously mentioned, three CoS are defined to accommodate different types of traffic regarding to the QoS requirements. (I) RTS, to support delay and jitter sensitive real time transmissions, (II) MS, comprehending real time services with high throughput but no strict delay and jitter, and (III) BES, designed to support best effort transmissions without strict QoS requirements. Based on the CoS requirements and on the amount of free resources of each operator a resources provisioning algorithm is implemented, as defined in algorithm (2).

Algorithm 2 Resources Provisioning Algorithm

Require: Θ ▷ A list of candidate resource providers
Require: r ▷ A resource request
Require: $A_{total}(o)$ ▷ The total amount of resources of each operator $\in \Theta$
Require: $A_{occupied}(o) = \int_a^b f(x)dx = \lim_{x\to\infty}[s^*(P, f)]$ ▷ Occupied resources of each operator $\in \Theta$
1: $selected\_operator = \varnothing$
2: $\Theta = sort(\Theta, cost, asc)$
3: $c \leftarrow r.CoS$; $d \leftarrow r.Delay$; $j \leftarrow r.Jitter$; $t \leftarrow r.Throughput$
4: **switch** c **do**
5:     **case** RTS:
6:         **for all** $o \in \Theta$ **do**
7:             $A_{free}(o) = 1 - \left(\frac{A_{occupied}(o)}{A_{total}(o)}\right)$
8:             $delay(o) = get\_resources\_level(o, delay)$
9:             $jitter(o) = get\_resources\_level(o, jitter)$
10:             **if** $A_{free}(o) \geq t$ & $delay(o) \leq d$ & $jitter(o) \leq j$ **then**
11:                 **return** o
12:             **end if**
13:         **end for**
14:     **end case**
15:     **case** MS:
16:         **for all** $o \in \Theta$ **do**
17:             $A_{free}(o) = 1 - \left(\frac{A_{occupied}(o)}{A_{total}(o)}\right)$
18:             **if** $A_{free}(o) \geq t$ & $delay(o) \leq d$ **then**
19:                 **return** o
20:             **end if**
21:         **end for**
22:     **end case**
23:     **case** else:
24:         **for all** $o \in \Theta$ **do**
25:             $A_{free}(o) = 1 - \left(\frac{A_{occupied}(o)}{A_{total}(o)}\right)$
26:             **if** $A_{free}(o) > max\_operator$ **then**
27:                 $selected\_operator = o$
28:             **end if**
29:         **end for**
30:     **end case**
31: **end switch**
32: **return** $selected\_operator$

The Resources Provisioning Algorithm receives a list of candidate resources providers (Θ) which is generated by the Resources Controller, by executing Algorithm 1. This list is sorted in line 3 of Algorithm 2 in order to prioritize service providers which are offering low cost resources. Based on this ordered list of candidate resources providers, on the CoS and QoS restrictions extracted from to the parameters received in the resources request, the algorithm is going to search for resources providers which are able to guarantee the QoS demanded by the resources renter.

In order to gather the update occupancy status of the network resources, a function called $get\_resources\_level$ is defined to access the databases in Update Level of the proposed architecture and retrieve the relevant information. The CoS based analysis is conducted in lined 4 - 31 of the proposed algorithm. The difference of approach for each CoS is related to the kind of QoS parameters that each class of service takes into account, as follows:

- **RTS:** throughput, delay, and jitter are considered (line 10);
- **MS:** throughput and delay are considered (line 18);
- **BES:** only throughput is considered (line 26).

The logical approach of the algorithm is to select the lowest cost among the candidate service providers which are able to guarantee the QoS requirements of the resources renter. Details on the implementation of the proposed architecture are provided in the next section.

4.2 Simulation Tool

In order to evaluate the concepts proposed in this thesis, a simulation tool was designed and implemented using Matlab. Figure 4.2 shows the UML use case diagram which describes the design of the simulation tool.

The inputs of a simulation run are defined in a scenario configuration file. The entries which compose the structure of this file are presented and explained in Table 4.6.

Table 4.6: Scenario Configuration File Structure

Entry	**Description**
N_BS	Number of network operators base stations in the simulation
N_Chann	Number of shared channels available for allocation
Alloc_Alg	Allocation algorithm to be used by the resources controller
N_VoIP	Number of VoIP clients
N_Video	Number of Video clients
N_HTTP	Number of HTTP clients
SNR	Signal to Noise Ratio of the user during the simulation
Log	Log file used to store the results and statistics

The Simulation Control class reads the configuration file and sets up the simulation environment. Another important feature of this class is the one which allows multiple scenarios to be executed during a simulation run. Each line of the configuration file corresponds to a single simulation scenario. These lines are read one by one and after the execution of the scenario corresponding to a given line, the simulation control resets the simulation environment and reads the next line for a new scenario setup. This feature is important to improve the efficiency of results gathering, since all the desired simulation scenarios can be previously configured,

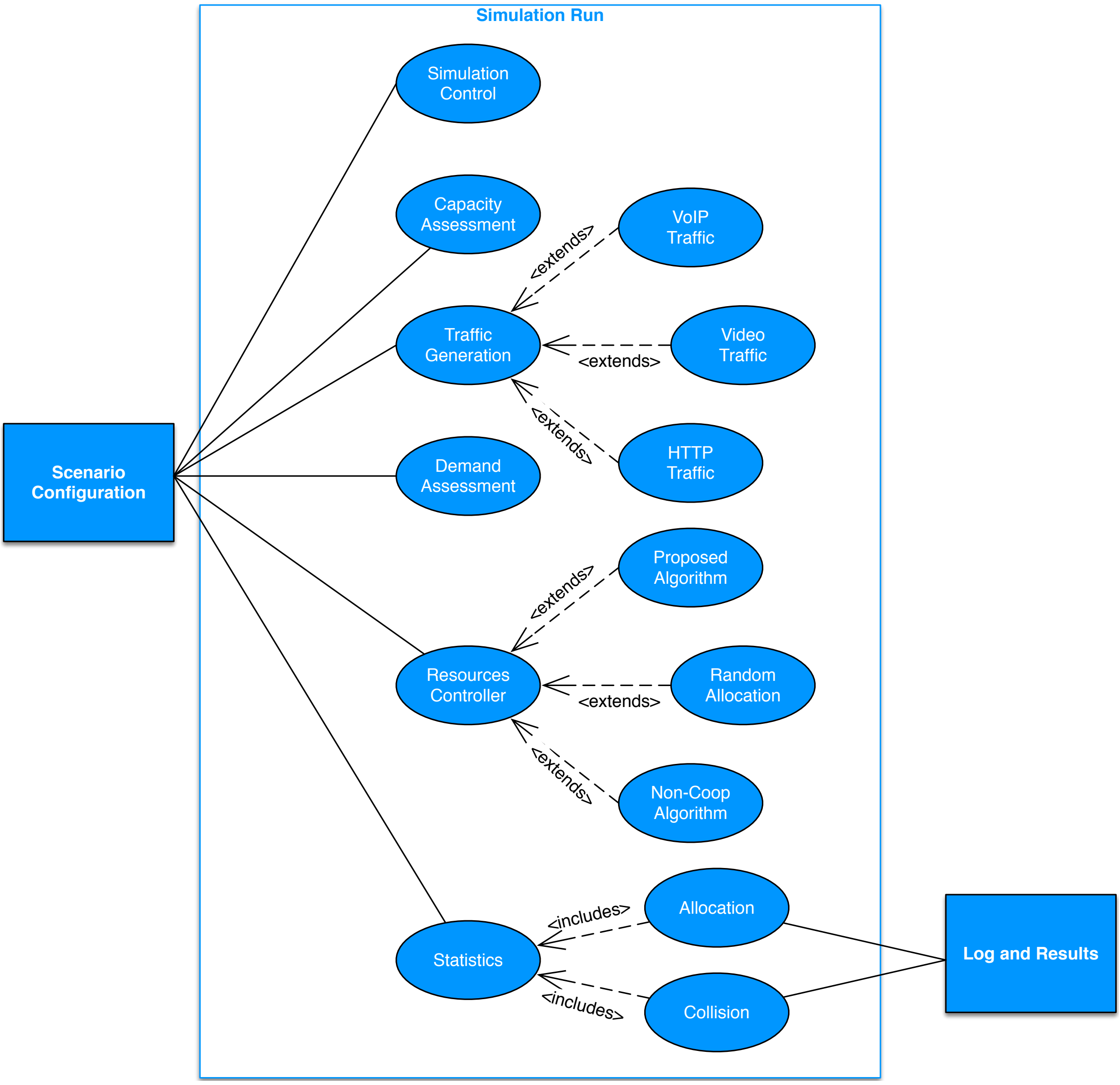


Figure 4.2: Simulation Tool Use Case Diagram

allowing all results to be gathered at once.

The first step towards resources allocation and consequently to obtain the results of each simulation scenario is the channel capacity assessment. This class implements the theoretical channel capacity formula specified in equation (4.2). This information is stored in the databases which compose the Resources Level of the proposed Architecture, as previously explained in Figure 4.1. The theoretical channel capacity also feeds the Demand Assessment class of the simulation tool, since this value along with the demand of the users is mandatory in order to take decisions regarding resources allocation.

The behavior of the network clients is simulated considering the traffic demands of each client. The Traffic Generation class is responsible for analyzing the traffic demands and instantiate the classes which extend the traffic generation. In the proposed architecture, three kinds of traffic are generate to simulate the behavior of three classes of service: (I) VoIP traffic, to simulate RTS; (II) Video traffic, to simulate MS; and (III) HTTP traffic to simulate BES. Each kind of traffic is implemented by a class which extend Traffic Generation and follow the specifications for traffic generation of the *System Evaluation Methodology* document, published by the WiMAX Forum (WiMAX Forum, 2008) as well as realistic traces (SCHULMAN; LEVIN; SPRING, 2009). Further details on the traffic generation models will be presented in Section 5.1.1.

Based on the outputs of the Traffic Generation, the Demand Assessment class is implemented according to the definition of equation (4.3). The output of such equation is used along with the output of the Capacity Assessment class to calculate the resources occupation factor (δ) as defined in (4.4). δ is then used as one of the inputs of the Resources Controller class, which is responsible for taking the decisions regarding the resources allocation of the proposed architecture. This resources controller class is extended by classes which implement the concepts of different scheduling and allocation algorithms. In order to analyze the performance of the architecture, considering the case studies of this thesis, three algorithms have been considered: (I) the novel algorithm proposed in this thesis, (II) the random allocation algorithm, and (III) the non-cooperative algorithm proposed in a related work.

After resources are allocated, a specific class is used to collect statistics regarding the simulation, which will lead to the results used for analyzing the performance of the proposed solution. This Statistics class includes two sub-classes, aimed at collecting Allocation related and Collision related statistics, respectively. The outputs of the sub-classes are stored in a log file, which is later parsed in order to analyze the results and produce the graphs used to explain the obtained results.

5 PERFORMANCE EVALUATION

In this chapter, the performance of the proposed solution is evaluated considering three different case studies to cover both homogeneous and heterogeneous network scenarios. Towards the performance evaluation, in Section 5.1 the simulation scenario is explained. Then, in Section 5.2 results regarding all three scenarios are presented and discussed.

5.1 Simulation Scenario

In this section, details on the simulation model are presented. First, the traffic models used for simulation are described. After, the simulation parameters setup of the scenario used to evaluate the proposed solution are discussed.

5.1.1 Traffic Models

In order to simulate the behavior of the proposed architecture it is mandatory to properly model the traffic demands of the MNOs which are participating in the resources sharing initiative. The traffic model must consider the connection arrival and the amount of traffic demanded per connection. The traffic models used in the simulation scenarios are based on the *System Evaluation Methodology* document, published by the WiMAX Forum (WiMAX Forum, 2008). This model was selected because it is based on realistic measurements and provides a solid base to estimate the actual traffic demanded by the different users.

In the simulations, three different kinds of traffic are considered to meet the CoS defined in the proposed architecture. The amount of traffic generated for each CoS is a parameter of the simulation tool. Based on the typical traffic load of current networks, as described by the UMTS forum (UMTS Forum, 2011a), the distribution of the load was considered to be 20% of RTS, 20% of MS, and the remaining 60% was classified as BES traffic in most scenarios. An exception is allowed for the third case study, because in this case, only video traffic needs to be simulated. The following subsections describe the characteristics of the traffic models for each CoS.

5.1.1.1 Real Time Services Traffic Model

The traffic generation for RTS CoS is modeled by VoIP transmissions encoded using the Adaptive Multi Rate (AMR) codec. This codec considers realistic conversations, which present ON/OFF behavior. This behavior is modeled considering the activity of the speech in the conversations using the aforementioned codec. The ON/OFF behavior can be modeled using a simple two-state Markov chain, as shown in Figure 5.1.

In this case, the states of the Markov Chain are called Talk (ON) and Silence (OFF). The

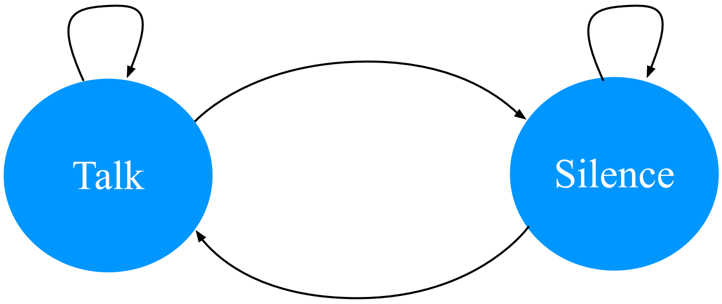


Figure 5.1: ON/OFF Behavior Markov Chain

mean duration of each state is defined as the mean duration of talk (t_t) and the mean duration of silence (t_s), where only a background noise is transmitted. In terms of implementation of the traffic model, it is necessary to calculate the number of packets that might be generated in each state. The number of packets during Talk state N_t, is defined in (5.1).

$$N_t = \frac{t_t}{t_f} + h \tag{5.1}$$

In this case, t_f represents the duration of each VoIP frame and h is the hangover time. The amount of packets transmitted as background noise (N_s) is defined as (5.2).

$$N_s = \frac{\left(\frac{t_s}{t_f} - h\right)}{8} \tag{5.2}$$

In (5.2), t_s represents the duration of the silence during a conversation. For simulation purposes, AMR codec is used to model conversations with the duration of each period modeled according to an exponential distribution with mean of 1026 ms of talk (ON period) and 1171 ms of silence (OFF period). The parameters of the simulation are presented in Table 5.1. These parameters are the same which were defined in the *System Evaluation Methodology* document, published by the WiMAX Forum (WiMAX Forum, 2008).

5.1.1.2 Multimedia Services Traffic Model

The second CoS considered in the proposed architecture is MS. To model MS, a traffic model was used to consider the transmission of on demand videos encoded with MPEG-4, since this coded is one of the most accepted and efficient for video encoding. The parameters of a video transmission may differ from one trace to another. Therefore, for simulation purposes, a trace obtained from a talk show (Kim, Bong Ho, 2007) is considered. In this trace, two resolutions are available 176x144 pixels and 320x240 pixels.

For the sake of simplicity, in this thesis proposal, only the first one was considered. Each of the videos has variable length, varying exponentially from 15 s to 60 s. The selected display

Table 5.1: VoIP traffic parameters

Parameter	Value
Call Holding	Exponential: μ = 210s
Codec	AMR
Frame duration	20ms
Talk duration (ON)	Exponential: μ = 1026ms
Silence duration (OFF)	Exponential: μ = 1071ms
Silence suppression	ON
Embedded Protocols	RTP/UDP/IP
Speech Activity	47.17%
MAC Header (ON)	42 bytes
MAC Header (OFF)	16 bytes

size of the video leads to a mean frame size of 2.725 Kbytes after the video is compressed. Considering 8 bits of color depth, each transmission demands an average channel capacity of 7.6 Mbps. The parameters used for simulations are summarizes in Table 5.2 (WiMAX Forum, 2008).

Table 5.2: Video traffic parameters

Parameter	Value
Video length	Truncated Exponential: 15 - 60s
Video resolution	176x144 pixels
Codec	MPEG-4
Protocol	TCP
Direction	Downlink
Color Depth	8 bit
Mean uncompressed frame size	38.016 kbytes
Compression ratio	13.95
Mean compressed frame size	2.725 kbytes
Frames per second	25

Let f_c be compressed frame size, t_{fs} the amount of frames transmitted per second, and d_v the duration of a transmitted video. The file size corresponding to the transmission of a given video (T_v) can be calculated using 5.3.

$$T_v = f_c.t_{fs}.d_v \tag{5.3}$$

5.1.1.3 Best Effort Services Traffic Model

The traffic model for BES CoS is defined considering the characteristics of HTTP, which is one of the most used protocols in the Internet. In order to carry out simulations, in this thesis, HTTP is modeled to reflect the behavior or web browsing users. The literature shows, based on realistic measurements, that typical web pages are composed of large amounts of objects with variable size (Kim, Bong Ho, 2007) (3GPP, 2008). HTTP traffic model is complex because it must consider two levels of behaviors: (I) user level and (II) IP level.

Web pages are composed of a main object which carries the HTTP code. This main object makes references to other objects, called embedded objects, in order to properly format and present the web page in the browser. By definition, every web page is composed of one main object and a variable amount of embedded objects. The size of each type of object may vary according to the purpose of the web site and the amount of information displayed.

In the user level model, the main concern is to model the ON/OFF behavior of human interaction with the web pages. The ON periods represent the transmission and parsing time of objects which are embedded in the web page. On the other hand, OFF periods are related to the activity of reading the contents of a web page. The values of the user level HTTP modeling are defined based on realistic measurements found in the literature and are summarized in Table 5.3.

Table 5.3: HTTP traffic parameters

Component	**Distribution**	**Parameters**	**PDF**
Main object size	Truncated Lognormal	Mean = 10710 bytes SD = 25032 bytes Min = 100 bytes Max = 2 Mbytes	$\sigma = 1.37$ $\mu = 8.37$
Embedded object size	Truncated Lognormal	Mean = 7758 bytes SD = 126168 bytes Min = 50 bytes Max = 2 Mbytes	$\sigma = 2.36$ $\mu = 6.17$
Number of embedded objects	Truncated Pareto	Mean = 5.64 Max = 53	$\sigma = 1.1$ $\mu = 55$
Reading time	Exponential	Mean = 30 s	$\mu = 0.033$
Parsing time	Exponential	Mean = 0.13 s	$\mu = 7.69$

At the IP level, the most important feature to be considered is the distribution of packet sizes, which are related to the network's Maximum Transmit Unit (MTU). 3GPP has observed, based on realistic measurements (3GPP, 2008), that approximately 76% of the packets follow the default MTU for Ethernet, which is 1500 bytes. The remaining packets typically use a MTU of 576 bytes. These values were considered for simulation purpose. It is important to highlight that 40 bytes of IP packet header were considered in every packet. Therefore, the actual packet payloads were set to 1460 and 536 bytes, respectively.

5.1.2 Simulation Parameters Setup

The proposed architecture simulation tool is defined analytically. Therefore, to guarantee a simple and precise implementation, Matlab was chosen as the most suitable tool to analyze the performance of the proposal. The simulation model was implemented using the specification described in 5.1 and aforementioned traffic models and traffic traces to describe the traffic classified into each CoS. The stochastic nature of the traffic models assures that even when the same parameters are considered for traffic generation, the resulting traffic will be different for each network operator which is simulated. This leads to different instantaneous traffic loads and therefore, it is possible to simulate resources sharing among these operators.

LTE and IEEE 802.22 MNOs follow basic configuration parameters. The frame duration is 10 ms and the transmissions are carried out in a 10 MHz wireless channel. Wi-Fi operators are implemented according to the specifications of IEEE 802.11 standard. LSA channels are considered to allow the usage of 10MHz channels. The number of connections was varied between 50 and 500 to allow the evaluation of the performance of the architecture in different traffic load scenarios. This load is allocated in the frames using a simple allocation algorithm which is based on First In First Out (FIFO) queuing discipline. This choice of allocation algorithm is due to the desire of simplicity, since details on the frame allocation are not on the scope of this thesis. A link efficiency of 80% was considered based on the traffic analysis conducted by Mogensen *et al.* (MOGENSEN et al., 2007).

More generic parameters were also defined to evaluate the performance of the proposed architecture in different scenarios. The traffic rate per CoS reflects the models typically used in the literature, *i.e.* RTS = 20%, MS = 20%, BES = 60% in most scenarios. An exception is conceded in the last evaluated scenario, where only RTS traffic is considered due to the characteristics of the application. Since the traffic models are stochastic, it is considered a confidence interval of 95% for traffic generation. The measured SINR varied from 10 dB to 30 dB with an antenna gain of 14 dB (MOGENSEN et al., 2007). The key parameters used to setup the simulation scenarios are summarized in Table 5.4.

All simulations were implemented and executed using Matlab version 2014a running on 2.7 GHz Intel Core i5 processor with 8 GB, 1600 MHz, DDR3 memory. Five different machines using MAC OS 10 operating system were used to run the simulation model and obtain the results logs.

Table 5.4: Simulation Setup Parameters

Parameter	**Value**
Frame duration	10 ms
Channel Bandwidth	10 MHz in most scenarios
Number of Connections	variable from 50 to 500
Classes of service	RTS, MS, and BES
Traffic rate per class of service	RTS = 20%, MS = 20%, BES = 60%
Traffic allocation in the frame	FIFO queuing discipline
SINR	variable: 10 dB to 30 dB
Antenna Gain	14 dBi
Link efficiency	80%
Confidence Interval (traffic models)	95%

5.2 Case Studies

Three case studies are conducted to evaluate the performance of the proposed resources sharing architecture in different network scenarios. The first one is discussed in 5.2.1 and consists in a homogeneous scenario composed of different LTE MNOs which coexist in a given geographical area. The scenario, presented in 5.2.2, comprises a heterogeneous network environment composed of LTE MNOs, LSA incumbent users, and Wi-Fi access points. Finally, the third evaluation scenario is inserted in the context of video surveillance in smart cities and is composed of a pool of MNOs which implement various technology. In this scenario, presented in 5.2.3, the goal is to evaluate whether the proposed architecture is able to select which operator should handle the transmission of the videos taking into account both the required QoS and the costs involved in the resources sharing.

The performance evaluation is presented based on the Goals/Questions/Metrics (GQM) model, in order to clarify how the objectives and contributions of this thesis are addressed. Figure 5.2 illustrates groups of goals of this thesis, the corresponding questions, and the metrics used to answer each question.

For the sake of performance evaluation, the goals of this thesis are organized in three groups: (I) to analyze whether or not the proposed architecture impacts on the QoS provided to the network clients; (II) how to reduce the overhead generated during message exchanging between the resources broker and the MNOs; and (III) to improve the resources sharing in comparison to related works.

To deal with the first group of goals, two questions are defined:

1. Which are the most relevant QoS metrics? To answer this question, three important QoS metrics are analyzed: throughput, delay, and jitter.
2. Is the number of collisions and issue to be considered? The number of collisions directly

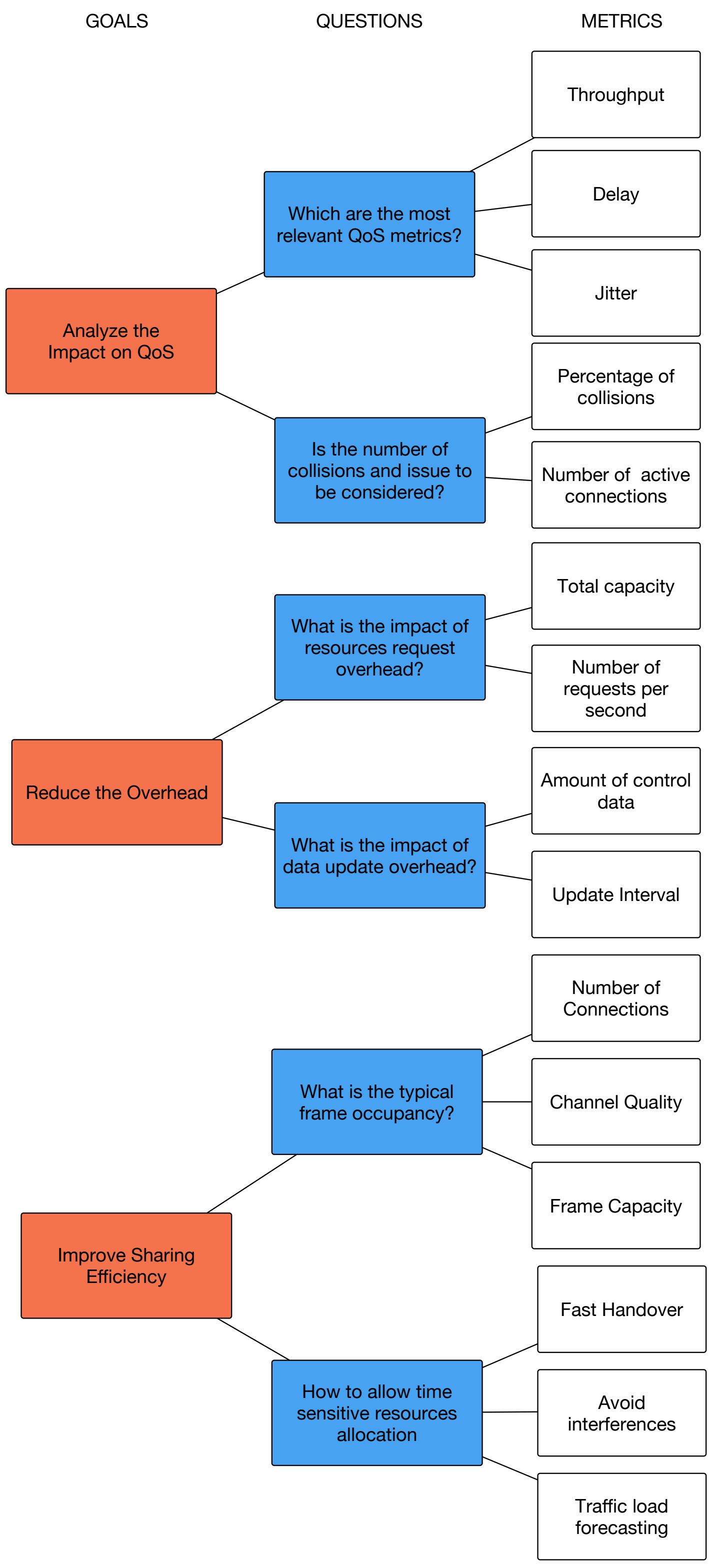


Figure 5.2: GQM model of the proposed solution

affect the transmission quality and therefore is a very important issue regarding QoS analysis. Thus, the percentage of collisions is analyzed taking into account the number of active connections.

The second group of goals is addressed considering the following set of questions:

1. What is the impact of resources request overhead? The control message exchange overhead is considering relating the overall network capacity and the amount of requests sent and received by the resources broker.
2. What is the impact of data update overhead? In this case, the size of structure used to send updated information about the MNOs which are participating in the resources sharing initiative is considered. Moreover, the interval between updates is also an important metric to properly calculate this kind of overhead.

Two questions and six metrics are considered to deal with the final goals of this thesis, which are related to improving resources sharing efficiency.

1. What is the typical frame occupancy? This question is addressed by taking into account the amount of client devices using shared resources. The channel quality experienced by these devices and the frame capacity of the MNOs are also considered.
2. How to allow time sensitive resources allocation? This is achieved by implementing a fast handover solution to reduce the probability of interfering with other devices, especially those belonging to the incumbent user of a given spectrum frequency. The solution presented in this thesis uses MLRM to forecast the traffic load of potential traffic steering routes.

The GQM model and related questions and metrics are considered in the following sections in order to evaluate the performance of the proposed solution.

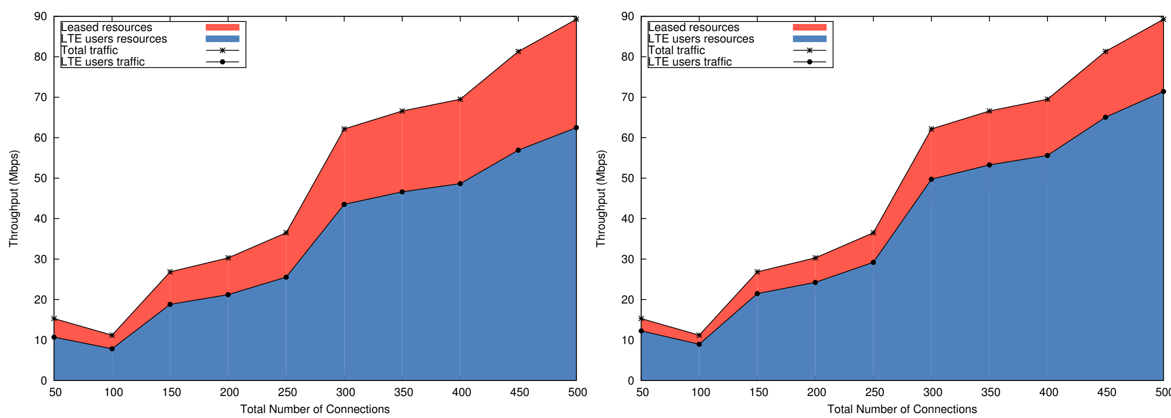
5.2.1 LTE-Advanced Operators Coexisting in the same Geographical Area

In this scenario, the performance of the proposed architecture is evaluated considering a network topology where four LTE MNOs coexist in the same geographical area. Three types of results are presented. First, simulation results are presented to analyze whether underutilized resources exist in different network traffic loads. In the sequence, the results are compared with two related works. In the third approach, the control information overhead caused by the proposed architecture is evaluated. The second kind of results consists in a comparative performance evaluation considering the proposed architecture and two related works.

In all simulations related to this case study, the traffic demand is originated from a variable number of connections generated by different network users. These connections are modeled according to the traffic models previously explained and represent the aggregated traffic which may belong to three CoS defined in the architecture. The aggregated traffic presents a behavior in which peaks that surpass the frame capacity may occur. To deal with this situation, a simple algorithm was implemented to allocate traffic within the frame considering the theoretical frame capacity as a limitation. For the sake of simplicity, this algorithm is based on the concept of FIFO, meaning that data is allocated as it is generated until the frame is full. In this case, an allocation efficiency of 80% is considered (MOGENSEN et al., 2007) and the theoretical frame capacity is calculated using equation (4.2). It is important to emphasize that the proposed architecture allows the substitutions of such algorithm.

The first kind of results shows the amount of resources used by LTE primary users and consequently, the amount of underutilized resources, which may be leased to secondary users. This analysis takes into account two scenarios to represent different loads in the network managed by a LTE MNO. The second analysis considers the amount of unused frame capacity when the amount of connections is varied in a given LTE network. The aim of this analysis is to show that in certain scenarios, the strategy of leasing resources can be very advantageous to both the resources server and the resources renter.

The first aspect analyzed is the network throughput considering a variable total number of connections. The outcomes of the simulations are presented in Figure 5.3. The results show the amount of resources used by LTE primary users and the amount of resources that may be leased in a homogeneous networking scenario that considers a theoretical capacity of 113Mbps, calculated taking into account the resources assessment model presented in 4.1.3.3. The traffic load of LTE network was also varied to consider a total load of 70% in Figure 5.3 (a) and a total load of 80% in Figure 5.3 (b).



(a) LTE traffic load of 70%

(b) LTE traffic load of 80%

Figure 5.3: Underutilized resources considering a total frame capacity of 113Mbps

Results show that the amount of used resources is proportional to the total number of connections until about 80% of the network capacity is allocated. This value is very close to the maximum theoretical capacity, once the allocation efficiency is 80% and the maximum network throughput is 113Mbps in this scenario. This result is important because it shows that even in situations where the network load is high, the theoretical limit of the network is not fully allocated to primary users. This result shows that resources underutilization is a reality in LTE and consequently MNOs have the opportunity of leasing unused resources.

As can be seen in the figure, the amount of resources allocated to LTE primary users varies from 9% to about 54 % of the available network throughput when the proposed resources sharing architecture is not implemented, *i.e.* when only resources demanded by LTE network primary users are considered. On the other hand, results obtained after the proposed architecture is implemented show that up to 25% of the resources can be leased to opportunistic secondary users when the LTE-Advanced traffic load is 70%. When this load increases to 80%, the gain is still observed, reaching values of up to 18% of the resources which may be leased to secondary users. This leads to advantages to both MNOs, since underutilized resources of a given network operator can be used by another operator in exchange for a profit.

The second analysis related to this case study shows the normalized frame capacity of LTE considering a variable amount of connections. In this scenario, the proportion of RTS, MS, and BES traffic follows a proportion of 20%, 20%, and 60%, respectively. These values were based on a recent report published by CISCO, which aims to forecast the network traffic between 2015 and 2020 (CISCO, 2016). Another important aspect of a cellular network that is considered in this second analysis refers to the propagation conditions of the radio frequency channel. These conditions are considered by analyzing the behavior of the LTE network in situations where the channels present different SINR values. The results presented in Figure 5.4 show that according to the number of connections, a proportion varying from 20 % to 85 % of the frame capacity is underutilized.

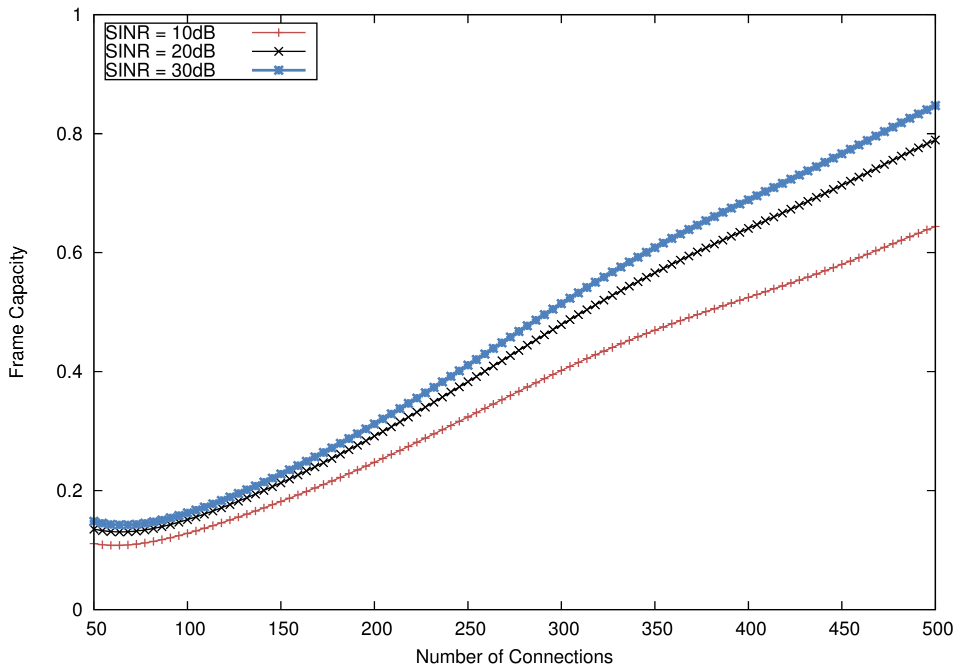


Figure 5.4: Frame capacity

This behavior may be observed in every scenario of traffic load. However, it may become even more aggressive in off pick hours, when the traffic load is naturally reduced. These underutilized resources can potentially be leased to secondary users which may belong to other LTE operators. The resources leasing may be implemented especially when the radio frequency channel conditions are favorable, for example with SINR varying from 20 to 30 dB, since in this case the amount of available resources can be increased by up to 20 %.

The performance of the proposed architecture is compared with two different resource allocation algorithms found in the Literature. Both algorithms were originally analyzed by Gardellin *et al.* (GARDELLIN; DAS; LENZINI, 2013) and are called Random Channel Allocation and Non-cooperative Channel Allocation. Both approaches consider that the shared resource is channel capacity and are applied to IEEE 802.22 networks but are general enough to be adapted to other network scenarios, such as the one under analysis.

Both algorithms used by Gardellin *et al.* allow the coexistence of multiple users given a set of radio frequency channels. Therefore, the coexistence is modeled as a channel assignment problem aiming at an efficient use of the available frequency spectrum. The first algorithm, called Random Channel Allocation is very simple. The algorithm is provided with knowledge about the number of channels that can be allocated and randomly selects one of these channels to allocate to the cellular network users. In the approach considered in this evaluation scenario, the Random Channel Allocation algorithm is adapted to the scenario where multiple LTE MNOs coexist in the same geographical area.

The second algorithm is part of the main proposal of Gardellin *et al.* (GARDELLIN; DAS; LENZINI, 2013) and is called Non-Cooperative Channel Allocation algorithm. In this algorithm, each network operator is responsible for selecting a channel for transmission. The nature of this algorithm may lead to the occurrence of collisions when two or more devices choose to transmit in the same wireless channel. To deal with collisions, a backoff mechanism is implemented considering a backoff window and a counter. The backoff window is defined as a value between the minimal viable SINR ($minSINR$) for transmission and this value added by the number of neighboring channels sensed by the network device (n_c). Therefore, the backoff window can be defined as $[minSINR, minSINR + n_c]$. The counter is randomly chosen within the backoff window range.

The first metric considered in this evaluation is the throughput. Figure 5.5, illustrates the throughput of the LTE network in conditions in which the number of connections is varied. In this case, results are obtained considering three wireless channels and four LTE operators in four different scenarios:

1. **Random channel allocation:** in this scenario, the Random Channel Allocation algorithm is applied to distribute the shared channels;
2. **Gardellin's proposal:** in this case, the Cooperative Channel Allocation algorithm is used for deciding channel allocations;
3. **Proposed solution with Random Channel Allocation:** simulation of the proposed ar-

chitecture. The channel allocation, in this case, is random;

4. **Proposed solution with backoff-based allocation:** the proposed architecture is complemented by a Non-Cooperative Channel Allocation Algorithm, based on the one proposed by Gardellin *et al.* (GARDELLIN; DAS; LENZINI, 2013).

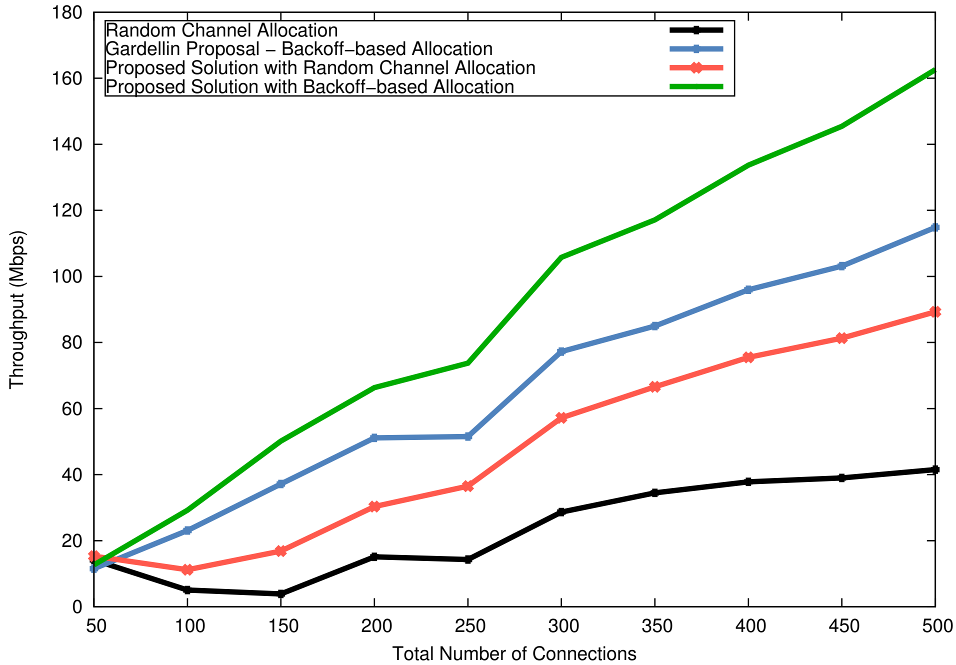


Figure 5.5: Throughput behavior with variable number of connections

As can be observed in Figure 5.5, the random channel allocation is the worst case scenario. This behavior is explained due to the nature of a random algorithm. Since no channel access control is implemented in this case, an uncontrolled number of collisions may occur, leading to a decrease in the overall throughput, what will impact negatively on the QoS. The throughput significantly improves when LTE resources sharing is implemented. In this case, it is important to highlight that although the channel allocation is random, when LTE resources are demanded, a SLA is established, what leads to more efficient resources allocation. A similar behavior is observed when the backoff-based cooperative channel allocation algorithm is analyzed. However, in this case, the nature of the algorithm reduces the number of collisions and consequently improves the network throughput. The gains provided by the proposed approach surpass those obtained by Gardellin *et al.* proposal by up to 28%. Therefore it shows to be the most effective solution to improve the throughput QoS metric.

The second QoS metric that is taken into account is the average delay, which is compared with related work in Figure 5.6. The average delay is affected by the number of active connections and consequently by the overall network demand. In this scenario, the delay of MS and RTS classes of services is measured considering a Weighted Fair Queuing (WFQ) scheduling algorithm. Results show that the implementation of the proposed architecture guarantees a reduction of more than 50% on the average delay in situations where the network is saturated.

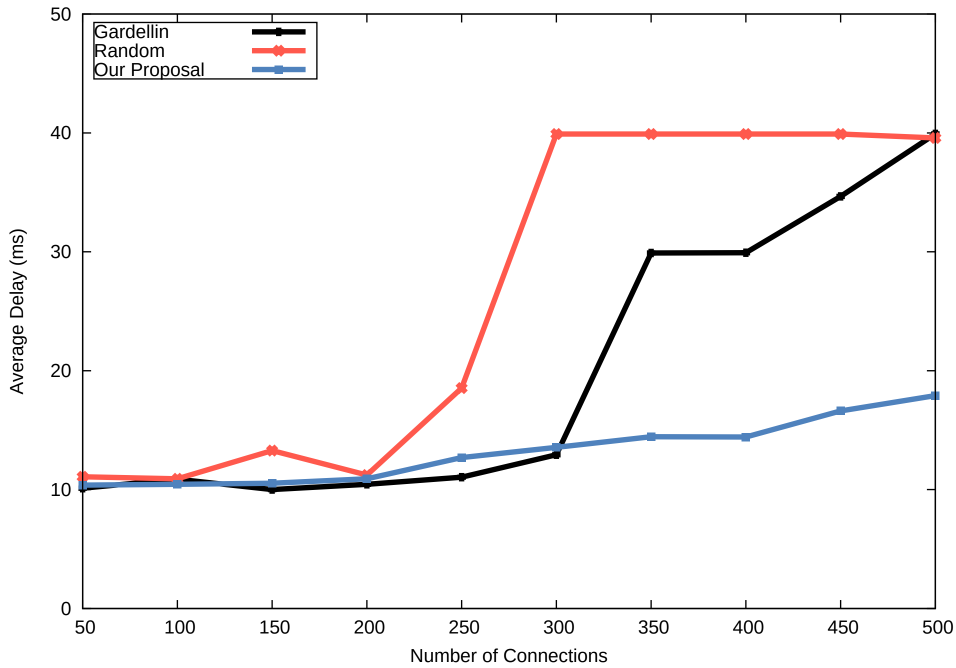


Figure 5.6: Delay Comparison

Another important aspect related to QoS provisioning is to analyze the amount of collisions, since this metric affects all the QoS metrics, especially the packet loss. In Figure 5.7, the percentage of collisions is related with the number of active connections. It is clear that the amount of collisions increases with the number of connections and stabilizes after a certain amount of connections, what is due to the characteristics of the Connection Admission Control (CAC) implemented by the MNOs to guarantee a certain QoS level. However, when the proposed approach is implemented, the amount of collisions, in the worst case is less than the half of the amount of collisions observed in the related approaches.

Another important aspect which is analyzed refers to the overhead generated by the transmission of control data. Two kinds of control data are considered in this analysis: (I) information exchanged by the polling/reply mechanism implemented in the update level of the resources broker and (II) information exchanged by the opportunistic users to request resources to the decision level of the broker. The first type of control data is analyzed in Figure 5.8, where the overhead is presented considering update intervals varying from 1 minute to 1 hour. This range of intervals allows the analysis of both very aggressive update strategies and of more conservative ones.

The amount of network operators that are sharing the same geographical area is also varied in the simulations. The overhead generated by the presence of only one operator, although not realistic, is evaluated as a baseline for comparison. The simulation scenarios implemented to evaluate the overhead of the architecture are designed to simulate a common situation where

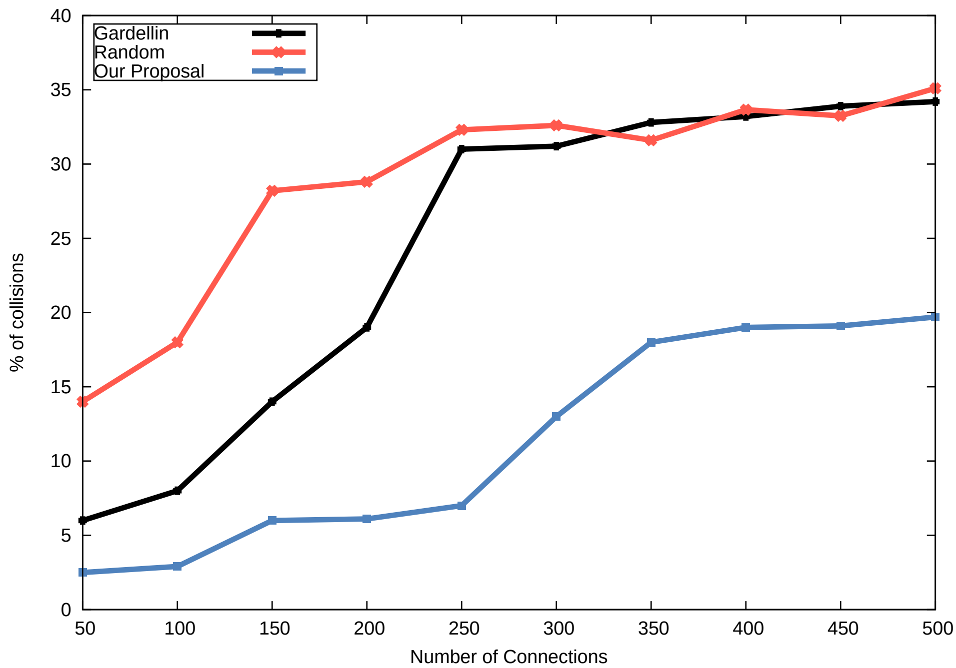


Figure 5.7: Percentage of collisions with variable number of connections

four LTE MNOs share the same geographical area, following the allocation model adopted by the Brazilian government (Anatel, 2015). The coexistence of six and ten operators were also simulated in order to evaluate the behavior of the architecture where the amount of operators is extrapolated, what may occur in the future.

The behavior of the update overhead follows the expectation that it may reduce as the interval between updates increase. Moreover, the amount of network operators sending updates directly affects the amount of control data generated during an update process. In terms of network load generated by the control data, it is possible to conclude that the update overhead can be neglected by the network operators. This is justified because, in a worst case scenario, where 10 network operators are updating the broker every minute, only about 0.12 Mbps of control traffic is generated. Considering that in good wireless channel conditions, each network operator may transmit approximately 120 Mbps, this value is very small, corresponding to 0.001% of the whole traffic.

Other kind of overhead generated by the proposed architecture is correlated with the request of resources, which is necessary every time a new resources sharing operation is about to begin. The behavior of this overhead is shown in Figure 5.9. In the graph, the percentage of the network capacity used for transmission of resources request is analyzed in relation to the amount of requests received per second.

Results show that even when a very large amount of request messages is exchanged, less then 5% of the MNO resources are compromised with control data transmission. In a realistic situation, the amount of requests per second should lead to an occupation of less then 1% of

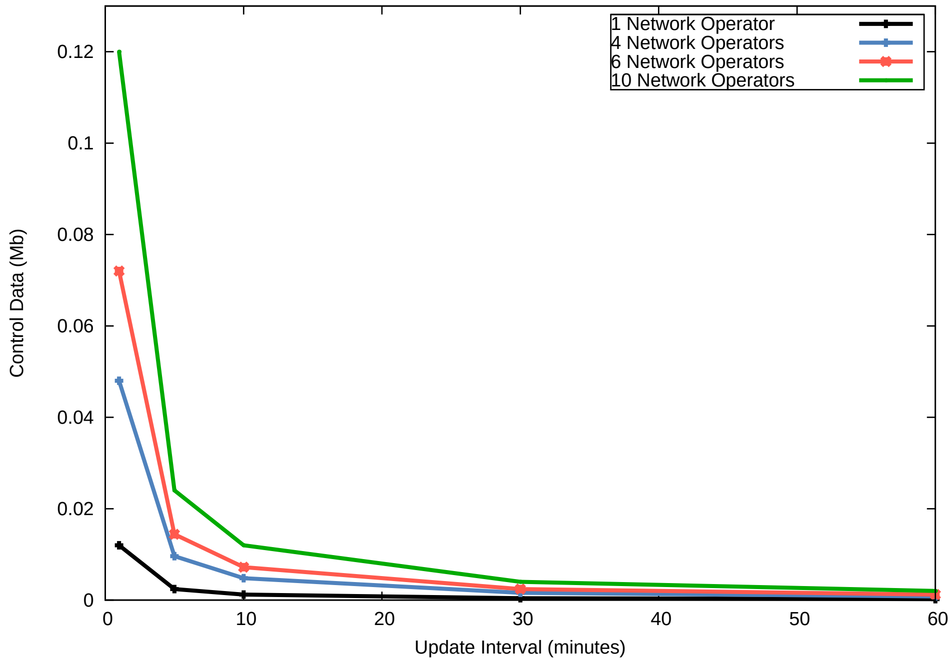


Figure 5.8: Update Overhead

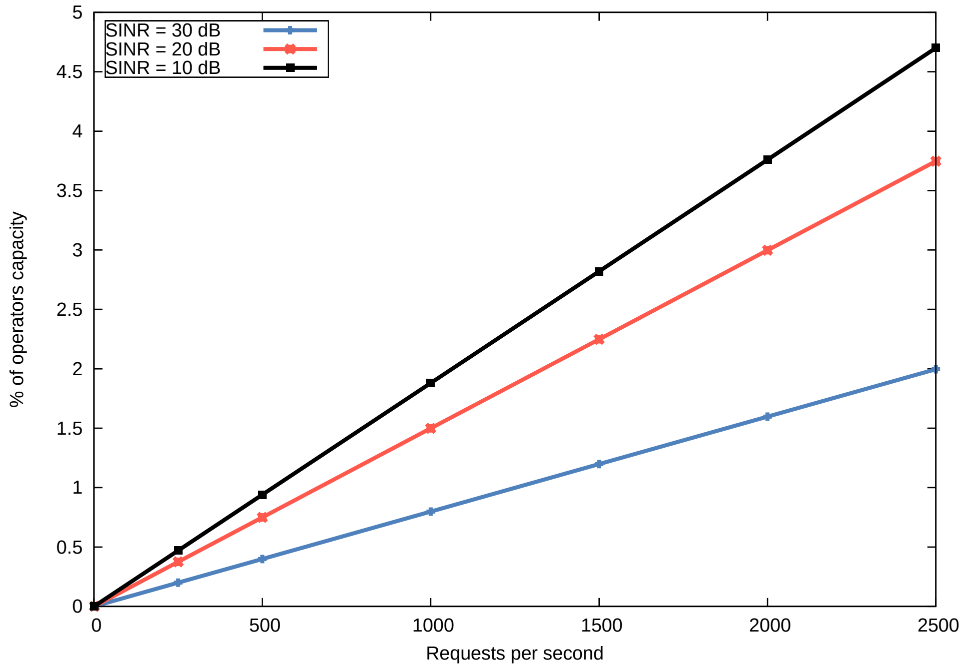


Figure 5.9: Request Overhead

the network resources. These results show that the request overhead can also be neglected by the network operators. Therefore, by analyzing the overhead generated by the proposed architecture, it is possible to conclude that the amount of control information that is exchanged among the broker and the network operators does not affect the overall performance of the network.

5.2.2 Fast Handover in heterogeneous network scenarios

The main goal of this study case scenario is to analyze the proposed architecture when a cognitive mechanism is implemented to perform in advance decisions to allow traffic steering in unscheduled evacuation of LSA bands. This scenario is justified by the fact that using LSA bands is one of the emerging solutions to deal with the resources scarcity problem. In this case, spectrum sharing is allowed by the spectrum rights holder (*i.e.*, the incumbent user) which temporarily provides access to the LSA licensees (PONOMARENKO-TIMOFEEV A. PYATTAEV; KARLS, 2016). However, the incumbent user is eligible to dynamically request the resources back at any time. Such request compels the LSA licensees to promptly evacuate the spectrum to avoid interfering with the incumbent services. In order to vacate the resources in a timely manner, LSA licensees must implement fast handover strategies and consequently manage to steer the traffic of evacuees to available portions of the spectrum.

In this specific case, the resources sharing topology is composed of one LTE network operator accessing LSA frequencies, which coexists, in the same geographical area, with three IEEE 802.11 networks operating in no interfering channels (1, 6, and 11). The IEEE 802.11 networks play the role of possible evacuation routes for the LTE clients, whenever the incumbent user requests its resources back. The broker of the proposed architecture is therefore responsible for coordinating resources sharing. Besides all simulation parameters already explained, in this case, realistic traces were also used to model the behavior of the IEEE 802.11 networks. The traces were obtained from CRAWDAD database (SCHULMAN; SPRING, 2009).

In this case study scenario, the resources broker is demanded to take fast decisions when an evacuation is required. The evacuation request is composed of a struct which informs the CoS and the QoS requirements of the client. The Decision Level is constantly running, with the goal of taking in advance decisions regarding the traffic steering, which is used to promptly vacate the LSA band when required. This in advance decision demands this level to forecast the traffic of the LTE-LSA and Wi-Fi operators in order to identify viable evacuation route. Such forecast requires knowledge about the historical traffic load of LTE and Wi-Fi networks, which is stored in the Resources Level. Later, the historical traffic load is processed and normalized by the Decision Controller in the Decision Level. The resulting values serve as inputs to the Provisioning Algorithm.

With regard to the performance of the proposed solution under this specific scenario, the first factor to analyze is the accuracy of the traffic load forecasting model. The forecasting follows three key phases. The first is the time series extraction of traffic data from LTE-LSA and Wi-Fi networks. The second consists of fitting the polynomial curve of traffic data of both LTE-LSA and Wi-Fi networks. In the third phase, the forecasting is carried out by means of the MLRM as detailed in equation 4.7.

Considering that the time series is a sequence of data points, that generally consists of successive measurements made in a time interval (HAMILTON, 1943), these data points are divided into three data sets: training, validation, and testing. The training data set contains the traffic load measurement that corresponds to the first 15 minutes of the time series. The validation data set consists of 10 percent of the testing data set which is used to analyze the outcomes of the prediction, by taking account of metrics such as accuracy and processing time.

The MLRM processes the trained data set of simulated traffic demands for the LTE-LSA network, as well as the aggregate traffic of the Wi-Fi networks. The simulated traffic in the LTE-LSA network and the traffic traces of Wi-Fi are computed in units of seconds to improve the accuracy of the model. The first step of the analytical methodology involves calculating the polynomial curve fitting for smoothing out the peaks and noise of the network traffic. The polynomial was fixed at 10 degrees for curve fitting analysis of traffic of each network. The classification is then performed again and includes the new data points obtained from the ten degrees polynomial for the training, validation, and testing datasets. After this, the MLRM carries out the traffic load forecasting and the validation data set is used to evaluate its accuracy for each network.

The MLRM accuracy is evaluated by the cross-validation method which involves the comparing the forecast values with the current values. At this point, the MLR model can be adjusted to improve the accuracy of the upcoming predictions. Equation (4.9) is also used to calculate the accuracy obtained after running the MRLM. Figure 5.10 shows the analysis of the accuracy of traffic load forecasting, which examines three Wi-Fi networks as possible traffic steering routes. As can be inferred analyzing the graph, the traffic load forecasting was very accurate, reaching levels of 96.18%, 93.61%, and 94,20% of accuracy, considering the traffics of Wi-Fi networks 1 (operating in channel 1), 2 (operating in channel 6), and 3 (operating in channel 11), respectively.

Every time the traffic load forecasting is performed, the values of the time series data points prediction are updated and input into the Provisioning Algorithm, which is responsible for selecting the traffic steering routes. The first step taken by the provisioning algorithm is to estimate the availability and occupation of bandwidth for each target network taking into account the previous forecasting. The second step involves selecting the Wi-Fi networks which can guarantee the same level of QoS as that offered in LTE-LSA network. This kind of decision is taken based on the predicted availability of network resources, considering the traffic load, the QoS requirements and the costs involved in the resources sharing initiative. The traffic load forecasting starts from the 15 minutes in Figures 5.10 and 5.11 because it requires the historical traffic load measurements of the last 15 minutes to train the MLRM and predict the next 15 minutes of traffic load trend with an accuracy close to 95% and to guarantee a fast response.

The provisioning algorithm conducts the analysis of future network capacity of each overlapping Wi-Fi network by relying on the trapezoidal numerical integration to calculate the area

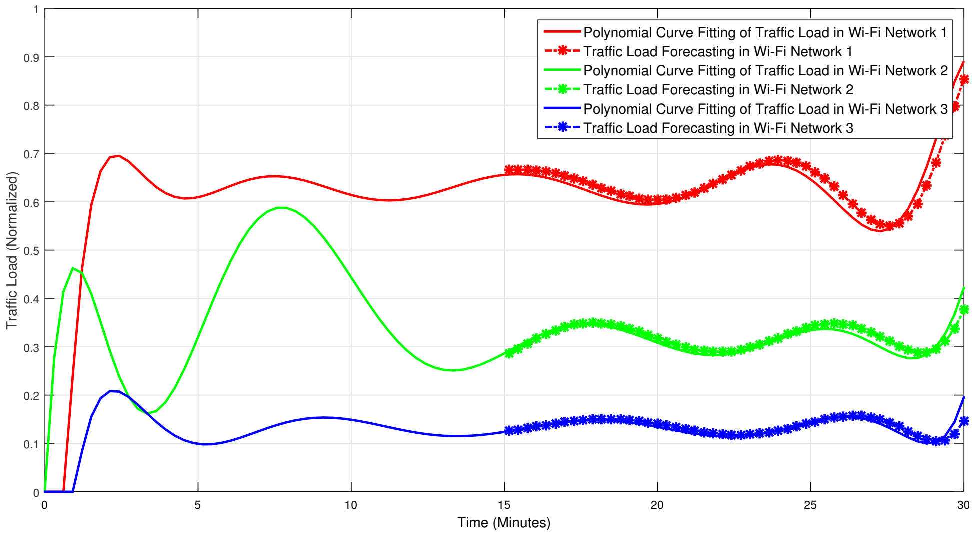


Figure 5.10: Actual vs. Predicted Traffic Load for each Wi-Fi Network

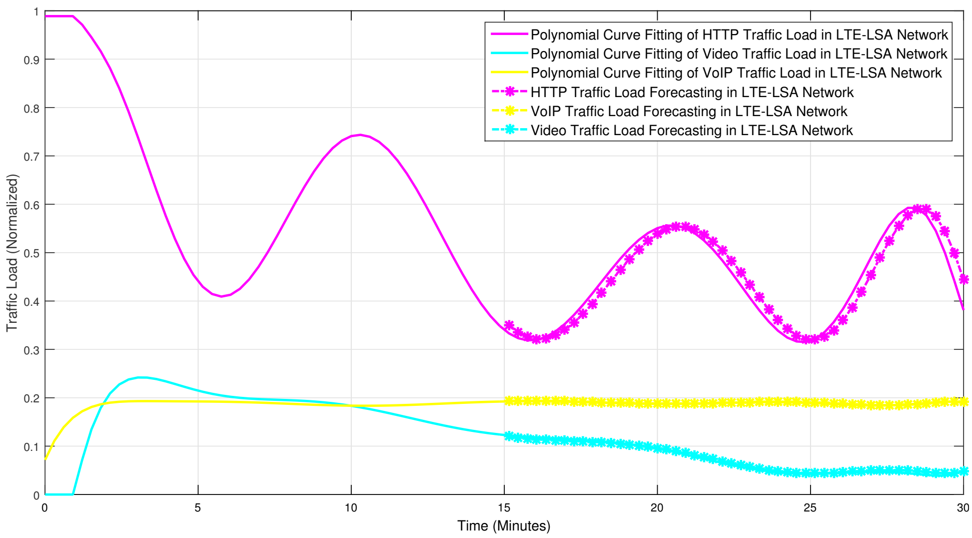


Figure 5.11: Actual vs. Predicted Traffic Load for each class of service in LTE-LSA Network

under the curve of the MLRM forecast. The area under the curve is equivalent to the percentage of occupied resources for each Wi-Fi network. Analyzing Figure 5.10, it is possible to observe that the percentage of forecast occupied bandwidth 65.8%, 31.6%, and 15.4%, for the Wi-Fi networks 1, 2, and 3, respectively. Based on these values, the in advance decision algorithm determines Wi-Fi 1 as a low priority route for traffic steering because of its considerably high traffic load. On the other hand, the provisioning algorithm defines Wi-Fi 2 and 3, as high priority traffic steering routes. After this initial analysis, when an evacuation is required, the provisioning algorithm associates the CoSs with the previous information to perform the traffic offloading while taking account of the QoS requirements of the evacuees.

The bandwidth occupation in Wi-Fi 1 oscillates close to 95% with its original users, making this network unavailable. Figures 5.12 and 5.13 show the occupation of Wi-Fi 2 and 3 after the traffic steering. The bandwidth occupation in Wi-Fi 2 fluctuates between 40% and 70% after the offloading of Video and VoIP traffic demands from the LTE-LSA network, while Wi-Fi 3 network bandwidth occupation is around 80%. These results show that all the traffic was accommodated in the destination networks without overloading them. Thus, the QoS of the evacuees can be guaranteed without interfering with the original Wi-Fi users in terms of network capacity.

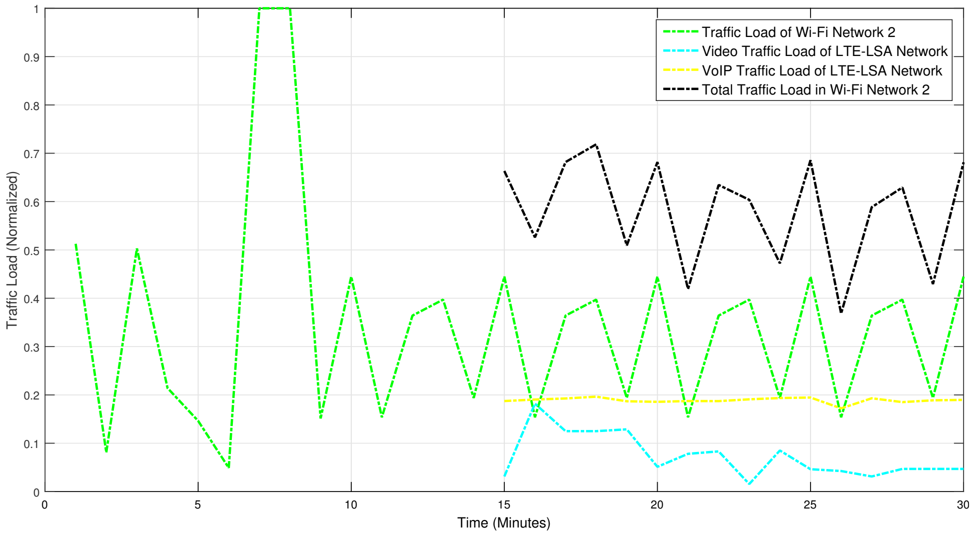


Figure 5.12: Video and VoIP Traffic Steering from LTE-LSA to Wi-Fi 2

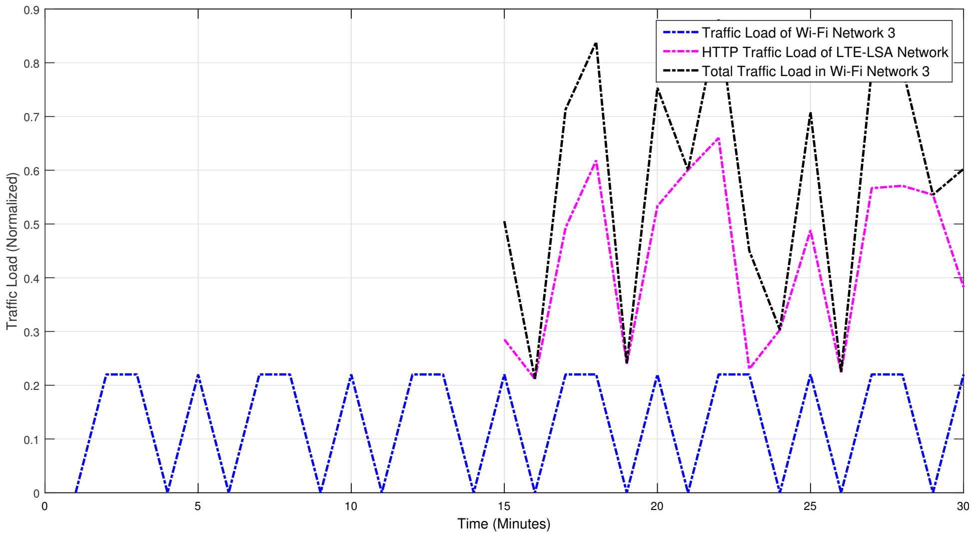


Figure 5.13: HTTP Traffic Steering from LTE-LSA to Wi-Fi 3

Another important QoS metric is the delay. Figure 5.14 shows the behavior of this metric considering a variable amount of connections accommodated by each Wi-Fi network. As can be seen in the graph, Wi-Fi 1 has the smallest delay value because it is a low-priority traffic steering route and thus the provisioning algorithm does not make it eligible to receive traffic from delay-sensitive applications. Wi-Fi networks 2 and 3, on the other hand, receive QoS sensitive traffic and are capable of keeping the average delay below 30ms. This value is sufficient to guarantee the QoS of multimedia traffic, which generally requires the delay to be between 100 and 200 ms (CISCO, 2016).

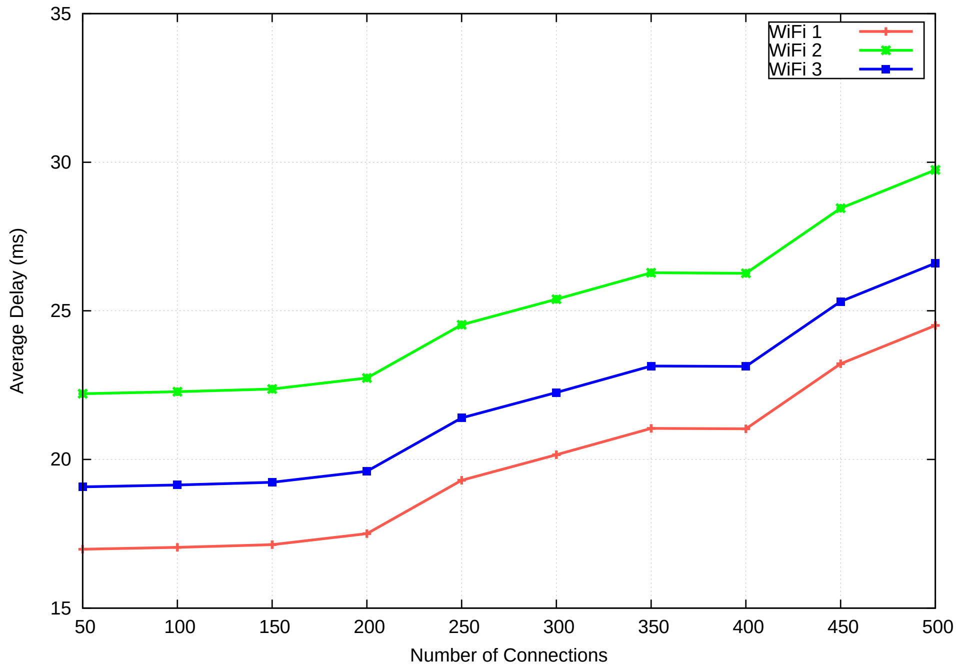


Figure 5.14: Average delay for differente Wi-Fi networks

Another crucial factor that must be covered by the provisioning algorithm is to avoid interfering with the incumbent services in the event of an unscheduled evacuation. For this reason, the traffic steering must occur as fast as possible. The outcomes of this approach are similar to those of related work in the literature. Mustonen et al. (MUSTONEN et al., 2014) was able to perform the decision in approximately 0.9 seconds on average, while Palola et al. (PALOLA et al., 2014) designed an algorithm which was able to carry out the decision in 0.624 seconds. The cognitive mechanism proposed in this thesis, which is based on accurate forecasts, reduces the average decision time to values as low as 0.0371 seconds.

The processes related to the overall time required by the proposed solution to evacuate the LSA band and hence to offload the traffic to the selected Wi-Fi network, are outlined in Table 5.5. Since the proposed approach involves taking in advance decisions, the duration of both the decision process and the overall evacuation can be reduced. A CEPT Report 159 (CEPT, 2011) stated that the duration for turning off an LTE BS with one sector must take at most 20.620. This constraint imposes the need of reducing the evacuation duration to avoid interfering with the incumbent services in the LSA frequency and ensure the QoS of the evacuees. The results of the simulations show that the proposed solution allows the overall evacuation to be conducted in about 11.3 seconds, which represents a value that is around 46% below the specified limit.

Table 5.5: Evacuation Duration

Process	Average Duration [s]	Standard Deviation [s]
Traffic Load Forecasting	3.8267	0.2161
Cognitive Decision	0.0371	0.0051
Traffic Steering	7.3962	0.9477
Total Duration	11.2698	3.0163

5.2.3 Video Surveillance in Smart Cities

The third case study considers the context of Internet of Things (IoT), which aims to connect objects to enable the exchange of information (ZOU et al., 2017). IoT has become particularly popular with the rapid development of small low-cost sensors, wireless communication technologies, and new Internet techniques. Typical applications of IoT include surveillance, health care, transportation, start cities infrastructure, and sensors technology. Among these applications, a very important one is surveillance, both for smart cities and country borders.

Considering this context, this case study aims at analyzing the performance of the proposed architecture in a scenario where multiple video cameras are used for the surveillance of borders in a military scenario. Figure 5.15 shows the structure of the IoT system and how the proposed architecture fits into the existing IoT architecture.

The IoT features can be divided in four layers (CHEN et al., 2017). The sensors layer is located at the bottom of the architecture and is composed of the IoT sensors used for gathering data. In the specific case of the analyzed scenario, both fixed video cameras and smartphones are considered to coexist in order to generate videos which must be processed later by the IoT architecture. The second layer is called Networking and represents all the network operators available for data transmission coexist in the same geographical area. Since the analyzed scenario is military, all kinds of resources defined in the Resources Level of the proposed Broker are present. For the sake of simulation, 4G and 5G LTE operators are considered primary users, IEEE 802.11 and IEEE 802.22 are considered secondary users, and LSA military frequencies are considered as the LSA Pool. The third layer of the architecture, presented in Figure 5.15 represents the IoT Platform, which typically implements the concepts of big data in order to process the information gathered by the sensors. In this platform, the proposed resources broker is included to manage the resources sharing among the heterogeneous networks of the Networking Layer. In the top layer of the model, Surveillance is considered as the IoT application for this scenario.

In this case study, the Resources Broker acts to balance the load among the available MNOs. Therefore, it aims to improve the amount of resources available for the transmission of the video captured by the cameras interconnected along a country border. Different situations are consid-

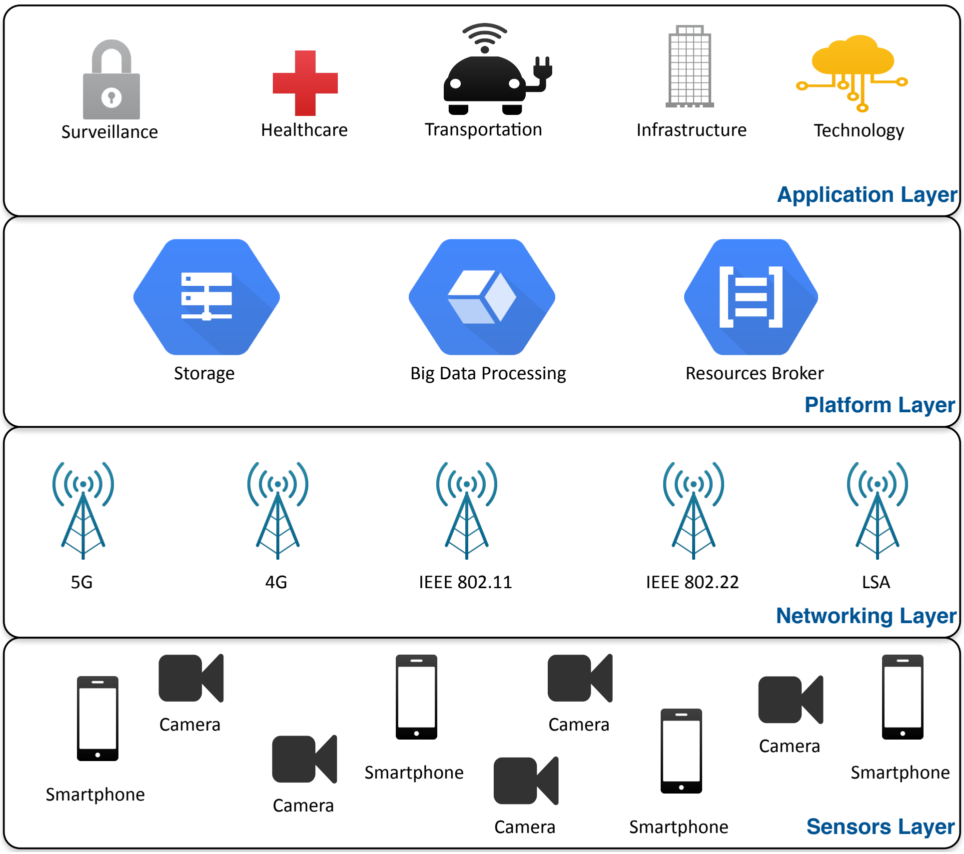


Figure 5.15: Video Surveillance IoT Architecture

ered regarding the amount of sensors (cameras and smartphones) used for video surveillance, since a variation between 50 and 500 sensors is considered to analyze the behavior of the proposed solution in both low and high traffic loads. The first performance aspect analyzed is the traffic load distributed to each MNO. Figure 5.16 illustrates the outcomes of the simulations for this scenario.

The graph presented in 5.16 considers only the traffic belonging to the sensors installed along the country border, without considering the MNO clients traffic. This limited scenario in terms of MNO traffic was considered because the aim is to analyze the capability of the resources broker to properly distribute the traffic demand among the MNOs, considering the rules imposed by Algorithms 1 and 2, which implement the resources controller and the resources provisioning mechanisms, respectively. Therefore, the presence of the original MNO traffic would not allow a precise analysis of the sensors traffic balancing.

Analyzing the graph, it is possible to realize that the majority of the traffic was steered to Wi-Fi and IEEE 802.22 networks. This occurs because one of the rules of the resources controller algorithm is to prioritize cheap network resources. Since these networks belong to shared use regime, when Equation (4.5) is applied, it leads to a better cost-benefit, considering that the target MNO is able to guarantee the QoS requirements of the resources renter. Therefore, nearly 70% of the traffic generated by the IoT sensors is directed to these networks. The remaining

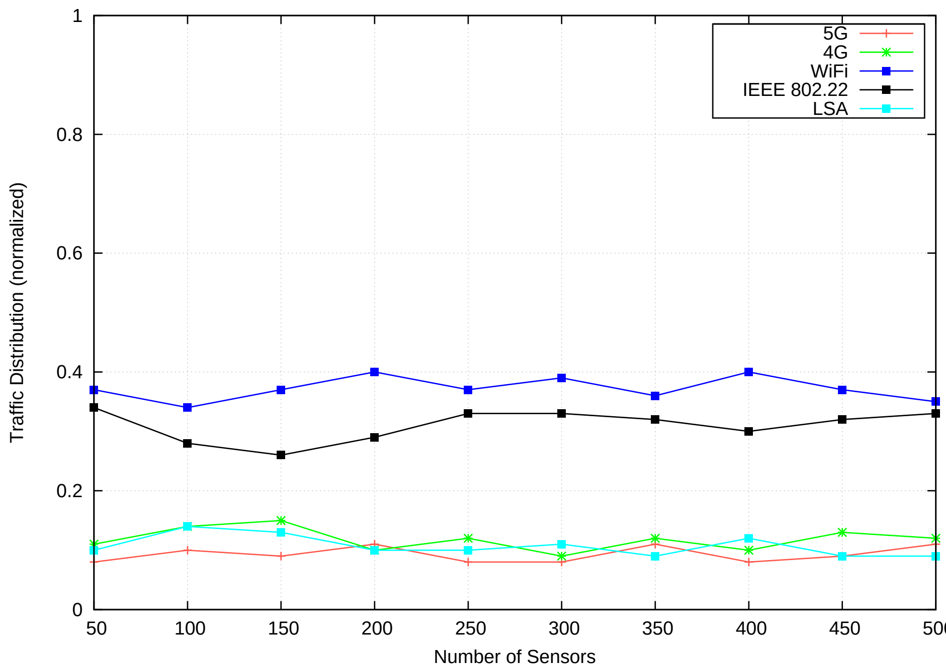


Figure 5.16: Traffic load in each MNO - IoT scenario

traffic is accommodated by more expensive resource providers, *e.g.* LSA, 4G, and 5G. This occurs because the transmission of Real Time Video is classified into the RTS CoS and therefore, the requests belonging to this CoS are eligible to access more expensive resources in cases that the cheaper resources are not enough to guarantee QoS.

Another important factor to be analyzed in this IoT scenario is whether or not the QoS is guaranteed by the proposed resources broker. Two QoS metrics are considering the transmission of RTS video traffic. The first metric is the average delay measured in the destination MNOs. The results regarding the delay are shown in Figure 5.17, where this metric is analyzed considering a variable number of sensors and a delay limit to guarantee QoS of 150ms (CISCO, 2016).

It is important to emphasize that the QoS is guaranteed in terms of average delay, since in no case this value overpasses the delay limit for this kind of application. In the worst case observed in the graph, *i.e.* 4G LTE MNO in a situation where 500 sensors coexist, the average delay is around 140ms. Analyzing situations with lower network traffic, it is possible to observe that all MNOs are able to guarantee similar delays for the video transmission, with values between 100ms and 120ms, what is acceptable in terms of QoS. This controlled average delay values result from the execution of both resources controller and resources provisioning algorithms, which gather historical and current information about the MNO average delay prior to taking an allocation decision.

The second important QoS metric for the RTS CoS is the jitter. The values for this metric were also measured in the resource providers considering the coexistence of up to 500 sensors. In this specific case, the threshold value to guarantee QoS is 50ms, according to the definitions

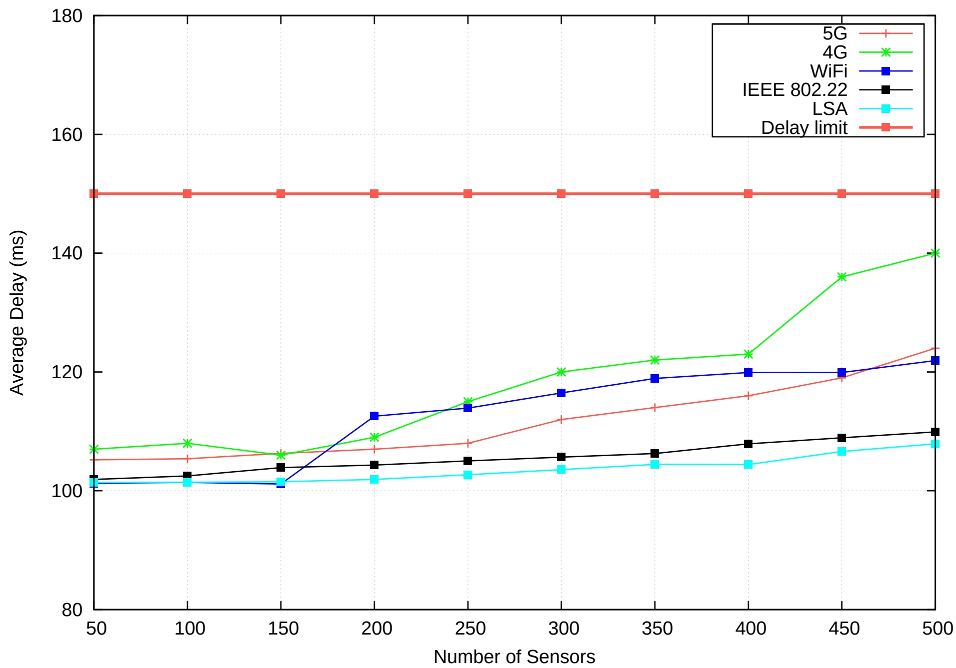


Figure 5.17: Average Delay in each MNO - IoT scenario

of commercial equipment manufacturers (CISCO, 2016). The results regarding the jitter metric are presented in Figure 5.18.

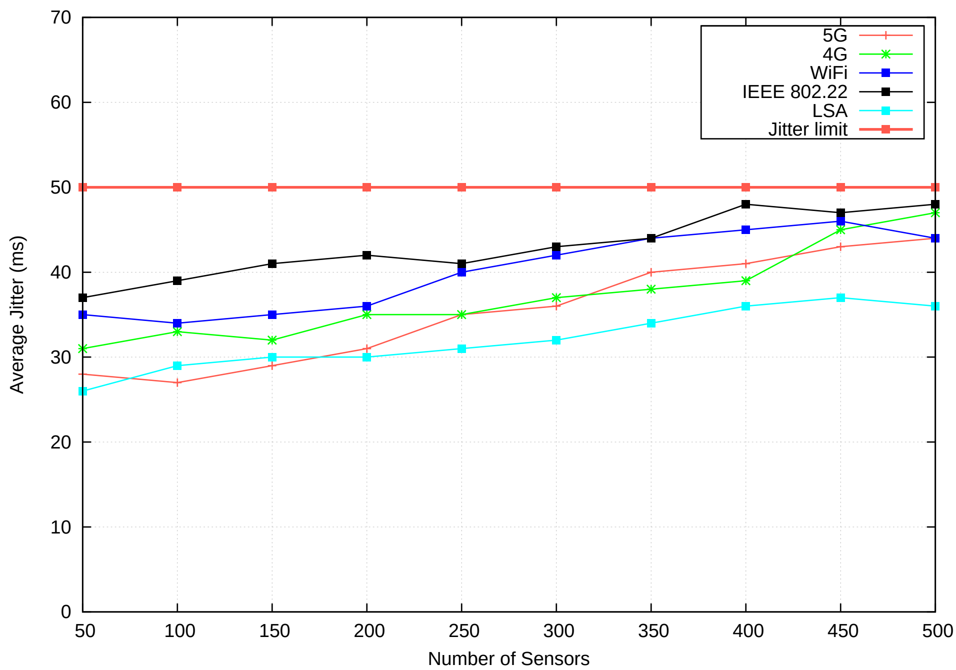


Figure 5.18: Average Jitter in each MNO - IoT scenario

In this case, the obtained results present a behavior which is similar to that observed when conducting the analysis of the delay metric. The jitter values are maintained between 25ms and

47ms in all MNOs. At no point of the graph, the jitter is over the threshold of 50ms. Therefore, it is possible to conclude that, considering the evaluated scenario, the resources broker was able to guarantee QoS considering both delay and jitter metrics.

6 CONCLUSION AND FUTURE WORK

This thesis dealt with the spectrum scarcity problem by implementing QoS-aware resources sharing in both homogeneous and heterogeneous network scenarios. A simulation tool was designed to obtain results which confirmed the hypothesis that it is possible to implement a resources sharing mechanism to tackle the resources scarcity problem and guarantee the QoS of network clients. Three simulations scenarios have been defined to evaluate the performance of the proposed resources sharing architecture, aiming at answering the research questions. The results obtained via simulation show that the overhead was kept sufficiently low to avoid interfering with the network traffic. Furthermore, it was possible to implement a fast handover mechanism, which conducted the handover process 46% faster than the maximum time allowed to avoid interfering with the incumbent user. Further results showed that the QoS is guaranteed for the resources renter, considering the coexistence of up to 500 sensors in an IoT-based video surveillance scenario. The summary of the contributions and conclusions obtained in this thesis are presented and discussed in Section 6.1. Final remarks and direction for future work based on issues not completely tackled in this thesis are presented in Section 6.2.

6.1 Summary of Contributions

In this work, an architecture was designed and implemented to allow resources sharing involving MNOs which coexist in the same geographical area. The proposed solution allows the sharing of different kinds of network resources, for example, spectrum and capacity, in both homogeneous and heterogeneous network scenarios. The most important component of the proposed solution is a resources broker which takes into account the QoS requirements and the cost of the resources in order to take decisions regarding resources sharing.

The proposed architecture contributes with the state of the art on resources sharing since the comparison with two related works shows that in terms of throughput, the gains provided by the proposed approach surpass those obtained by Gardellin et al. proposal by up to 28% and up to three times if compared with the random allocation approach. In terms of delay metric, the implementation of the proposed architecture guarantees a reduction of more than 50% on the average delay in situations where the network is saturated.

Another important contribution of the proposed architecture is the efficient mechanism implemented to update the information needed to manage the resources.The update overhead does not surpass 0.001% of the whole network traffic. Other kind of overhead analyzed considered the amount of control information exchanged to request resources. In this case, even when a very large amount of request messages is exchanged, less then 5% of the operator's resources are compromised with control data transmission.

A fast evacuation solution was also proposed to take time sensitive resources allocation decisions without interfering with the incumbent users traffic in heterogeneous network scenarios.

This solution is based on a very accurate traffic forecasting MLRM, which reached a level of 96.18% of correction in the traffic forecasting. Moreover, the results of the simulation show that the proposed solution allows the overall evacuation to be conducted in about 11.3 seconds, which represents a value that is around 46% below the specified time limit for this kind of handover.

Finally, the behavior of the proposed architecture in a video surveillance scenario was analyzed due to the strict QoS requirements imposed by this kind of application. In this case, the cost-benefit is taken into account by selecting the lowest cost available among the operator which claim to fulfill the QoS requirements. Nearly 70% of the traffic generated in this scenario was steered to WiFi and IEEE 802.22 networks, which belong to shared use regime and therefore provide cheaper resources. Moreover, both delay and jitter QoS metrics were kept under the specified limit and therefore, QoS was guaranteed.

6.2 Final Remarks and Suggestions for Future Work

Future investigations should conduct an in-depth analysis of the performance of the proposed solution. This analysis could include the execution of the proposed resources sharing architecture in realistic testbeds. A few existing testbeds are compatible with the proposed architecture. The most important one is the Cognitive Radio Trial Environment (CORE) from VTT Technical Research Center of Finland, since it is suitable to implement the coexistence of technologies in LSA scenarios.

Moreover, other kinds of provisioning algorithms can be implemented, both via simulation and testbeds, and evaluated to deal with specific network scenarios. Another possible future work is to improve the favor balance between MNOs. Although the solution presented in this thesis considers the cost of the resources in the decision process, a simple algorithm was implemented without considering the favors balance. Improvements could be focused on the definition of a threshold related to the maximum cost that may be leased before the favor is paid back.

Finally, another relevant topic which can be addressed in future works, is the proposal of a communication protocol to replace the one used in this thesis, which is based on LTE-Advanced standardization. A specific protocol may reduce even more the update overhead of the proposed solution.

APPENDIX A ORIGINAL SLR PAPERS SET

	Title	Publication	Year	Q1	Q2	Q3	Grade
1	Toward spectrum sharing: opportunities and technical enablers	IEEE Communications Magazine	2016	1	1	1	3
2	Coordination protocol for inter-operator spectrum sharing in co-primary5G small cell networks	IEEE Communications Magazine	2016	1	1	1	3
3	Synergistic spectrum sharing in 5G HetNets: A harmonized SDN-enabled approach	IEEE Communications Magazine	2016	1	1	1	3
4	Beyond Coexistence: Traffic Steering in LTE Networks with Unlicensed Bands	IEEE Wireless Communications	2016	1	1	1	3
5	Spectrum access system for the citizen broadband radio service	IEEE Communications Magazine	2015	1	1	1	3
6	Self-coexistence in cellular cognitive radio networks based on the IEEE802.22 standard	IEEE Communications Magazine	2013	1	1	1	3
7	Radio access network virtualization for future mobile carrier networks	IEEE Communications Magazine	2013	1	1	1	3
8	Secondary User Access in LTE Architecture Based on a Base-Station-Centric Framework With Dynamic Pricing	IEEE Transactions on Wireless Communications	2013	1	1	1	3
9	Wireless resource sharing for multiple operators: Generalization, fairness, and the value of prediction	Computer Networks	2016	1	1	0.5	2.5
10	Advanced spectrum sharing in 5G cognitive heterogeneous networks	IEEE Wireless Communications	2016	1	1	0.5	2.5

11	CPC-based backward-compatible network access for LTE cognitive radio cellular networks	IEEE Communications Magazine	2015	1	1	0.5	2.5
12	Cooperative Spectrum Sharing Between Cellular and Ad-Hoc Networks	IEEE Transactions on Wireless Communications	2014	1	1	0.5	2.5
13	Coexistence Decision Making for Spectrum Sharing Among Heterogeneous Wireless Systems	IEEE Transactions on Wireless Communications	2014	1	1	0.5	2.5
14	Proportional Fair Traffic Splitting and Aggregation in Heterogeneous Wireless Networks	IEEE Communications Letters	2016	0.5	1	0.5	2
15	Optimized Resource Management in Heterogeneous Wireless Networks	IEEE Communications Letters	2016	0.5	1	0.5	2
16	Coordinated Resource Partitioning and Data Offloading in Wireless Heterogeneous Networks	IEEE Communications Letters	2016	0.5	1	0.5	2
17	Public safety networks evolution toward broadband: sharing infrastructures and spectrum with commercial systems	IEEE Communications Magazine	2016	1	0	1	2
18	Cellular Meets WiFi: Traffic Offloading or Resource Sharing?	IEEE Transactions on Wireless Communications	2016	1	1	0	2
19	QoS-Aware Tethering in a Heterogeneous Wireless Network using LTE and TV White Spaces	Computer Networks	2015	0.5	1	0.5	2
20	Coexistence of Wi-Fi and heterogeneous small cell networks sharing unlicensed spectrum	IEEE Communications Magazine	2015	1	1	0	2

21	Service provider competition and cooperation in cloud-based software defined wireless networks	IEEE Communications Magazine	2015	1	0	1	2
22	Resource Allocation for Heterogeneous Applications With Device-to-Device Communication Underlaying Cellular Networks	IEEE Journal of Selected Areas in Communications	2015	0.5	1	0.5	2
23	Quantifying Benefits in a Business Portfolio for Multi-Operator Spectrum Sharing	IEEE Transactions on Wireless Communications	2015	1	1	0	2
24	Resource sharing in heterogeneous cloud radio access networks	IEEE Wireless Communications	2015	1	1	0	2
25	Load Distribution in Heterogeneous Cellular Networks	IEEE Communications Letters	2014	1	1	0	2
26	Software-defined-networking-enabled capacity sharing in user-centric networks	IEEE Communications Magazine	2014	1	1	0	2
27	Coopetition Methodology for Resource Sharing in Distributed OFDM-Based Cognitive Radio Networks	IEEE Transactions on Communications	2014	1	0.5	0.5	2
28	Dynamic Cooperative Secondary Access in Hierarchical Spectrum Sharing Networks	IEEE Transactions on Wireless Communications	2014	1	0.5	0.5	2
29	A Cooperative Transmission Scheme for Improving the Secondary Access in Cognitive Radio Networks	IEEE Transactions on Wireless Communications	2014	1	0.5	0.5	2

30	Spectrum Allocation and QoS Provisioning Framework for Cognitive Radio With Heterogeneous Service Classes	IEEE Transactions on Wireless Communications	2014	1	0	1	2
31	An Asynchronous Fixed-Point Algorithm for Resource Sharing With Coupled Objectives	IEEE/ACM Transactions on Networking	2014	1	0	1	2
32	Energy-Efficient Cell Activation, User Association, and Spectrum Allocation in Heterogeneous Networks	IEEE Journal of Selected Areas in Communications	2016	0	1	0.5	1.5
33	Energy Efficient Resource Allocation for Mixed RF/VLC Heterogeneous Wireless Networks	IEEE Journal of Selected Areas in Communications	2016	0	1	0.5	1.5
34	Relay-Based Spectrum Sharing With Secondary Users Powered by Wireless Energy Harvesting	IEEE Transactions on Communications	2016	1	0	0.5	1.5
35	Resource Partitioning and User Association With Sleep-Mode Base Stations in Heterogeneous Cellular Networks	IEEE Transactions on Wireless Communications	2015	0	1	0.5	1.5
36	An energy-efficient QoS-based network selection scheme over heterogeneous WLAN – 3G networks	Computer Networks	2014	0	1	0.5	1.5
37	Aspiration Level-Based Strategy Dynamics on the Coexistence of Spectrum Cooperation and Leasing	IEEE Communications Letters	2014	1	0	0.5	1.5
38	On the coexistence of IEEE 802.11ac and WAVE in the 5.9 GHz Band	IEEE Communications Magazine	2014	0.5	1	0	1.5

39	ITLinQ: A New Approach for Spectrum Sharing in Device-to-Device Communication Systems	IEEE Journal of Selected Areas in Communications	2014	1	0	0.5	1.5
40	Quality of Service Games for Spectrum Sharing	IEEE Journal of Selected Areas in Communications	2014	1	0	0.5	1.5
41	Resource Allocation in Spectrum-Sharing OFDMA Femtocells With Heterogeneous Services	IEEE Transactions on Communications	2014	1	0	0.5	1.5
42	Opportunistic Spectrum Sharing With Limited Feedback in Poisson Cognitive Radio Networks	IEEE Transactions on Wireless Communications	2014	1	0	0.5	1.5
43	Two Dimension Spectrum Allocation for Cognitive Radio Networks	IEEE Transactions on Wireless Communications	2014	0.5	0.5	0.5	1.5
44	Spectrum Sharing in MIMO Cognitive Radio Networks Based on Cooperative Game Theory	IEEE Transactions on Wireless Communications	2014	1	0	0.5	1.5
45	Improving Spectrum Efficiency via In-Network Computations in Cognitive Radio Sensor Networks	IEEE Transactions on Wireless Communications	2014	1	0	0.5	1.5
46	Cognitive radio resource management for future cellular networks	IEEE Wireless Communications	2014	1	0	0.5	1.5
47	Mobile traffic offloading by exploiting social network services and leveraging opportunistic device-to-device sharing	IEEE Wireless Communications	2014	1	0	0.5	1.5

48	Distributed joint resource and power allocation in self-organized femtocell networks: A potential game approach	Journal of Network and Computer Applications	2014	1	0	0.5	1.5
49	Cross-layer QoE-driven admission control and resource allocation for adaptive multimedia services in LTE	Journal of Network and Computer Applications	2014	1	0	0.5	1.5
50	Next generation public safety networks: A spectrum sharing approach	IEEE Communications Magazine	2016	1	0	0	1
51	Pervasive spectrum sharing for public safety communications	IEEE Communications Magazine	2016	1	0	0	1
52	Global Optimal Rate Control and Scheduling for Spectrum-Sharing Multi-Hop Networks	IEEE Transactions on Wireless Communications	2016	1	0	0	1
53	Proactive Spectrum Sharing Incentive for Physical Layer Security Enhancement Using Outdated CSI	IEEE Transactions on Wireless Communications	2016	1	0	0	1
54	Coupled Detection and Estimation Based Censored Spectrum Sharing in Cognitive Radio Networks	IEEE Transactions on Wireless Communications	2016	1	0	0	1
55	A Dynamic Resource Sharing Mechanism for Cloud Radio Access Networks	IEEE Transactions on Wireless Communications	2016	1	0	0	1
56	Multichannel Analysis of Cell Range Expansion and Resource Partitioning in Two-Tier Heterogeneous Cellular Networks	IEEE Transactions on Wireless Communications	2016	0	1	0	1

57	On the limits of predictability in real-world radio spectrum state dynamics: from entropy theory to 5G spectrum sharing	IEEE Communications Magazine	2015	1	0	0	1
58	Spectrum and license flexibility for 5G networks	IEEE Communications Magazine	2015	1	0	0	1
59	Joint Spectrum Sharing and Power Allocation for OFDM-Based Two-Way Relaying	IEEE Transactions on Wireless Communications	2015	1	0	0	1
60	Network Layer Scheduling and Relaying in Cooperative Spectrum Sharing Networks	IEEE Transactions on Wireless Communications	2015	1	0	0	1
61	Resource allocation in heterogeneous cloud radio access networks: advances and challenges	IEEE Wireless Communications	2015	0	1	0	1
62	Accelerating 5G QoE via public-private spectrum sharing	IEEE Communications Magazine	2014	1	0	0	1
63	Enabling the coexistence of LTE and Wi-Fi in unlicensed bands	IEEE Communications Magazine	2014	0	1	0	1
64	Spectrum Sharing for Device-to-Device Communication in Cellular Networks	IEEE Transactions on Wireless Communications	2014	1	0	0	1
65	Green radio communications in a heterogeneous wireless medium	IEEE Wireless Communications	2014	0	1	0	1
66	Spectrum sharing using licensed shared access: the concept and its workflow for LTE-advanced networks	IEEE Wireless Communications	2014	1	0	0	1
67	Network resource management in support of QoS in ubiquitous learning	Journal of Network and Computer Applications	2014	0.5	0	0.5	1

68	On the Feasibility of Sharing Spectrum Licenses in mmWave Cellular Systems	IEEE Transactions on Communications	2016	0.5	0	0	0.5
69	Share communication and computation resources on mobile devices: a social awareness perspective	IEEE Wireless Communications	2016	0.5	0	0	0.5
70	Load Balancing in Two-Tier Cellular Networks With Open and Hybrid Access Femtocells	IEEE/ACM Transactions on Networking	2016	0.5	0	0	0.5
71	Load balancing mechanisms and techniques in the cloud environments: Systematic literature review and future trends	Journal of Network and Computer Applications	2016	0.5	0	0	0.5
72	Dynamic traffic steering based on fuzzy Q-Learning approach in a multi-RAT multi-layer wireless network	Computer Networks	2014	0.5	0	0	0.5
73	Energy-efficient non-cooperative cognitive radio networks: micro, meso, and macro views	IEEE Communications Magazine	2014	0	0.5	0	0.5
74	A Cooperative Matching Approach for Resource Management in Dynamic Spectrum Access Networks	IEEE Transactions on Wireless Communications	2014	0.5	0	0	0.5
75	Adaptive Network Coding for Spectrum Sharing Systems	IEEE Transactions on Wireless Communications	2014	0.5	0	0	0.5
76	An efficient location aware distributed physical resource block assignment for dense closed access femtocell networks	Computer Networks	2016	0	0	0	0

77	Adaptive Resource Allocation Algorithm of Lyapunov Optimization for Time-Varying Wireless Networks	IEEE Communications Letters	2016	0	0	0	0
78	Resource Allocation for Wireless Powered Multi-Pair Massive Antenna Relaying With Zero Forcing Beamforming	IEEE Communications Letters	2016	0	0	0	0
79	Convexity of Fairness-Aware Resource Allocation in Wireless Powered Communication Networks	IEEE Communications Letters	2016	0	0	0	0
80	SDN meets SDR in self-organizing networks: fitting the pieces of network management	IEEE Communications Magazine	2016	0	0	0	0
81	Spectrum Management for Proactive Video Caching in Information-Centric Cognitive Radio Networks	IEEE Journal of Selected Areas in Communications	2016	0	0	0	0
82	Spectrum Sharing in mmWave Cellular Networks via Cell Association, Coordination, and Beamforming	IEEE Journal of Selected Areas in Communications	2016	1	0	0	0
83	Resource Allocation and Fairness in Wireless Powered Cooperative Cognitive Radio Networks	IEEE Transactions on Communications	2016	0	0	0	0
84	A Batch-Based MAC Design With Simultaneous Assignment Decisions for Improved Throughput in Guard-Band-Constrained Cognitive Networks	IEEE Transactions on Communications	2016	0	0	0	0

85	A generalized resource allocation framework in support of multi-layer virtual network embedding based on SDN	Computer Networks	2015	0	0	0	0
86	Resource Allocation for Device-to-Device Communication Underlaying Cellular Networks: An Alternating Optimization Method	IEEE Communications Letters	2015	0	0	0	0
87	Resource allocation in full-duplex communications for future wirelessnetworks	IEEE Wireless Communications	2015	0	0	0	0
88	QoE-driven spectrum assignment for 5G wireless networks using SDR	IEEE Wireless Communications	2015	0	0	0	0
89	QoE/QoS-aware LTE downlink scheduler for VoIP with power saving	Journal of Network and Computer Applications	2015	0	0	0	0
90	JMB: scaling wireless capacity with user demands	Communications of the ACM	2014	0	0	0	0
91	A survey on resource allocation techniques in OFDM(A) networks	Computer Networks	2014	0	0	0	0
92	Network resource control for Xen-based virtualized software routers	Computer Networks	2014	0	0	0	0
93	User-in-the-loop: spatial and temporal demand shaping for sustainable wireless networks	IEEE Communications Magazine	2014	0	0	0	0
94	Spectrum- and energy-efficient D2DWRAN	IEEE Communications Magazine	2014	0	0	0	0
95	Context-aware quality of service in wireless sensor networks	IEEE Communications Magazine	2014	0	0	0	0

96	Efficient Spectrum Utilization on TV Band for Cognitive Radio Based High Speed Vehicle Network	IEEE Transactions on Wireless Communications	2014	0	0	0	0
97	Cooperative Secure Resource Allocation in Cognitive Radio Networkswith Guaranteed Secrecy Rate for Primary Users	IEEE Transactions on Wireless Communications	2014	0	0	0	0
98	Coalitional Games for Resource Allocation in the Device-to-Device Uplink Underlaying Cellular Networks	IEEE Transactions on Wireless Communications	2014	0	0	0	0
99	Proportionally Fair Distributed Resource Allocation in Multiband WirelessSystems	IEEE/ACM Transactions on Networking	2014	0	0	0	0
100	Double Auctions for Dynamic Spectrum Allocation	IEEE/ACM Transactions on Networking	2014	0	0	0	0
101	An enhanced fast handover with seamless mobility support for next-generation wireless networks	Journal of Network and Computer Applications	2014	0	0	0	0
102	A novel resource allocation scheme for LTE network in the presence of mobility	Journal of Network and Computer Applications	2014	0	0	0	0
103	Reliable multi-channel scheduling for timely dissemination of aggregated data in wireless sensor networks	Journal of Network and Computer Applications	2014	0	0	0	0

Table A.1: SLR papers set

APPENDIX B TRAFFIC LOAD FORECASTING MODELS EVALUATION

B.1 Evaluation of Traffic Load Forecasting Models

With regard to the performance of the proposed solution, it is important to analyze is the accuracy of the traffic load forecasting for each model. The forecasting procedure follows three key phases. The first is the time series extraction of traffic data from LTE-LSA and Wi-Fi networks. The second consists of fitting the polynomial curve of traffic data of both LTE-LSA and Wi-Fi networks. In the third phase, the forecasting is carried out by means of the machine learning models, as outlined in the following sections.

The normalization of the traffic load for each CoS of the LTE-LSA network and the aggregate traffic of each Wi-Fi network is based on time series with treatment effect. According to Hamilton (HAMILTON, 1943) the traffic load measurement that corresponds to a time series of 900 seconds. The validation dataset consists of 10 percent of the testing dataset which is used to analyze the outcomes of the prediction, by taking into account metrics such as accuracy and processing time.

B.1.1 Traffic Load Forecasting with the Multiple Linear Regression Model

The MLRM processes the trained dataset of each simulated traffic demand per CoS *i.e.,* Video, VoIP, and HTTP from the LTE-LSA network, as well as the aggregate traffic of Wi-Fi networks. The simulated traffic in the LTE-LSA network and the traces in Wi-Fi are computed in units of seconds to improve the accuracy of the model. The first step of the analytical methodology involves calculating the polynomial curve fitting for the smoothing out the peaks and noise of overall traffic *i.e.,* it is used for training, validation and testing datasets. The polynomial was fixed at 10 degrees for curve fitting analysis of the traffic of each network. The classification is then performed again and includes the new data points obtained from the ten degrees polynomial for the training, validation, and testing datasets. After this, the MLRM is applied to carry out the traffic load forecasting and the validation dataset is used to evaluate its accuracy for each network.

The MLRM accuracy is evaluated by the cross-validation method which involves comparing the forecast values with the current values. At this point, the MLR model can be adjusted to improve the accuracy of the upcoming predictions. The MAPE Equation (4.9) is also used to measure the accuracy of the MLRM. Figure B.1 enables the analysis of the accuracy of the traffic load forecasting of the three overlapping Wi-Fi networks in relation to current traffic load. As can be seen in the graph, the traffic load forecasting was very accurate, and reached levels of 96.179%, 93.607%, and 94.197% degree of accuracy, for Wi-Fi networks 1, 2, and 3, respectively.

Another analysis that is conducted for traffic load forecasting concerns the LTE-LSA fre-

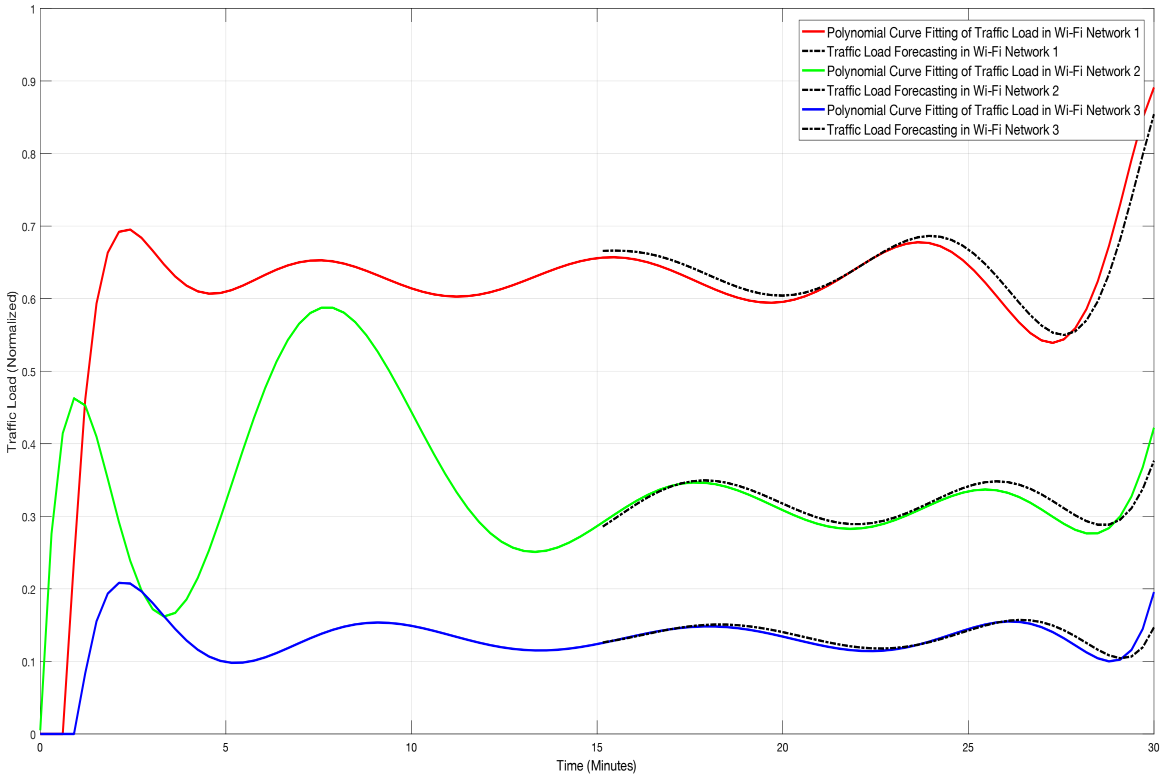


Figure B.1: Current vs. Predicted Traffic Load for each Wi-Fi Network with the Multiple Linear Regression Model

quencies. This analysis is needed to forecast the behavior of the incumbent user, which may request an evacuation. The outcomes of the simulation related to this scenario are shown in Figure B.2 which shows the traffic load forecasting of VoIP, Video, and HTTP in the LTE-LSA network. In this case, the levels of accuracy are up to 95.724%, 98.473%, and 94.764% respectively, for each class of service.

B.1.2 Traffic Load Forecasting using the Neural Network Model

The NN is modeled by the Cascade Forward Back Propagation (CFBP) and carries out the following steps for traffic load forecasting of each CoS in the LTE-LSA network, and the aggregate traffic of each Wi-Fi network. As depicted in Figure B.4, the forecasting of HTTP, Video, and VoIP traffic loads change according to the historical traffic load of each network. The trained data of the NN model is collected by the top levels of the resources broker. The first step involves the normalization of the training, validation, and testing datasets in time series. The second step consists on carry out the current implementation of CFBP model in Matlab NN toolbox for traffic load forecasting.

The CFBP model processes the trained dataset of each simulated traffic demands *i.e.,* Video, VoIP, and HTTP from the LTE-LSA network, as well as of the aggregate traffic of Wi-Fi networks. The simulated traffic in the LTE-LSA network and the traffic traces of Wi-Fi are com-

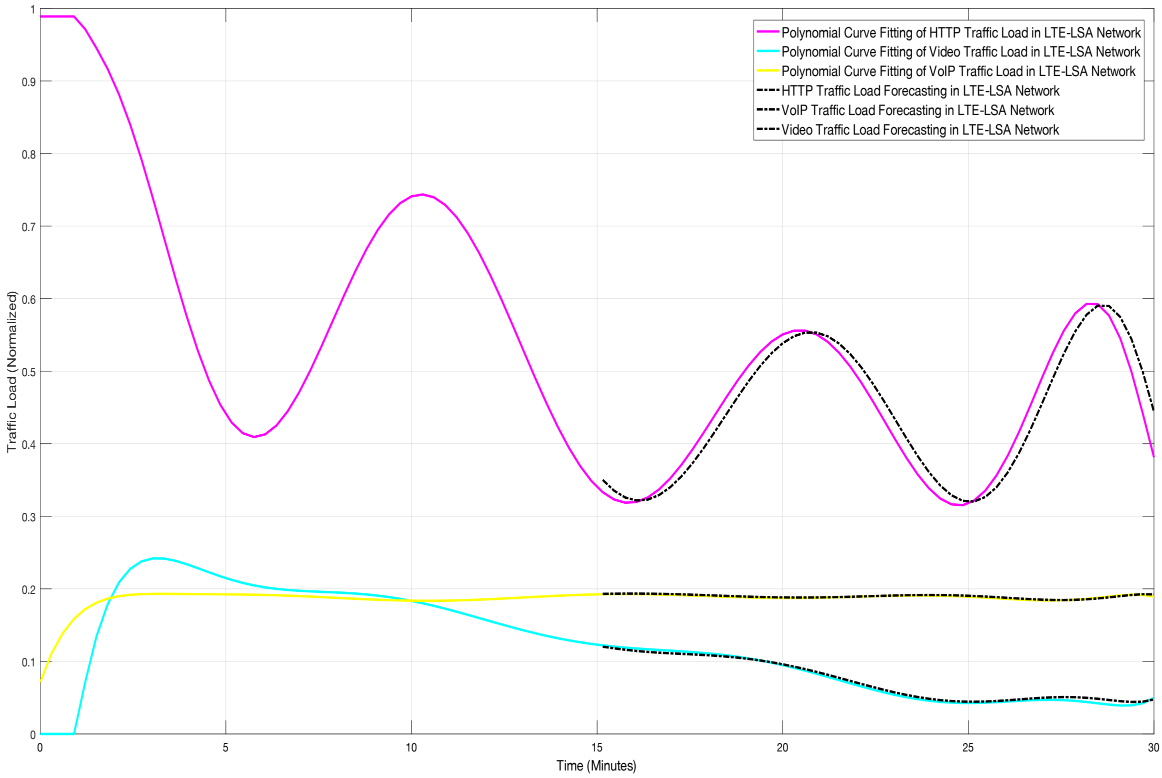


Figure B.2: Current vs. Predicted Traffic Load for each CoS in the LTE-LSA Network with the Multiple Linear Regression Model

puted in units of seconds to improve the accuracy of the results. The first step of the analytical methodology implicates calculating the polynomial curve fitting tor smoothing out the peaks and noise of the traffic in the LTE-LSA and Wi-Fi networks. The polynomial was also fixed at 10 degrees to achieve the best curve fitting of the traffic load of each network. The second step consists in the classification for the training, validation, and testing datasets which are composed by the new data points obtained from 10 degrees polynomial. The third step regards the CFBP model carrying out the traffic load forecasting. Then, the prediction accuracy of each network is evaluated by means of the validation dataset.

The predictions accuracy of CFBP model are calculated by means of the MAPE Equation (4.9). As well as, the cross-validation method enables to compare the current values with the predicted ones. Thus, the CFBP model can be auto adjusted to improve the accuracy of the future predictions. Figure B.7 shows the analysis of the accuracy degree of the traffic load forecasting which examines three Wi-Fi networks as possible targets for traffic steering procedures. Such analysis is on the basis of the relation among the current and predicted traffic load for each network. In fact, the traffic load forecasting for Wi-Fi networks 1, 2, and 3 achieved the degree average accuracy of 96.727%, 95.924%, and 98.633%, respectively. Also, the traffic load forecasting of each CoS from the LTE-LSA network is used by the provisioning algorithm to estimate the load of traffic that will be steered towards overlapping Wi-Fi networks and thus prevent future network congestion.

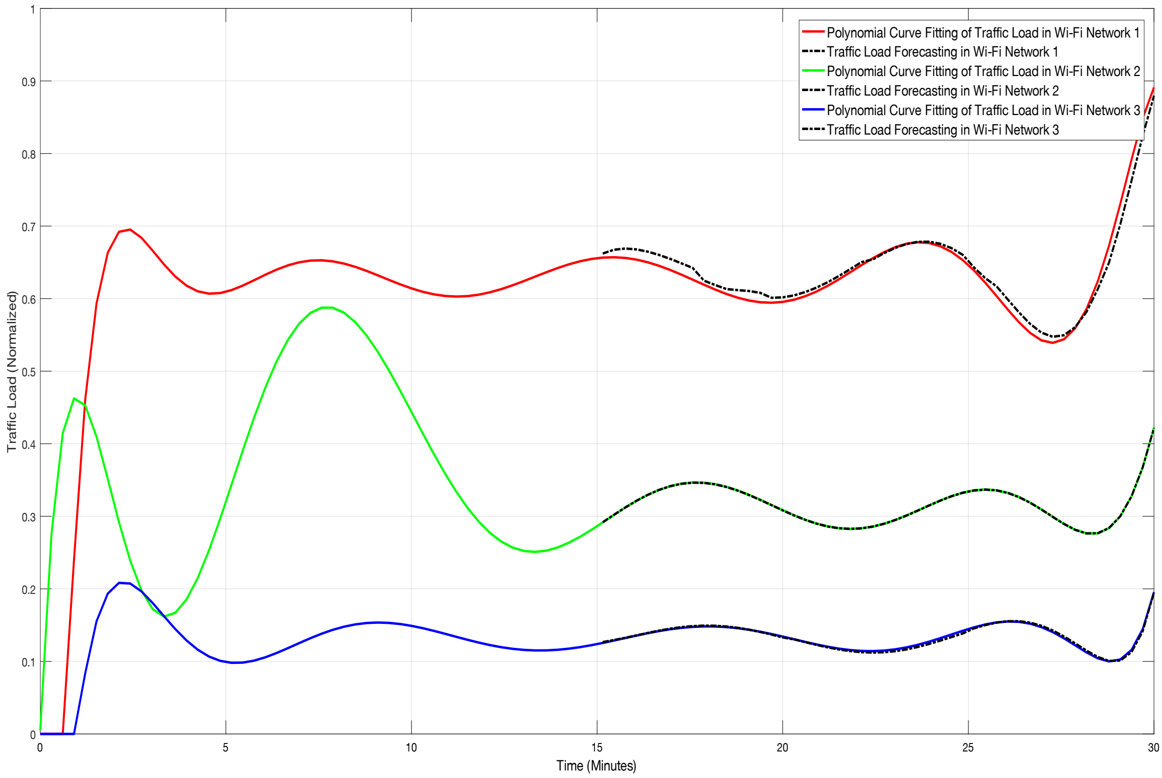


Figure B.3: Current vs. Predicted Traffic Load for each Wi-Fi Network using the Neural Network Model

Figure B.4 depicts the traffic load forecasting of VoIP, Video, and HTTP in the LTE-LSA network. In this case, the average accuracy degree achieved by the CFBP prediction model is 99.547% (VoIP), 98.559% (HTTP), and 98.249% (Video), respectively.

Whenever the CFBP model carry out the traffic load forecasting, the updated data points values of the time series prediction can be used as input to the provisioning algorithm. This algorithm is responsible for finding the best network for handover and traffic steering procedures. Thus, the resources broker has as first action to estimate the availability and the resources occupation of each target network based on previous forecasting results obtained by the implementation of the CFBP model. The second action involves selecting the Wi-Fi networks which are able to guarantee the same level of QoS as that offered in the LTE-LSA network. This kind of decision is taken considering the predicted availability of network resources. The main resource, in this case, is the network capacity for next 15 minutes. However, in order to guarantee the same QoS level, the proposed solution also considers the delay and jitter QoS metrics, when applicable. The third action which is performed by the provisioning algorithm consists on the association of the CoS to the decision process.

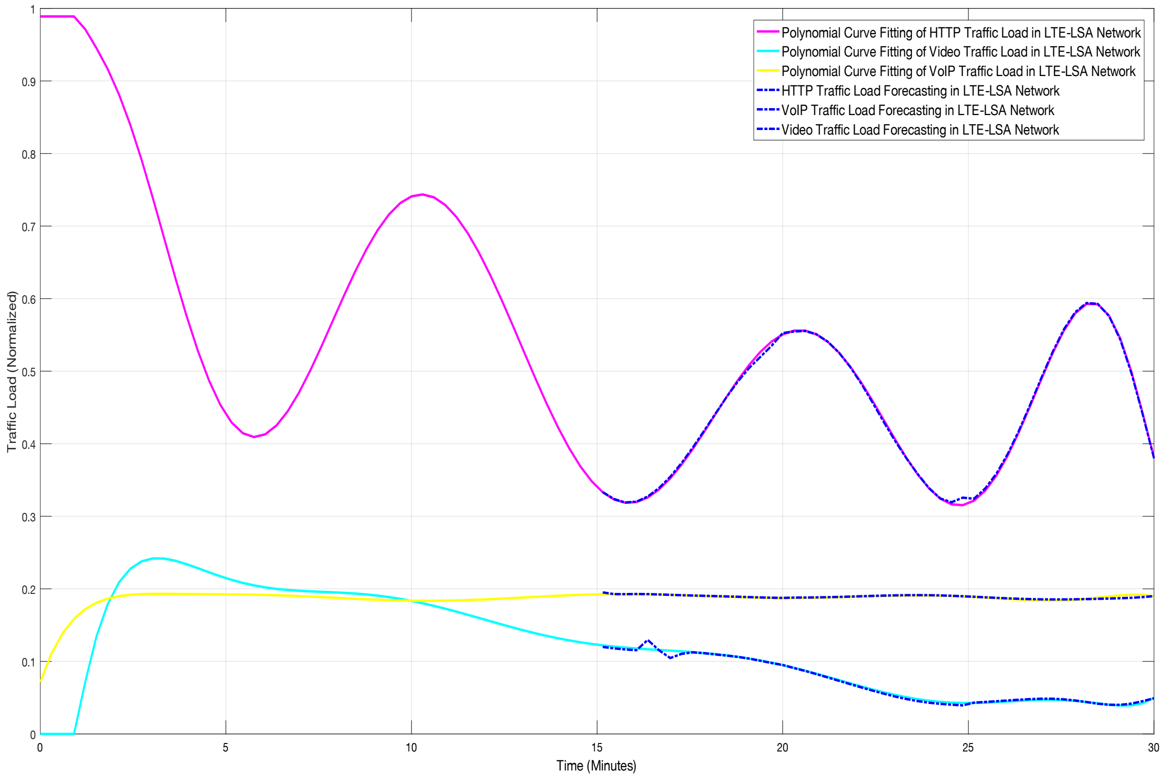


Figure B.4: Current vs. Predicted Traffic Load for each CoS in the LTE-LSA Network using the Neural Network Model

B.1.3 Traffic Load Forecasting using the Regression Tree Model

The traffic predictions carried out by the RTM are also evaluated through MAPE Equation (4.9) to calculate the prediction accuracy degree, as well as to compare the current values with the predicted by means of the cross-validation method. Figure B.5 shows the analysis of the accuracy degree of the traffic load forecasting which examines three Wi-Fi networks as possible best target network(s) for vertical handover and traffic steering procedures. Such analysis is based on the relation among the current and predicted traffic load for each network. Indeed the traffic load forecasting for Wi-Fi networks 1, 2, and 3 achieved the degree average accuracy of 96.549%, 28.305%, and 94.830%, respectively. Also, the traffic load forecasting of each CoS from the LTE-LSA network is used by the proposed decision algorithm to estimate the load of traffic that will be steered towards overlapping Wi-Fi networks and thus prevent future network congestion.

The outcomes of the simulation related to this scenario are depicted in Figure B.6 which shows the traffic load forecasting of VoIP, Video, and HTTP in the LTE-LSA network. In this case, the forecasting accuracy degree was in average around to 89.394% (VoIP), 97.025% (HTTP), and 96.064% (Video), respectively.

Whenever the RTM carries out the traffic load forecasting, the data points values of time series prediction are updated and entered as input to the provisioning algorithm. For this purposes,

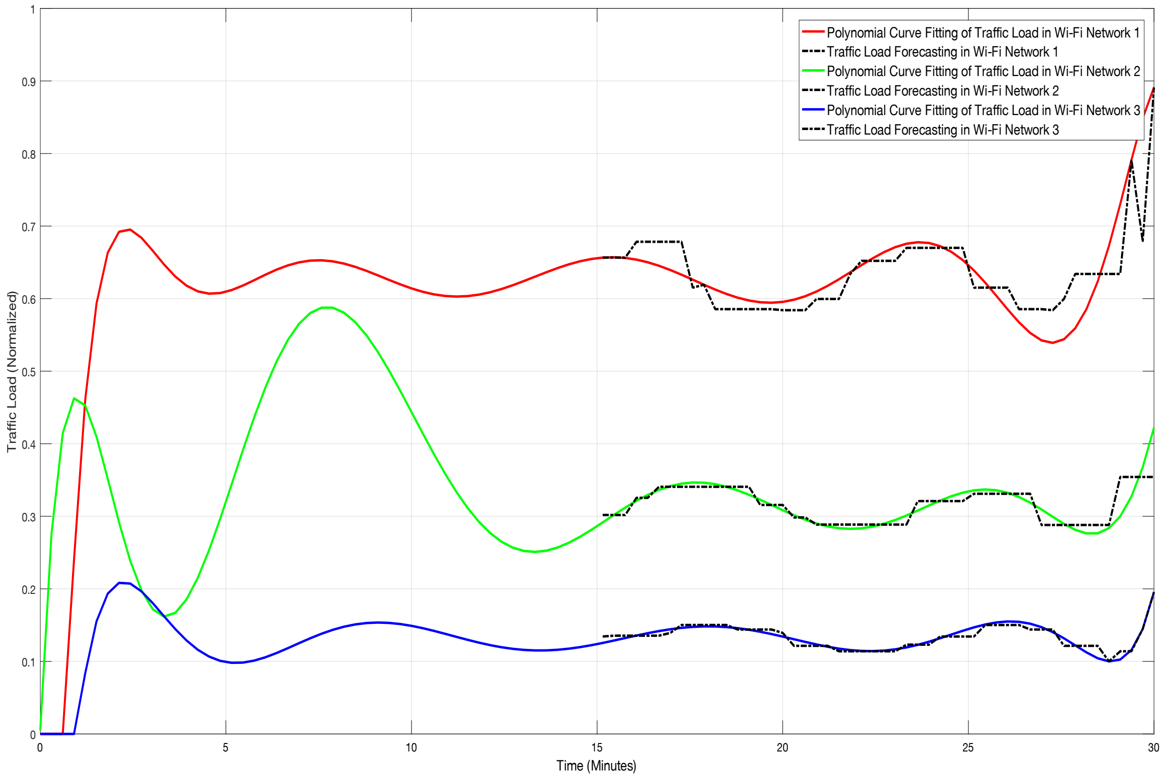


Figure B.5: Current vs. Predicted Traffic Load for each Wi-Fi Network using the Regression Tree Model

the resources occupancy is estimated for each target network

B.1.4 Traffic Load Forecasting using the Fourier Model

The traffic predictions carried out by the FM are evaluated using the MAPE Equation (4.9) to calculated the accuracy degree, as well as to compare the current values with the predicted is used the cross-validation method. Figure B.7 shows the analysis of the accuracy degree of the traffic load forecasting.The traffic load forecasting achieved an average accuracy degree of 96.327%, 94.936%, and 92.310% for Wi-Fi networks 1, 2, and 3, respectively. The traffic load forecasting of each CoS of the LTE-LSA network is used by the provisioning algorithm to estimate the load of traffic that will steer toward the overlapped Wi-Fi networks. The outcomes of the simulation related to this scenario are depicted in Figure B.8 which shows the traffic load forecasting of VoIP, Video, and HTTP in the LTE-LSA network. In this case, the degree of forecasting accuracy is in average to 91.015% (VoIP), 99.198% (HTTP), and 90.091% (Video), respectively.

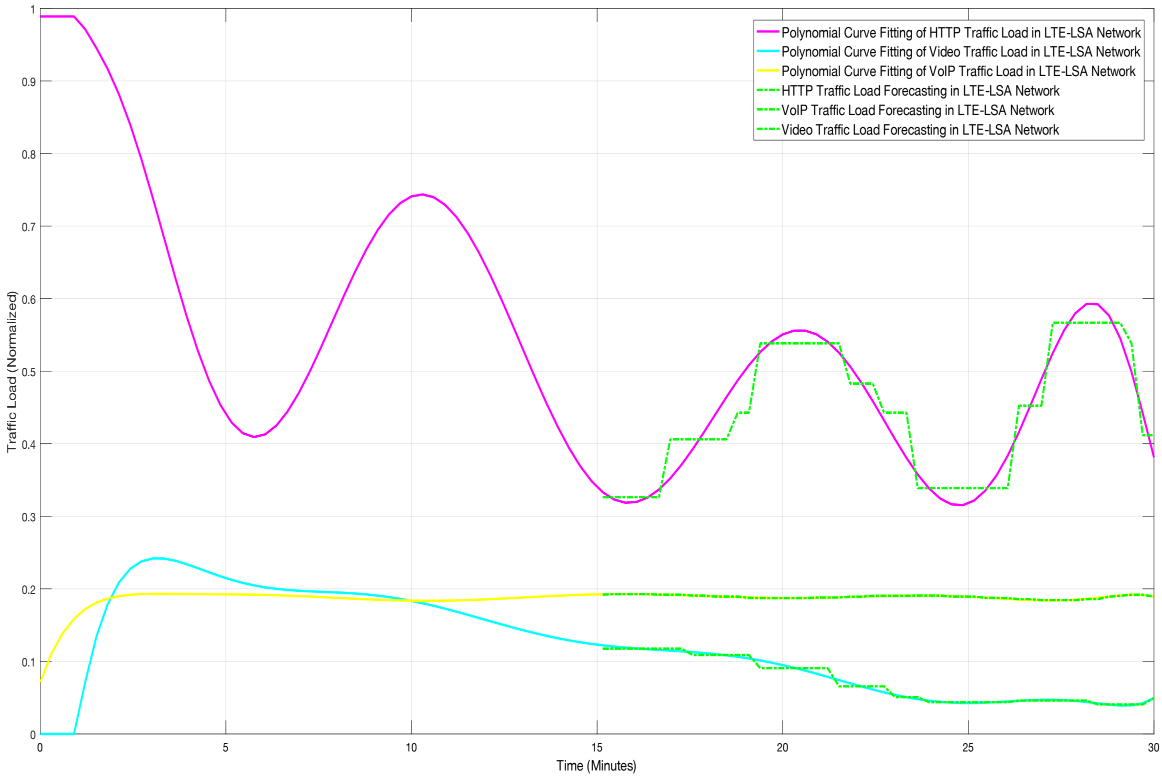


Figure B.6: Current vs. Predicted Traffic Load for each CoS in the LTE-LSA Network using the Regression Tree Model

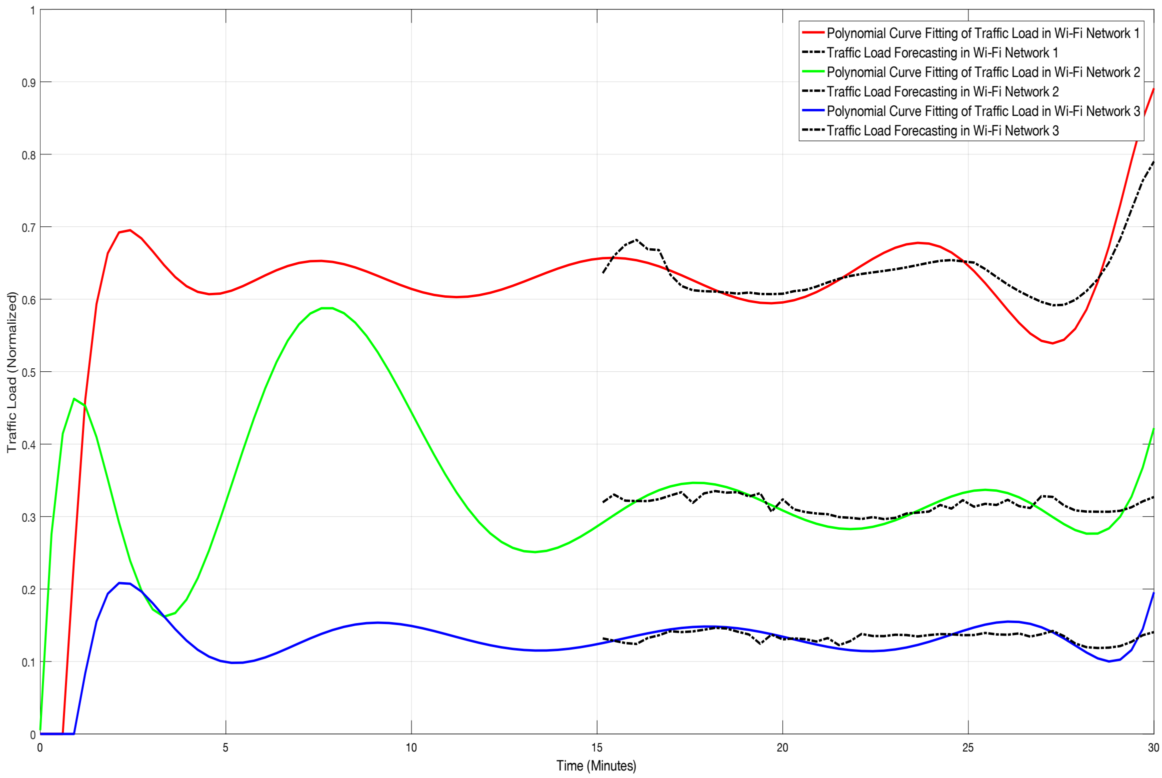


Figure B.7: Current vs. Predicted Traffic Load for each Wi-Fi Network using the Fourier Model

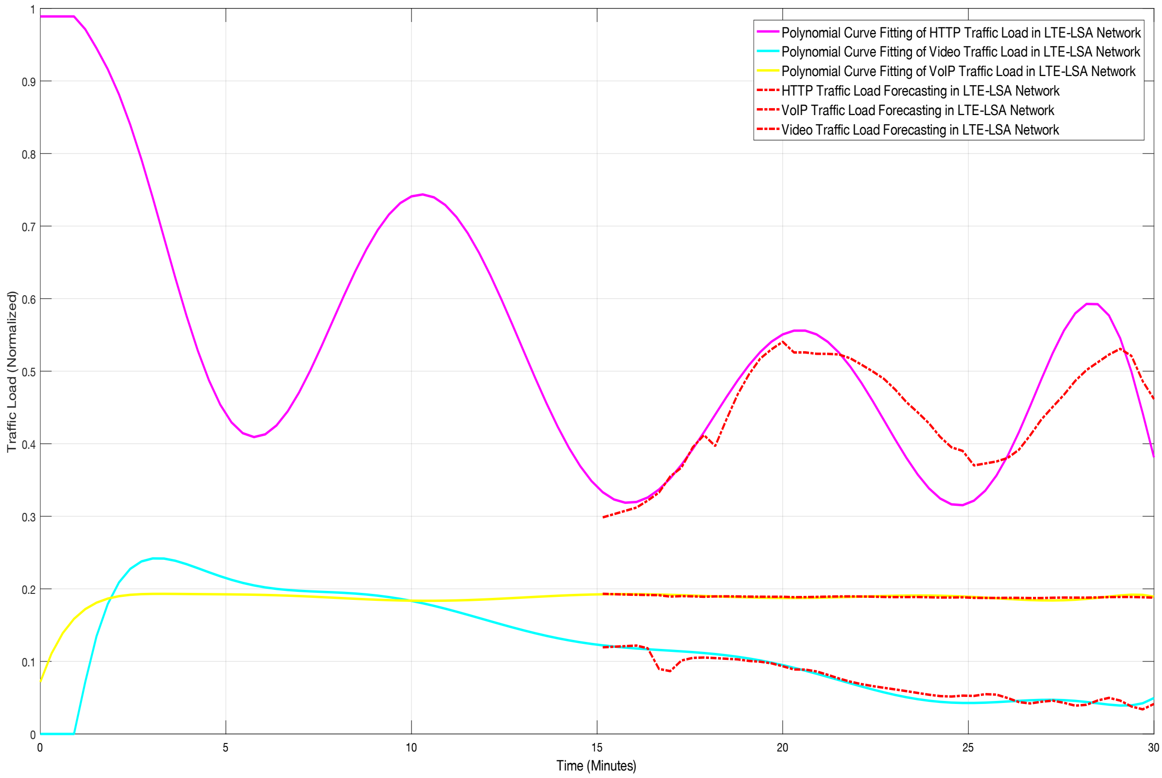


Figure B.8: Current vs. Predicted Traffic Load for each CoS in the LTE-LSA Network using the Fourier Model

B.2 Comparison Table of Performance Metrics for each Traffic Load Forecasting Model

The evaluation of traffic load forecasting performance in time and accuracy metrics are fundamental for the proposed solution. According to Table B.1, the RTM is the fastest model to predict the future 15 minutes of traffic load for the Wi-Fi 1 and 2. However, the MLR is the fastest model to predict the next 15 minutes of traffic load for Wi-Fi 3 in the analyzed scenario. In the case of Table B.1, the results obtained represent the average time for each model that carry out the prediction of aggregate traffic load for each Wi-Fi network. Furthermore, the Table B.2 contains the output of the mean absolute percentage error for each Wi-Fi network, through these results can be estimated the accuracy rate of the prediction.

According to Table B.2, the NN model has the lowest MAPE and thus the most accurate prediction model for the aggregate traffic load of each Wi-Fi network in the proposed scenario. In addition, the MLRM and FM can be considered for traffic load forecasting in Wi-Fi networks due to the MAPE outputs are very closed to the NNM.

According to Table B.3, MLRM has the fastest average time for traffic load forecasting for each CoS in the LTE-LSA network.

able B.4, shows that NN obtains the best MAPE for each CoS in the LTE-LSA network. In addition, for traffic load forecasting of each CoS in the LTE-LSA network can also be considered the MLRM since this model obtained stable MAPE results.

Table B.1: Average Time for Traffic Load Forecasting in Wi-Fi Networks

Forecasting Model	**Wi-Fi Network 1 [s]**	**Wi-Fi Network 2 [s]**	**Wi-Fi Network 3 [s]**
Multiple Linear Regression	6.525	0.072449	0.04258
Neural Network	12.289	383.088	5.564
Regression Tree	0.024	0.02478	0,106
Fourier	4,909	0,696	0,733737

Table B.2: Mean Absolute Percentage Error of Wi-Fi Traffic Load Forecasting Models

Forecasting Model	**Wi-Fi Network 1 [%]**	**Wi-Fi Network 2 [%]**	**Wi-Fi Network 3 [%]**
Multiple Linear Regression	3.821%	6.393%	5.803%
Neural Network	3.273%	4.076%	1.367%
Regression Tree	3.451%	71.695%	5.170%
Fourier	3.673%	5,064%	7,690%

Table B.3: Average Time for Traffic Load Forecasting for Each CoS in the LTE-LSA Network

Forecasting Model	**HTTP [s]**	**VoIP [s]**	**Video [s]**
Multiple Linear Regression	0.0118	0.0156	0.0170
Neural Network	12.289	383.088	5.564
Regression Tree	7.6403	211.350	82.013
Fourier	0.7579	0.18967	0.1223

Considered all aspects evaluated regarding the traffic forecasting models, in average, MRLM is the one which better fits to the solutions proposed in this thesis. Therefore, this model was selected to be considered in the implementation of the simulations. However, the proposed resources broker is modular and allows the substitution of the forecasting model.

Table B.4: Mean Absolute Percentage Error of Traffic Load Forecasting Models for each CoS in the LTE-LSA Network

Forecasting Model	**HTTP [s]**	**VoIP [s]**	**Video [s]**
Multiple Linear Regression	4.276%	1.527%	5.236%
Neural Network	0.453%	1.441%	1.751%
Regression Tree	10.606%	2.975%	3.936%
Fourier	8.985%	0.802%	9.909%

APPENDIX C PUBLISHED PAPER – IEEE WIMOB 2016

Providing high quality network access is challenging for network operators in the current static model used for allocating the spectrum of frequencies. Dealing with this challenge demands optimized resources allocation. One of the ways of providing this optimization is by allowing resources sharing among network operators which share the same geographical area. To allow and control the sharing of resources in such network scenarios, in this paper, a multilevel broker is presented to allow network operators to share their underutilized resources. This broker dynamically establishes a service level agreement that takes into account the quality of service requirements of the resources renters. A performance evaluation conducted in a scenario composed of multiple LTE-Advanced network operators shows that the implementation of the proposed architectures leads to more efficient allocation of underutilized network resources compared to two algorithms found in the literature.

- **Title –**
  Improving QoS in Multi-operator Cellular Network
- **Conference –**
  IEEE International Conference on Wireless and Mobile Computing, Networking and Communications
- **Date –**
  October, 2016
- **Held at –**
  New York, USA

Improving QoS in Multi-operator Cellular Networks

Rafael Kunst, Leandro Avila, Edison Pignaton, Sergio Bampi, Juergen Rochol
Federal University of Rio Grande do Sul (UFRGS)
Email: {rkunst, laavila, edison.pignaton, bampi, juergen}@inf.ufrgs.br

***Abstract*— Providing high quality network access is challenging for network operators in the current static model used for allocating the spectrum of frequencies. Dealing with this challenge demands optimized resources allocation. One of the ways of providing this optimization is by allowing resources sharing among network operators which share the same geographical area. To allow and control the sharing of resources in such network scenarios, in this paper, a multilevel broker is presented to allow network operators to share their underutilized resources. This broker dynamically establishes a service level agreement that takes into account the quality of service requirements of the resources renters. A performance evaluation conducted in a scenario composed of multiple LTE-Advanced network operators shows that the implementation of the proposed architectures leads to more efficient allocation of underutilized network resources compared to two algorithms found in the literature.**

I. Introduction

Providing high quality network access is a current challenge for network operators, due to the increasing demand for both wireless applications and services. This scenario becomes even more challenging if the current spectrum allocation model is taken into account. The spectrum allocation in this model is based on exclusive licenses issued to determined network operators, what may lead to resources underutilization and ultimately cause a resources scarcity problem. Considering this scenario, to offer quality of service (QoS) enabled solutions to the customers, a network operator needs to be able to deal with the resources scarcity problem.

The main motivation of this paper is to provide a solution to improve the QoS offered to the network users by dealing with the spectrum scarcity problem. The proposed approach considers the implementation of resources sharing among network operators, supporting both homogeneous and heterogeneous network scenarios. This motivation is based on the premise that both commercial and non-commercial wireless network operators have a certain amount of unused resources during off-peak hours. By sharing unused resources with coexisting operators in the same geographical area, a network operator can improve its profit, while the resources renter is able to improve the QoS offered to its customers.

Many related work have been proposed recently to allow resources sharing in both homogeneous and heterogeneous network scenarios. These works typically consider the spectrum availability in two different ways. In situations where licenses are not required for accessing a given frequency range, the concept of Collective Use of Spectrum (CUS) applies. Otherwise, when licenses are necessary, Licensed Shared Access (LSA) is considered to implement resources sharing [1]. Most of the current research, although very relevant, generally identify the need for additional resources and search for available resources without taking into account the QoS requirements of the resources renter. This common approach of related works may lead to the allocation of network resources that are not suitable for the users.

Considering the aspects not fully covered by the related work, in this paper, a multilevel broker is implemented to allow resources sharing among network operators which coexist in a given geographical area. This broker is able to operate considering both homogeneous and heterogeneous network environments. QoS requirements of the resources renter are considered via the establishment of a dynamic Service Level Agreement (SLA). This multilevel broker provides interfaces and a control protocol to allow communication of different network operators. The control protocol has been designed to exchange a small amount of control information to prevent the overhead from significantly interfering with the network traffic.

The multilevel broker is modeled and simulated using Matlab to evaluate its behavior in a network scenario where four LTE-Advanced network operators share the same geographical area. Although the proposed approach allows the coexistence of heterogeneous network technologies, the evaluation was performed using only LTE-Advanced to reflect a scenario typically found in the reality. The main contributions of the paper are listed as follows:

- Design and simulation of a multilevel broker, capable of allowing resources sharing among network operators which coexist in the same geographical area;
- Analysis of the impact caused by the multilevel broker on the QoS offered by the network operators to its customers in a typical commercial scenario;
- Analysis of the overhead involved in the implementation of the multilevel broker.
- Advancements on spectrum sharing research, since the simulation results prove that the proposed solution allows more efficient resources allocation in comparison with related work.

The remainder of this paper is organized as follows. Related works are presented in section II. The resources sharing architecture and the multilevel broker are discussed in III. The simulation model and the performance evaluation are presented in section IV. The paper is concluded in section V, which also brings directions for future work.

II. Related work on Wireless Resource Sharing

Gardellin *et al.* [2] dealt with the coexistence of different cognitive radio cellular networks. The coexistence problem is stated in terms of channel assignment between the cells, where the cooperative and non-cooperative schemes are compared using a fairness index, which is based on the throughput of each Customer Premises Equipment (CPE). The channels are assigned based on their quality, measured considering the Signal to Noise plus Interference Ratio (SINR). Using this parameter, the authors attempted to find an appropriate set of channels without harmful interference, making the coexistence possible. The authors obtained results that allowed to conclude that the cooperative operation achieves a better fairness index compared to non-cooperative and random methods.

Dixit *et al.* [3] proposed a framework to implement cooperative coexistence between primary (licensed) and secondary (unlicensed) users of LTE networks. The main goal of the work was to optimize spectrum utilization, bringing as an advantage to the primary network operator the possibility of earning profits by leasing white spaces. A pricing model was introduced as a way to allow temporary access for secondary users when the network resources are underutilized.

Chatzikokolakis *et al.* [4] analyzed the requisites and technical enablers of spectrum sharing in the context of heterogeneous networks and different frequency allocation regimes. Beyond the theoretical analysis, the authors also propose a simple spectrum sharing mechanism based on fuzzy logic. The proposed algorithm was used to match the needs of the spectrum renter by selecting the more suitable spectrum frequencies to serve its demands. A functional architecture was used to allow communication among primary and secondary users. The results obtained via simulation showed that the proposed architecture and the artificial intelligence algorithm together provide an increase on the spectrum allocation efficiency.

Spectrum sharing among co-primary 5G small cell networks was investigated by Singh *et al.* [5]. A non-cooperative protocol was proposed to keep the overhead low. Such a protocol is based on minimizing the cost involved for a given network operator to rent resources from another 5G network operator. The model used to describe the costs is based on spectrum favors. Two approaches were proposed to meet various network operation scenarios. An instantaneous reciprocity model was applied in situations where the operators are considered impatient. On the other hand, a long-term reciprocity was proposed to be used when operators have persistent and publicly known identity, so the operators can learn from each other behavior. In both approaches, the cost spectrum favors were calculated based on a repeated non-cooperative game.

Table I summarizes the relevant previous works considering whether or not five different aspects are covered by each proposal. The first aspects refer to the type of network resources that can be shared, *i.e.* the support for sharing CUS and/or LSA resources. The support for heterogeneous networks is also considered to compare the proposals. Finally, the cost and the QoS support are also analyzed.

TABLE I

COMPARISON BETWEEN THIS PROPOSED APPROACH AND RELATED WORK

Work	**CUS**	**LSA**	**Heterogeneous Networks**	**Cost**	**QoS**
[2]	X				X
[3]		X		X	
[4]	X	X	X		X
[5]	X	X	X	X	
Mutilevel Broker	X	X	X	X	X

III. RESOURCES SHARING ARCHITECTURE

In this section, the concepts of the proposed architecture are discussed. The designed solution is presented in III-A. In III-B, the system model used to assess the amount of resources available to each network operator is presented. Another important feature of the proposed approach is the signaling protocol provided to allow intercommunication among the network operators. This protocol is discussed in III-C.

A. Design of the Resources Sharing Architecture

The design of the proposed architecture is presented in Fig. 1. The illustration is divided in two parts which communicate through a polling and reply mechanism. In the left side, the spectrum users coexist in a geographical area considering a scenario that allows one operator to communicate with all neighboring network operators. In the right side of the figure, the structure of a broker is represented. This broker is responsible for coordinating resources sharing among the spectrum users.

In the illustration, an example scenario is presented to reflect the behavior of typical spectrum users. In this scenario, four LTE-Advanced network operators are represented to illustrate the reality of LTE-Advanced frequencies allocation in Brazil. IEEE 802.22 CR and IEEE 802.11 network operators were also represented to indicate that the proposed approach allows the coexistence and resources sharing between primary and secondary users belonging to heterogeneous networks.

Another important aspect to highlight is related to the direction of the resources sharing. The proposed architecture allows resources sharing in two ways, *i.e.* each network operator can dynamically assume the role of a resources provider or the role of a resources renter. In Fig. 1, the direction of resources provider is represented by a straight connector, while resources renting is represented by a dashed connector. Although, for the sake of simplicity in the representation each secondary user is communicating with only one LTE-Advanced network operator, the architecture, indeed, allows the secondary users to communicate with any other network operator within the same geographical area.

Different types of resources can be shared in the proposed approach. Further than allowing the cooperation between primary and secondary users, the architecture permits the cooperation between wireless network technologies which operate using diverse kinds of resources, *e.g.* spectrum of frequencies and channel capacity. In Fig. 1, LTE-Advanced is an example of technology based on channel capacity, while IEEE 802.22 conducts spectrum sensing to directly transmit over the spectrum of frequencies. To turn feasible the translation between two kind of resources, a centralized entity is necessary. Moreover, a centralized approach is indicated in situations where the primary user activity do not change constantly. The application scenario of the proposed approach fits this criteria in a long-term observation, *e.g* 24 hours, since wireless cellular network operators deal with predictable traffic in most situations [1].

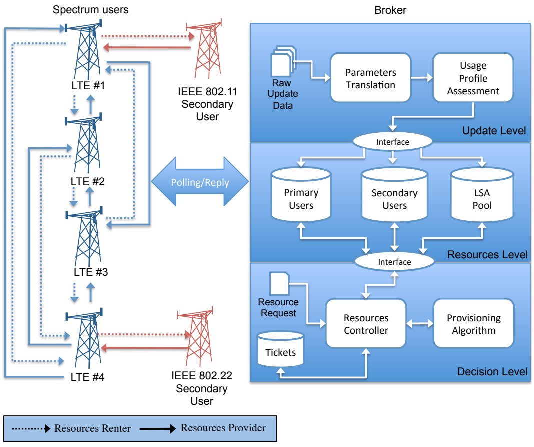


Fig. 1. Architecture Design

A novel multilevel broker is proposed to play the role of the centralized entity in the resources sharing architecture. Three levels were defined to provide independent and simultaneous control of different tasks of spectrum sharing management. These levels were named accordingly to the function executed by each one: (I) Update Level, (II) Resources Level, and (III) Decision Level. These levels are interconnected by interfaces which implement the flow of information that allows information exchange among the different levels of the broker.

Update Level
The update level is responsible for collecting operation parameters from the network operators which participate in the spectrum sharing initiative. The updating mechanism is based on the implementation of a polling-based technique, which is controlled by the Parameters Translation block of the broker. This block allows the configuration of the interval between polls. The precise definition of such interval is crucial to deal with the trade-off between having accurate information about the current resources usage profile of each network operator and the overhead generated by the control information transmitted to update the broker. Another important function performed by the Parameters Translation block is the translation of the raw update data into useful information to allow the architecture to take proper decisions regarding to resources sharing. Therefore, the definition of the structure used by the network operators to update the broker is very important. The structure is defined by the tuple: [Operator ID, Delay, Jitter, Throughput].

Upon receiving the raw update data, the Parameters Translation block performs a SINR estimation in the wireless channel and collects the timestamp of the instant when the raw update information was received. These two parameters complement the ones informed by the network operator and are used respectively to estimate the overall network load and to provide information for historical assessment of each QoS parameter.

The pre-processed raw data is sent to the Usage Profile Assessment block, which applies the concepts of cognition to keep track of the historical information provided by the network operators. This historical information is taken into account to define the current usage profile of the network to minimize the effect of abnormal behaviors of the traffic that may occur in realistic operation scenarios. The weight given to the historical information (α) and the weight considered to the most recent update $(1-\alpha)$ are parameters of this block. Equation 1 is applied to calculate the weighted load (ℓ) of each considered QoS parameter of the network operators.

$$\ell = \alpha . \sum_{i=1}^{n} \ell_{(t-i)} + (1-\alpha).\ell_{(t)} \tag{1}$$

This equation considers a pre-defined number (n) of historical evaluations of ℓ and performs an exponential smoothing to obtain the weighted load of a given QoS parameter. The same equation is applied to the remaining QoS parameters in order to obtain the complete assessment of the usage profile of each network operator. The processed usage profile information is then sent to the Resources Level using the proper interface.

Resources Level
For the sake of simplicity in dealing with resources allocation, two classes of users are consider to coexist in the same geographical area in the proposed approach. All the users within the specified geographical area are classified by the broker to fit into one of the classes. The first class is called primary users. This class is composed of users who hold a license issued by a regulatory agency given rights to occupy a pre-determined range of the spectrum of frequencies. The second class is called secondary users and comprehends all those network users who aim to opportunistically access the available network resources.

To help accommodating these two classes of users in the same geographical area, the broker must have knowledge about three types of frequencies [6]:

1) **Exclusive Use:** this kind of frequencies relies on licenses granted by regulatory bodies and is controlled by network operators who hold usage rights for a specified range of frequencies for a defined period of time.
2) **Shared Use:** refers to the range of spectrum frequencies which are license-exempt.
3) **Exclusive Shared Use:** the most recent model of spectrum access is the basis for the so called LSA regime.

The Resources level of the broker is responsible for providing information regarding the users currently operating in the geographical area as well as about the available ranges of frequencies of each type. Therefore, this level implements three databases which are often fed by the Update Level and feed the Decision Level with information about the resources allocation status.

A Primary Users database is specified to store regulatory information regarding the exclusive usage rights afforded to license holders. The Secondary Users database allows the broker to register opportunistic and license-exempt network operators. Finally, a LSA Pool database is defined to store information about Exclusive Shared Access frequencies.

Decision Level
Requests for resources renting are received and processed by the Decision Level of the multilevel broker. The Resource

Request contains all information demanded by the broker to decide which resources will be designated for sharing, taking into account the QoS requirements and the cost. A structure is defined to format such requests: [Operator ID, CoS, Delay, Jitter, Thoughput, Duration, Priority]

Two fields defined in the resources request structure deserve further explanation. The Class of Service (CoS) field is 3 bits long to support three classes defined in the proposed architecture, plus one bit reserved for future use. These classes are specified as follows:

- **001 - Real Time Services (RTS):** supports delay and jitter sensitive real time transmissions.
- **010 - Multimedia Services (MS):** comprehends real time services with high throughput but no strict delay and jitter.
- **011 - Best Effort Services (BES):** designed to support best effort transmissions without strict QoS requirements.

The second field of the resources request structure that deserves special attention is Priority. The priority of a request is defined by the resources renter and is related to the amount of investment that such network operator is willing to make in order to rent resources. A high priority indicates that the network operator is able to rent more expensive resources than in a low priority situation. This field was defined to be 3 bits long to allow the setup of three values of priority currently defined in the architecture, but also to support future enhancements. The currently defined levels of priority are the following:

- **001 - High Priority:** all the available network operators and the LSA pool of frequencies are considered.
- **010 - Medium Priority:** will not take into account the more expensive network resources.
- **011 - Low Priority:** focuses on finding cheap resources options for renting.

Table II summarizes the features of each class of service and the corresponding priorities.

TABLE II

CLASSES OF SERVICES OF THE PROPOSED ARCHITECTURE

Class of Service	Supported Priorities	Exclusive Use	Shared Use	Exclusive Shared
RTS	High/Medium	X	X	X
MS	all		X	X
BES	Low/Medium		X	

Every time a resources request is received, it is processed by a Resources Controller. This entity of the broker has direct access to the Tickets database, which is responsible for controlling the cost of the resources sharing transactions. Through the proper interface it is also able to retrieve information from the databases in the Resources Level of the broker. The aim of the resources level is to obtain updated knowledge about the network resources status and feed the Provisioning algorithm with possible resources servers for a given request. Towards this aim, the execution of the Resources Controller follows the specification of algorithm 1.

The input of the Resources Controller algorithm is a resource request. Such algorithm interfaces with the Resources Level and therefore is able to access the Primary User, Secondary User, and LSA databases. In the first stage, the algorithm classifies the resource request according to the priority informed by the requesting operator considering the class of service (as defined in Table II). The function called $get\_mno(< Type\ of\ Resource >, < QoS\ Parameters >)$ is responsible for searching the databases of Resources Level to retrieve candidate resource providers which have enough resources to guarantee QoS. This retrieval of information takes into account the restrictions imposed by QoS parameters specified in the resources request, *i.e.* maximum delay, maximum jitter, and minimum throughput.

Algorithm 1 Resources Controller

Require: r ▷ A struct containing a resource request
Require: get_mno([databases], [QoS Requirements])
$p \leftarrow r.Priority$; $d \leftarrow r.Delay$; $j \leftarrow r.Jitter$; $t \leftarrow r.Throughput$
switch p **do**
    case High
$mno \leftarrow get\_mno([Primary, Secondary, LSA], [d, j, t])$
    case Medium
$mno \leftarrow get\_mno([Secondary, LSA], [d, j, t])$
    case Low $mno \leftarrow get\_mno([Secondary], [d, j, t])$
for all mno **do**
    $cost(i) \leftarrow [mno.Id, mno.Tickets]$
if $cost = \varnothing$ & $p =$ High **then**
    return 0
else if $cost = \varnothing$ & $p =$ Medium **then**
    $mno \leftarrow get\_mno([Primary], [d, j, t])$
    for all mno **do**
        $cost(i) \leftarrow [mno.Id, mno.Tickets]$
    if $cost = \varnothing$ **then return** 0
else if $cost = \varnothing$ & $p =$ Low **then**
    $mno \leftarrow get\_mno([Primary, Secondary], [d, j, t])$
    for all mno **do**
        $cost(i) \leftarrow [mno.Id, mno.Tickets]$
    if $cost = \varnothing$ **then return** 0
return provisioning($cost$)

After accessing the Resources Level databases, the algorithm calculates the cost of each resource available. The cost (ζ) follows the model of favors exchanged among resources providers. The cost of each favor is influenced by three main factors: (I) the type of service provider (ρ), (II) the amount of resources currently compromised by the selected resources provider (ℓ) at a given instant of time, and (III) the priority of the request (κ). ζ is calculated using (2).

$$\zeta = \rho.\kappa.\left(\frac{\ell_{RTS} + \ell_{MS} + \ell_{BES}}{L}\right) \tag{2}$$

In this equation, L represents the total amount of currently unused resources in a given resources provider. The values related to the priorities and types of service providers are summarized in Table III.

TABLE III

PARAMETERS USED TO CALCULATE THE COST

ρ		κ	
Type of Provider	**Value**	**Priority**	**Value**
Shared Use	1	Low	1
Exclusive Shared Use	2	Medium	2
Exclusive Use	3	High	3

It is important to highlight that the Broker estimates the initial cost without considering the duration of the loan, since this information is not accurate at this first stage of analysis, because the expected duration may differ from the real duration of a transaction in a realistic scenario. Therefore, after the transaction if finished, the initial cost is multiplied by the duration of the loan. Since the duration of a sharing transaction is computed in unit of hours by the broker, the final price of the favor, as a consequence, will be computed in an unit of tickets per hour. To guarantee fairness in the resources sharing transactions among network operators, the favor will be registered by the broker considering its final cost.

The Resources Controller algorithm generates an array of candidate resources providers. Each entry of the array is composed of the unique identification of the service provider and the cost of this transaction. The resulting array is used as the input to the Resources Provisioning algorithm. The aim of the Resources Provisioning algorithm is to take a decision on which resource providers is the best to serve a specific request.

B. Resources Assessment Model

The approach defined in the broker demands an accurate assessment of the amount of resources controlled by each operator. For the sake of simplicity, it is assumed that all operators, including the secondary users, are able to access pre-defined spectrum bands as their main resource. This assumption is close to reality, since this kind of allocation is standard for LTE-Advanced and LSA regime. In this situation, the spectrum frequency is always available to the network operator who is responsible for managing the access of the clients to the spectrum of frequency. The amount of resources is then correlated with the transmission capacity of each network operator.

The capacity is modeled considering the Shannon's model, as defined in (3). The channel bandwidth (B) is considered to calculate the theoretical channel capacity (C), which is the main resource shared in the proposed architecture.

$$C = B \log_2 \left(1 + \frac{P.g}{\sigma^2}\right) * \eta \tag{3}$$

In (3), P represents the transmission power, g is the gain provided by the transmitting antenna, and σ^2 is the noise power. Besides the SINR, the link efficiency (η) is considered to model a more realistic scenario.

The resources demand in a given instant of time ($d(t)$) takes into account the individual demand ($d_i(t)$) of the *ith* active connection of each network operator. The total number of active connections is represented by n. Moreover, the overhead, caused by both cyclic prefix insertion (ϑ_{CP}) and pilot subcarriers used for synchronization (ϑ_{PS}) is considered. Therefore, $d(t)$ is calculated as defined in 4.

$$d(t) = \left(\sum_{i=1}^{n} d_i(t)\right) + \vartheta_{CP} + \vartheta_{PS} \tag{4}$$

To simplify the decision process implemented by the Decision Level of the broker, the resources occupation factor ($\delta(t)$) in a given instant of time is calculated using (5). It is important to highlight that this equation correlates the current demand ($d(t)$) with the capacity of a network operator (C). The demand is originally calculated in unit of Mb, while the capacity is obtained in terms of Mbps. Therefore, to guarantee the consistency of $\delta(t)$ factor, the demand must be observed during the period of one second, to transform its unit into Mbps before applying the equation.

$$\delta(t) = \frac{\left(\sum_{i=1}^{n} d_i(t)\right) + \vartheta_{CP} + \vartheta_{PS}}{B \log_2 \left(1 + \frac{P.g}{\sigma^2}\right) * \eta} \tag{5}$$

C. Inter-Network Communication and Signaling

The specification of the signaling protocol is presented in Fig. 2. The protocol is based on the approach of the LTE-Advanced standard published by the 3GPP.

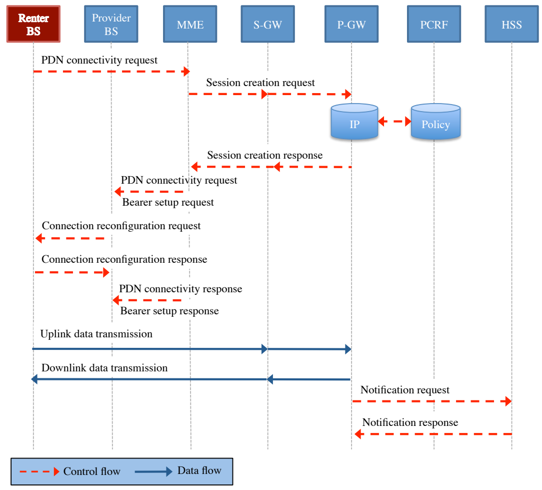


Fig. 2. Inter-network Signaling

Two paths are represented in Fig. 2 (i) control information (dashed line) and (ii) data flow (continuous line). The renter starts the process by communicating with a BS to dynamically establish a SLA to access shared resources, by sending a Packet Data Network (PDN) connectivity request to the resources provider. This message is automatically forwarded to the Mobility Management Entity (MME), which controls all signaling between the devices within the resources provider and its core network. To provide such control, MME receives information from a Home Subscriber Service (HSS), which holds information about authorized network users, as QoS profiles, roaming restrictions, and PDN that can be accessed by a given device. In the sequence, MME demands to Serving Gateway (S-GW) the creation of a transmission session.

This session is only created after Policy Control and Charging Rules Function (PCRF) verifies network, SLA, and QoS policies, and PDN Gateway (P-GW) provides IP connectivity. After receiving a successful response, MME sets the bearer up, allowing the renter to access the shared network resources. Finally, MME informs HSS about the new communication. The data transmissions pass through S-GW, which is responsible for controlling mobility of devices between different BSs as well as for administrative tasks such as collecting information for charging purposes, SLA compliance verification, and lawful inspections.

IV. Performance Evaluation

In this section, the performance of the proposed solution is evaluated considering a case study where four LTE-Advanced network operators coexist in the same geographical area, so unused resources from a given network operator can be opportunistically leased to devices belonging to other network operators. Towards the performance evaluation, in IV-A the simulation scenario is explained. Then, in IV-B the results obtained via Matlab simulations are presented and discussed.

A. Simulation Scenario

To simulate the behavior of the proposed architecture it is mandatory to properly model the traffic demands of the users. The traffic model must consider the connection arrival and the amount of traffic demanded per connection. The *System Evaluation Methodology* document, published by the WiMAX Forum [7] was selected to model the traffic because it is based on realistic measurements and provides a solid base to estimate the actual traffic demanded by the different users. In the simulations three different kinds of traffic are considered: HTTP (60% of the total traffic), VoIP (20%), and Video clip streaming (20%).

The first kind of traffic models best effort HTTP packets. The transmissions are composed of a main page, which has a given number of embedded objects, such as images, scripts, and other sorts of attached files. After requesting and receiving the files, the browser parses the page to make it readable to the user. The user then reads the page before making a new request. The values of each phase of the HTTP statistical model are presented in Table IV.

TABLE IV
HTTP TRAFFIC PARAMETERS

Component	Distribution	Parameters	PDF
Main page size	Truncated Lognormal	Mean = 10710 bytes SD = 25032 bytes Min = 100 bytes Max = 2 Mbytes	$\sigma = 1.37$ $\mu = 8.37$
Embedded object size	Truncated Lognormal	Mean = 7758 bytes SD = 126168 bytes Min = 50 bytes Max = 2 Mbytes	$\sigma = 2.36$ $\mu = 6.17$
Number of embedded objects	Truncated Pareto	Mean = 5.64 Max = 53	$\sigma = 1.1$ $\mu = 55$
Reading time	Exponential	Mean = 30 s	$\mu = 0.033$
Parsing time	Exponential	Mean = 0.13 s	$\mu = 7.69$

VoIP transmissions are modeled according to the parameters of Adaptive Multi Rate (AMR) codec, which presents ON/OFF behavior. The duration of each period is modeled using an exponential distribution with mean of 1026 ms of conversations (ON period) and 1171 ms of silence (OFF period). A Packet Data Unit (PDU) is generated every 20 ms.

The third traffic model considers the streaming of video clips encoded with MPEG-4. Each of the videos has variable length, varying from 15 s to 60 s. The display size of the video clip is 176x144, what leads to a mean frame size of 2725 Kbytes after the video clip is compressed.

Using the aforementioned traffic models, simulations were run on Matlab to evaluate the performance of the proposed architecture. The frame duration is 10 ms and the transmissions are carried out in a 10 MHz wireless channel. The number of connections was varied to evaluate the performance of the architecture in different traffic load scenarios.

B. Performance Evaluation

The performance evaluation is presented based on the Goals/Questions/Metrics (GQM) model, in order to clarify how the objectives and contributions of the paper are addressed in this section. Fig. 3 illustrates the three main goals of the paper, the corresponding questions, and the metrics used to answer each question.

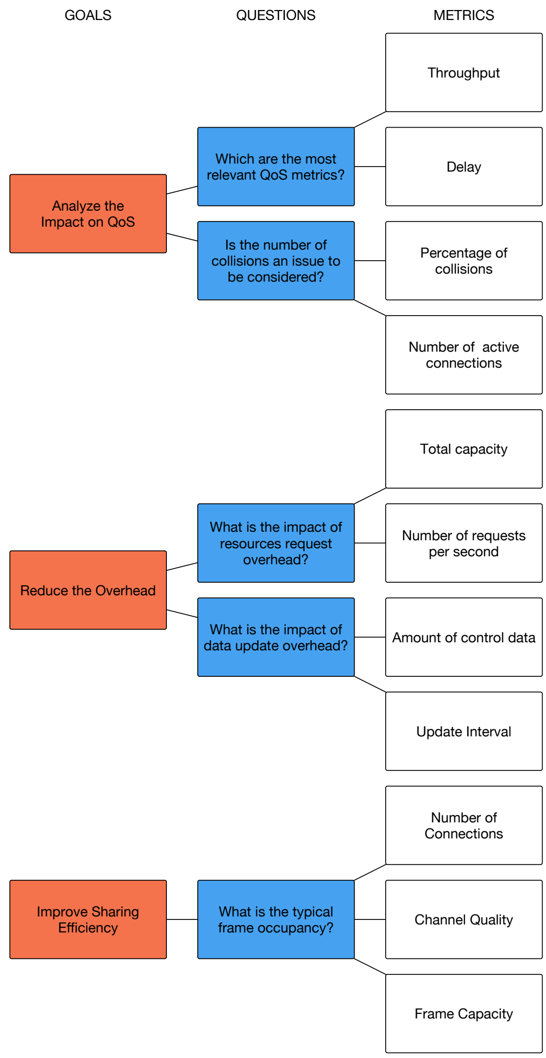


Fig. 3. GQM model of the proposed solution

The performance of the proposed multilevel broker was also compared with two different resource allocation algorithms found in the Literature. Both algorithms were originally analyzed by Gardellin *et al.* [2] and are called Random Channel Allocation and Non-cooperative Channel Allocation. Both approaches consider that the shared resource is channel capacity and are applied to IEEE 802.22 networks but are general enough to be adapted to other network scenarios, such as the one analyzed in this paper.

The first goal of the paper is to analyze the impact of the implementation of the proposed multilevel broker on the QoS offered by the network operators to the customers. The metrics considered in this evaluation are throughput and delay, since delay is very important to RTS and MS and a high throughput is desirable for all services. Fig. 4, illustrates the throughput of the LTE-Advanced network in conditions in which the number of connections is varied.

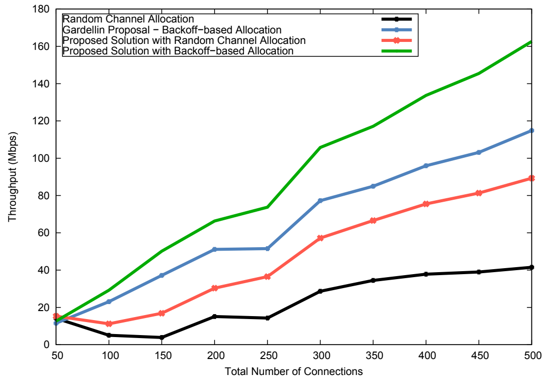


Fig. 4. Throughput behavior with variable number of connections

As can be observed in Fig. 4, the random channel allocation is the worst case scenario. This behavior is explained due to the nature of a random algorithm. Since no channel access control is implemented in this case, an uncontrolled number of collisions may occur, leading to a decrease in the overall throughput, what will impact negatively on the QoS. The throughput significantly improves when LTE-Advanced resources sharing is implemented. In this case, it is important to highlight that although the channel allocation is random, when LTE-Advanced resources are demanded, a SLA is established, what leads to more efficient resources allocation. A similar behavior is observed when the backoff-based cooperative channel allocation algorithm is analyzed. However, in this case, the nature of the algorithm reduces the number of collisions and consequently improves the network throughput. The gains provided by the proposed approach surpass those obtained by Gardellin *et al.* proposal by up to 28%. Therefore it shows to be the most effective solution to improve the throughput QoS metric.

The second QoS metric that is taken into account is the average delay, which is compared with related work in Fig. 5. The average delay is affected by the number of active connections and consequently by the overall network demand. In this scenario, the delay of MS and RTS classes of services is measured considering a Weighted Fair Queuing (WFQ) scheduling algorithm. Results show that the implementation of the proposed architecture guarantees a reduction of more than 50% on the average delay in situations where the network is saturated.

Another important aspect related to QoS provisioning is to analyze the amount of collisions, since this metric affects all the QoS metrics. In Fig. 6, the percentage of collisions is related with the number of active connections. It is clear that the amount of collisions increases with the number of connections and stabilizes after a certain amount of connections, what is due to the characteristics of the Connection Admission Control (CAC) implemented by the network operators to guarantee a certain QoS level. However, when the proposed approach is implemented, the amount of collisions, in the worst case is less than the half of the amount of collisions observed in the related approaches.

The second goal of the proposed solution is to keep the

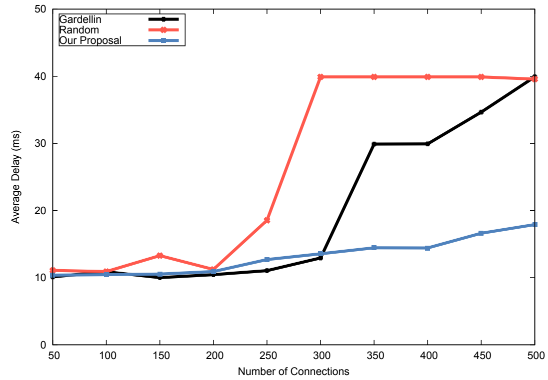


Fig. 5. Delay Comparison

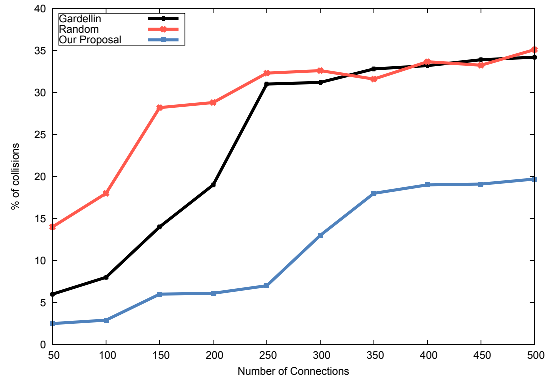


Fig. 6. Percentage of collisions with variable number of connections

overhead low. Two kinds of control data are considered in this analysis: (I) information exchanged by the polling/reply mechanism implemented by the update level of the broker and (II) information exchanged by the opportunistic users to request resources to the decision level of the broker. The first type of control data is analyzed in Fig. 7. In this figure, the overhead is presented considering update intervals varying from 1 minute to 1 hour. This range of intervals allows the analysis of very aggressive update strategies and of more conservative ones.

The amount of network operators that are sharing the same geographical area is also varied in the simulations. The overhead generated by the presence of only one operator, although not realistic, is evaluated as a baseline for comparison. Four operators are simulated in order to evaluate the overhead of the architecture in a common situation where four LTE-Advanced operators share the same geographical area, following the allocation model adopted by the Brazilian government. The coexistence of six operators was analyzed since it is one of the main scenarios in which the proposed architecture can be applied. Finally, the overhead generated during the process of resources sharing among ten network operators is analyzed as a worst case scenario.

The behavior of the update overhead follows the expectation that it may reduce as the interval between updates increase. Moreover, the amount of network operators sending updates directly affects the amount of control data that is generated during an update process. In terms of network load generated by the control data, it is possible to conclude that the update overhead can be neglected by the network operators. This is

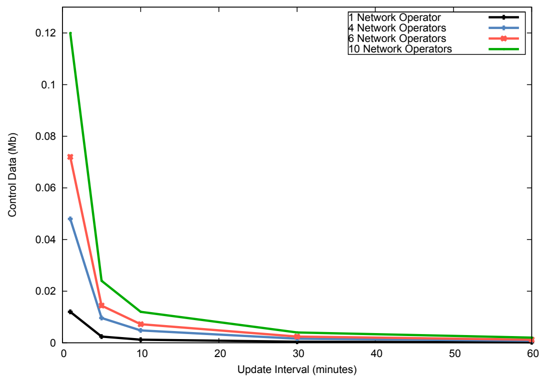


Fig. 7. Update Overhead

justified because, in a worst case scenario, where 10 network operators are updating the broker every minute, only about 0.12 Mbps of traffic is generated. Considering that in good wireless channel conditions, each network operator may transmit up to about 120 Mbps, this value is very small, corresponding to 0.001% of the whole traffic.

Another kind of overhead generated by the proposed architecture is correlated to the request of resources, which is necessary every time a new resources sharing operation is about to begin. The behavior of this overhead is shown in Fig. 8. In the graph, the percentage of a given operator network capacity used for transmission of resources requests is analyzed in relation to the amount of requests received per second.

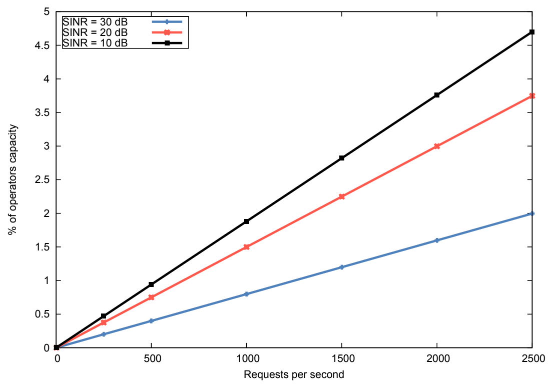


Fig. 8. Request Overhead

Results show that even when a very large amount of request messages is exchanged, less then 5% of the operator's resources are compromised for control data transmission. In a realistic situation, the amount of requests per second should lead to an occupation of less then 1% of the network resources. These results show that the request overhead can also be neglected by the network operators. Therefore, by analyzing the overhead generated by the proposed architecture, it is possible to conclude that the amount of control information that is exchanged among the broker and the network operators does not affect the overall performance of the network.

The third goal of the proposed approach is to improve the sharing efficiency. The accomplishment of this goal is measured by the frame occupancy. In Fig. 9, the amount of underutilized resources of a traditional, non cooperative LTE-Advanced network is analyzed. The occupation of the LTE-Advanced frame is normalized to the maximum capacity, already taking into account the overhead, calculated using (3). The total number of connections is also considered.

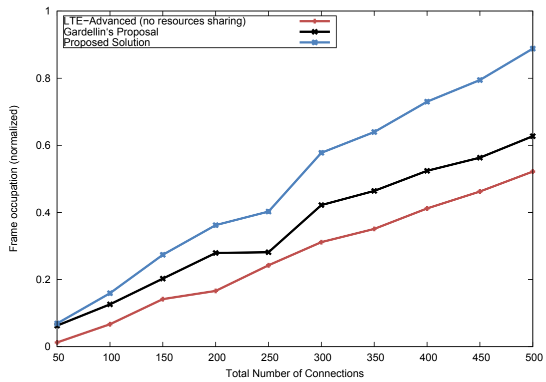


Fig. 9. Normalized Frame Occupation. Maximum capacity of 113Mbps

In a situation with no resources sharing, portions of more than 50% of the frame are underutilized. In situations in which resources sharing is implemented, the underutilization can be reduced to as low as 15% in the considered traffic demand scenarios. These results demonstrate that the concept of resources sharing is feasible for the scenario analyzed in this paper.

V. Conclusions and Future Work

This paper dealt with the spectrum scarcity problem in multi-operator wireless cellular networks. The results showed that the implementation of a multilevel broker that allows resources sharing improves the QoS provided to the network customers in comparison with two solutions found in the literature.

Directions for future work include the design novel provisioning algorithms at the Decision Level of the broker. Enhancements in such algorithm could improve even more the QoS delivered to the resources renter.

References

[1] M. Guizani, B. Khalfi, M. Ben Ghorbel, and B. Hamdaoui, "Large-scale cognitive cellular systems: resource management overview," *Communications Magazine, IEEE*, vol. 53, no. 5, pp. 44–51, May 2015.

[2] V. Gardellin, S. K. Das, and L. Lenzini, "Self-coexistence in cellular cognitive radio networks based on the IEEE 802.22 standard," *Wireless Communications, IEEE*, vol. 20, no. 2, 2013.

[3] S. Dixit, S. Periyalwar, and H. Yanikomeroglu, "Secondary User Access in LTE Architecture Based on a Base-Station-Centric Framework With Dynamic Pricing," *Vehicular Technology, IEEE Transactions on*, vol. 62, no. 1, pp. 284–296, 2013.

[4] K. Chatzikokolakis, P. Spapis, A. Kaloxylos, and N. Alonistioti, "Toward spectrum sharing: opportunities and technical enablers," *Communications Magazine, IEEE*, vol. 53, no. 7, pp. 26–33, July 2015.

[5] B. Singh, S. Hailu, K. Koufos, A. Dowhuszko, O. Tirkkonen, R. JaÌĹntti, and R. Berry, "Coordination protocol for inter-operator spectrum sharing in co-primary 5g small cell networks," *Communications Magazine, IEEE*, vol. 53, no. 7, pp. 34–40, July 2015.

[6] UMTS Forum, "Rspg opinion on licensed shared access," May 2013, SPG13-52.

[7] WiMAX Forum, "WiMAX system evaluation methodology version 2.1," July 2008, WiMAX Forum. in http://www.wimaxforum.org/documents. Accessed in March 2014.

APPENDIX D PUBLISHED PAPER – IEEE LCN 2016

The current static model used for allocating the spectrum of frequencies and the increasing demand for network resources imposed by modern applications and services may lead to a resources scarcity problem. Dealing with this problem demands optimized resources allocation. An alternative to provide this optimization is to allow resources sharing among heterogeneous network operators which may implement different technologies and kinds of resources. To allow and control the sharing of resources in heterogeneous network scenarios, in this paper, an architecture is presented to encourage network operators to share their underutilized resources. A multilevel broker is proposed to control the resources sharing. This broker dynamically establishes a service level agreement that takes into account the quality of service requirements of resources renters. A performance evaluation shows that the implementation of the proposed architectures leads to better allocation of underutilized network resources compared to two algorithms found in the literature.

- **Title –**
  A Resources Sharing Architecture for Heterogeneous Wireless Cellular Networks
- **Conference –**
  IEEE 41st Conference on Local Computer Networks
- **Date –**
  November, 2016
- **Held at –**
  Dubai, UAE

A Resources Sharing Architecture for Heterogeneous Wireless Cellular Networks

Rafael Kunst, Leandro Avila, Edison Pignaton, Sergio Bampi, Juergen Rochol
Federal University of Rio Grande do Sul (UFRGS)
Email: {rkunst, laavila, edison.pignaton, bampi, juergen}@inf.ufrgs.br

***Abstract*—The current static model used for allocating the spectrum of frequencies and the increasing demand for network resources imposed by modern applications and services may lead to a resources scarcity problem. Dealing with this problem demands optimized resources allocation. An alternative to provide this optimization is by allowing resources sharing among network operators. In this paper, an architecture is presented to encourage network operators to share their underutilized resources. A multilevel broker is proposed to control the resources sharing. This broker dynamically establishes a service level agreement that takes into account the quality of service requirements of resources renters and the cost of the resources. A performance evaluation shows that the implementation of the proposed architectures leads to better allocation of underutilized network resources.**

I. Introduction

The current spectrum allocation model prioritizes main network operators who purchase exclusive rights over portions of the spectrum of frequencies. This model makes it challenging to unlicensed network operators to provide high quality of service (QoS) to their users. Such challenge increases in the current traffic demand scenario, where network devices demand significant amounts of both spectrum and capacity resources to manage the traffic generated by both wireless applications and services. The described scenario may lead to a network resources scarcity problem in the near future.

The main goal of this paper is to present a solution to deal with this network resources scarcity problem. The proposed solution is based on the premise that both commercial and non-commercial network operators have underutilized resources during periods, especially in off-peak hours. The main idea is to allow these network operators to improve their profits by leasing the underutilized resources to opportunistic network operators that will be able to improve the QoS offered to their users.

Network resources can be shared according to two main approaches. The first one is called Collective Use of Spectrum (CUS) because specific licenses are not demanded to allow devices access to network resources. On the other hand, there are scenarios in which licenses are necessary to provide the devices with network resources access. In these scenarios, devices are submitted to a licensed access regime before accessing the resources, which is called Licensed Shared Access (LSA) [1].

Different approaches have been considered to implement resources sharing both in CUS and LSA regimes ([2], [3], [4], [5], [6], [7]). Most of these proposals involve the implementation of complex algorithms aiming to mitigate interference or to find spectrum access opportunities using artificial intelligence techniques. Although very relevant, these solutions generally identify the need for additional resources and search for available resources without taking into account the QoS requirements of the resources renter and the cost of these resources. This common approach of related works may lead to the allocation of network resources that are not suitable for the users. Considering the aspects not fully covered by the related work, in this paper, an architecture is presented to allow resources sharing among different network operators. The most important component of the proposed approach is a multilevel resources broker, which is designed to dynamically establish a service level agreement that takes into account the QoS requirements of resources renters.

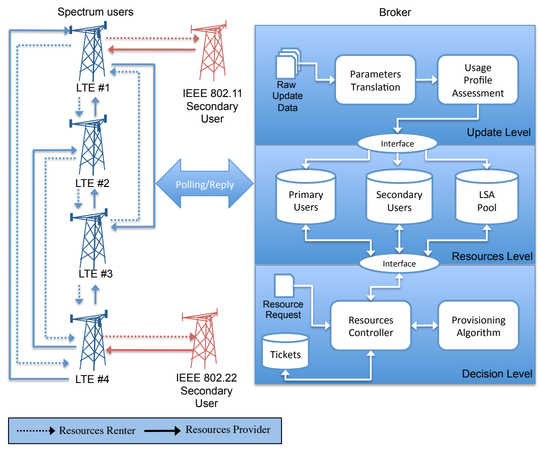


Fig. 1. Architecture Design

II. Resources Sharing Architecture

The design of the proposed architecture is presented in Fig. 1. The architecture is divided in two parts which communicate through a polling/reply mechanism. In the left side of the illustration, the spectrum users coexist in a geographical area considering a scenario that allows one operator to communicate with all neighboring network operators. In the right side of the figure, the structure of a broker is represented. This broker is responsible for coordinating resources sharing among the spectrum users.

A novel multilevel broker is proposed to play the role of the centralized entity in the resources sharing architecture. Three levels were defined to provide independent and simultaneous control of different tasks of spectrum sharing management. These levels were named accordingly to the function executed by each one: (I) Update Level, (II) Resources Level, and (III) Decision Level. These levels are interconnected by interfaces which implement the flow of information.

The **Update Level** is responsible for collecting operation parameters from the network operators which participate in the spectrum sharing initiative. The updating mechanism is based on the implementation of a polling-based technique, which is controlled by the Parameters Translation block of the broker. This block allows the configuration of the interval between polls and translation of the raw update data (operator ID, delay, jitter, and throughput data) into useful information to allow the architecture to take proper decisions regarding to resources sharing.

The pre-processed raw data is sent to the Usage Profile Assessment block, which applies the concepts of cognition to keep track of the historical information provided by the network operators. This historical information is taken into account to define the current usage profile of the network to minimize the effect of abnormal behaviors of the traffic that may occur in realistic operation scenarios. The weight given to the historical information (α) and the weight considered to the most recent update ($1-\alpha$) are parameters of this block. Equation 1 is applied to calculate the weighted load (ℓ) of each considered QoS parameter of the network operators.

$$\ell = \alpha.\sum_{i=1}^{n} \ell_{(t-i)} + (1-\alpha).\ell_{(t)} \quad (1)$$

This equation considers a pre-defined number (n) of historical evaluations of ℓ and performs an exponential smoothing to obtain the weighted load of a given QoS parameter. The same equation is applied to the remaining QoS parameter in order to obtain the complete assessment of the usage profile of each network operator.

The processed usage profile information is then sent to the **Resources Level** using the proper interface. For the sake of simplicity in dealing with resources allocation, two classes of users are considered to coexist in the same geographical area in the proposed approach. The first class, called primary users, is composed of users who hold a license issued by a regulatory agency concession to occupy a pre-determined range of the spectrum of frequencies. The second class is called secondary users and comprehends all those network users who aim to opportunistically access the available network resources. To help accommodating these two classes of users in the same geographical area, the broker must have knowledge about exclusive use, shared use, and exclusive shared use frequencies [8].

The Resources level of the broker is responsible for providing information regarding the users currently operating in the geographical area as well as about the available ranges of frequencies of each type. Therefore, this level implements three databases which are often fed by the Update Level and provide the Decision Level with information about the resource allocation status, called Primary Users, Secondary Users, and LSA Pool.

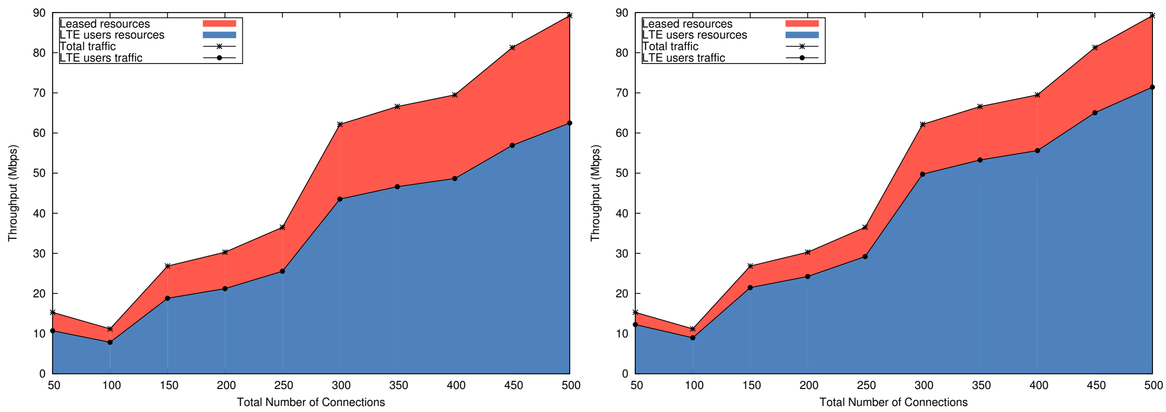
Request for resources renting are received and processed by the **Decision Level** of the multilevel broker. The Resource Request contains all information demanded by the broker to decide which resources will be designated for sharing, taking into account the QoS requirements (*i.e.* class of service, delay, jitter, and throughput requirements), the cost, the estimated duration of the lease, and the priority of the request. Three levels of priority and three classes of service are defined: Real Time Services (RTS), Multimedia Services (MS), and Best Effort Services (BES).

Every time a resource request is received, it is processed by a Resources Controller. This entity of the broker has direct access to the Tickets database, which is responsible for controlling the cost of the resources sharing transactions. Through the proper interface it is also able to retrieve information from the databases in the Resources Level of the broker. The aim of the resources level is to obtain updated knowledge about the network resources status and feed the Provisioning algorithm with possible resources servers for a given request. Towards this aim, the execution of the Resources Controller follows the specification of algorithm 1.

Algorithm 1 Resources Controller

Require: r ▷ A struct containing a resource request
Require: get_mno([databases], [QoS Requirements])
$p \leftarrow r.Priority$; $d \leftarrow r.Delay$; $j \leftarrow r.Jitter$; $t \leftarrow r.Throughput$
switch p **do**
  case High
$mno \leftarrow get\_mno([Primary, Secondary, LSA], [d, j, t])$
  case Medium
$mno \leftarrow get\_mno([Secondary, LSA], [d, j, t])$
  case Low $mno \leftarrow get\_mno([Secondary], [d, j, t])$
for all mno **do**
  $cost(i) \leftarrow [mno.Id, mno.Tickets]$
if $cost = \varnothing$ & $p =$ High **then**
  return 0
else if $cost = \varnothing$ & $p =$ Medium **then**
  $mno \leftarrow get\_mno([Primary], [d, j, t])$
  for all mno **do**
    $cost(i) \leftarrow [mno.Id, mno.Tickets]$
  if $cost = \varnothing$ **then return** 0
else if $cost = \varnothing$ & $p =$ Low **then**
  $mno \leftarrow get\_mno([Primary, Secondary], [d, j, t])$
  for all mno **do**
    $cost(i) \leftarrow [mno.Id, mno.Tickets]$
  if $cost = \varnothing$ **then return** 0
return provisioning($cost$)

The input of the Resources Controller algorithm is a resource request. In the first stage, the algorithm classifies the resource request according to the priority and the class of service informed by the requesting operator. The function called $get\_mno(< Type\ of\ Resource >, < QoS\ Parameters >)$ is responsible for searching the databases of Resources Level to retrieve candidate resource providers which have enough resources to guarantee QoS. After accessing the Resources Level databases, the algorithm calculates the cost of each resource available. The cost (ζ) follows the model of favors exchanged among resources providers. The cost of each favor is influenced by three main factors: (I) the type of service provider (ρ), (II) the amount of resources currently compromised by the selected resources provider (ℓ) at a given instant of time, and (III) the priority of the request (κ). ζ is calculated



(a) LTE traffic load of 70%

(b) LTE traffic load of 80%

Fig. 2. Underutilized resources considering a total frame capacity of 113Mbps

using (2).

$$\zeta = \rho.\kappa.\left(\frac{\ell_{RTS} + \ell_{MS} + \ell_{BES}}{L}\right) \quad (2)$$

In this equation, L represents the total amount of currently unused resources in a given resources provider. The values related to the priorities and types of service providers are summarized in Table I.

TABLE I
PARAMETERS USED TO CALCULATE THE COST

ρ		κ	
Type of Provider	**Value**	**Priority**	**Value**
Shared Use	1	Low	1
Exclusive Shared Use	2	Medium	2
Exclusive Use	3	High	3

The Resources Controller algorithm generates an array of candidate resources providers. Each entry of the array is composed of the unique identification of the service provider and the cost of this transaction. The resulting array is used as the input to a Resources Provisioning algorithm, which takes an optimized decision on which resource provider is the best to serve a specific request.

In order to deal with the resource request, it is fundamental to accurately assess the amount of resources (*i.e.* network capacity) controlled by each operator. The capacity is modeled according to Shannon's model. The channel bandwidth (B) is considered to calculate the theoretical channel capacity (C). In the proposed equation, P represents the transmission power, g is the gain provided by the transmitting antenna, and σ^2 is the noise power. The link efficiency (η) is considered to model a more realistic scenario. The resources demand in a given instant of time ($d(t)$) takes into account the individual demand ($d_i(t)$) of the *ith* active connection of each network operator. The total number of active connections is represented by n. Moreover, the overhead, caused by both cyclic prefix insertion (ϑ_{CP}) and pilot subcarriers used for synchronization (ϑ_{PS}) is considered.

To simplify the decision process implemented by the Decision Level of the broker, a resources occupation factor $\delta(t) = (d(t)/C)$ is defined. It is important to consider that this equation correlates the current demand ($d(t)$) with the capacity of a network operator (C). The demand is originally calculated in unit of Mb, while the capacity is obtained in terms of Mbps. Therefore, to guarantee the consistency of $\delta(t)$ factor, the demand must be observed during the period of one second, to transform its unit into Mbps before applying the equation. $\delta(t)$ follows equation 3.

$$\delta(t) = \frac{\left(\sum_{i=1}^{n} d_i(t)\right) + \vartheta_{CP} + \vartheta_{PS}}{B\log_2\left(1 + \frac{P.g}{\sigma^2}\right) * \eta} \quad (3)$$

III. PERFORMANCE EVALUATION

In this section, the performance of the proposed solution is evaluated considering a case study where four LTE-Advanced network operators coexist in the same geographical area. To simulate the behavior of the proposed architecture it is mandatory to properly model the traffic demands of the users. The traffic model must consider the connection arrival and the amount of traffic demanded per connection. The *System Evaluation Methodology* document, published by the WiMAX Forum [9] was selected to model the traffic because it is based on realistic measurements and provides a solid base to estimate the actual traffic demanded by the different users. In the simulations three different kinds of traffic are considered: HTTP (60%), VoIP (20%), and Video (20%).

The first kind of results show the amount of resources used by LTE-Advanced primary users and consequently, the amount of underutilized resources, which may be shared. This analysis takes into account two scenarios to represent different loads in the network managed by a LTE-Advanced operator. The results presented in Fig. 2 show the amount of resources that may be shared in an inter-networking scenario that considers a theoretical capacity of 113Mbps. The traffic load of LTE-Advanced network was also varied to consider a total load of 70% in Fig. 2 (a) and a total load of 80% in Fig. 2 (b).

As can be seen in Fig. 2, the amount of resources allocated to LTE-Advanced primary users varies from 9% to about 54% of the network throughput when the proposed resources sharing architecture is not implemented. Results obtained after the proposed architecture is implemented show that up to 25% of the resources can be shared with opportunistic users when the LTE-Advanced traffic load is 70%. When this load is increased to 80%, the gain is still observed, reaching values of up to 18% of the resources which may be shared with opportunistic users. This leads to advantages to both network operators, since underutilized resources of a given network operator can be used by another operator.

In Fig. 3, the amount of underutilized resources of a traditional, non cooperative LTE-Advanced network is analyzed. The occupation of the LTE-Advanced frame is normalized to the maximum capacity, already taking into account the overhead. The total number of connections is also considered. In a situation with no resources sharing, portions of more than 50% of the frame are underutilized. In situations in which resources sharing is implemented, the underutilization can be reduced as low as 15% in the considered traffic demand scenarios. These results demonstrate that the concept of resources sharing is feasible for the scenario analyzed in this paper.

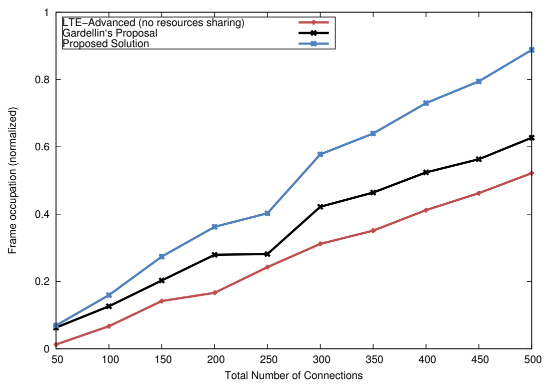


Fig. 3. Normalized Frame Occupation. Maximum capacity of 113Mbps

The performance of the proposed architecture was also compared with two different resource allocation algorithms found in the Literature. Both algorithms were originally analyzed by Gardellin *et al.* [2] and are called Random Channel Allocation and Non-cooperative Channel Allocation. Both approaches consider that the shared resource is the channel capacity and are applied to IEEE 802.22 networks but are general enough to be adapted to other network scenarios, such as the one analyzed in this paper. The evaluation scenario presented in Fig. 4, is related to the throughput supported by the LTE-Advanced network under conditions in which the number of connections is varied.

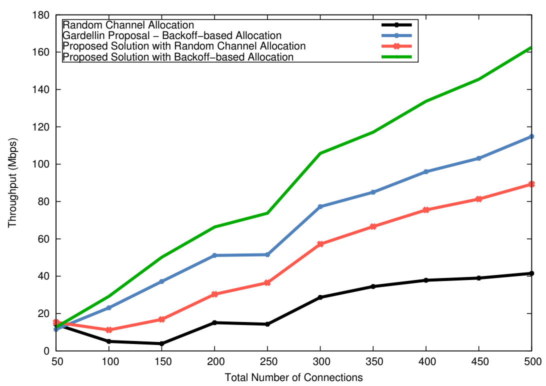


Fig. 4. Throughput behavior with variable number of connections

As can be observed in Fig. 4, the random channel allocation is the worst case scenario. This behavior is explained due to the nature of a random algorithm. Since no channel access control is implemented in this case, an uncontrolled number of collisions may occur, leading to a decrease in the overall throughput, which will impact negatively on the QoS. Results showed that up to 30% of the total number of transmissions typically result in collisions when this kind of allocation is applied. The throughput significantly improves when LTE-Advanced resources sharing is implemented. In this case, it is important to highlight that although the channel allocation is random, when LTE-Advanced resources are demanded, a SLA is established, what leads to more efficient resources allocation. A similar behavior is observed when the backoff-based non-cooperative channel allocation algorithm is analyzed. However, in this case, the nature of the algorithm reduces the number of collisions and consequently improves the network throughput. In the non-cooperative algorithm, collisions are significantly reduced, reaching values as low as 2.5% of the total number of transmissions.

IV. Conclusions and Future Work

This paper dealt with the spectrum scarcity problem by proposing a resources sharing architecture for wireless cellular networks. The results show that the implementation of the architecture is feasible, and the LTE-Advanced is able to serve up to 500 connections (60% http, 20% video, 20% VoIP) and yet share up to 25 % of its network capacity opportunistic users. The results were also compared with two solutions found in the literature. The gains surpass the related work by allowing the resources server to improve its resources utilization by as much as 28%. Directions for future work include the analysis of the overhead generated by the proposed architecture. Designing a novel provisioning algorithms at the Decision Level of the broker is also planned. Enhancements in such algorithm could improve the QoS delivered to the resources renters.

References

[1] M. Guizani, B. Khalfi, M. Ben Ghorbel, and B. Hamdaoui, "Large-scale cognitive cellular systems: resource management overview," *Communications Magazine, IEEE*, vol. 53, no. 5, pp. 44–51, May 2015.

[2] V. Gardellin, S. K. Das, and L. Lenzini, "Self-coexistence in cellular cognitive radio networks based on the IEEE 802.22 standard," *Wireless Communications, IEEE*, vol. 20, no. 2, 2013.

[3] S. Dixit, S. Periyalwar, and H. Yanikomeroglu, "Secondary User Access in LTE Architecture Based on a Base-Station-Centric Framework With Dynamic Pricing," *Vehicular Technology, IEEE Transactions on*, vol. 62, no. 1, pp. 284–296, 2013.

[4] X. Costa-Perez, J. Swetina, T. Guo, R. Mahindra, and S. Rangarajan, "Radio access network virtualization for future mobile carrier networks," *Communications Magazine, IEEE*, vol. 51, no. 7, 2013.

[5] M. Sohul, M. Yao, T. Yang, and J. Reed, "Spectrum access system for the citizen broadband radio service," *Communications Magazine, IEEE*, vol. 53, no. 7, pp. 18–25, July 2015.

[6] K. Chatzikokolakis, P. Spapis, A. Kaloxylos, and N. Alonistioti, "Toward spectrum sharing: opportunities and technical enablers," *Communications Magazine, IEEE*, vol. 53, no. 7, pp. 26–33, July 2015.

[7] B. Singh, S. Hailu, K. Koufos, A. Dowhuszko, O. Tirkkonen, R. JaĽntti, and R. Berry, "Coordination protocol for inter-operator spectrum sharing in co-primary 5g small cell networks," *Communications Magazine, IEEE*, vol. 53, no. 7, pp. 34–40, July 2015.

[8] UMTS Forum, "Rspg opinion on licensed shared access," May 2013, SPG13-52.

[9] WiMAX Forum, "WiMAX system evaluation methodology version 2.1," July 2008, WiMAX Forum. in http://www.wimaxforum.org/documents. Accessed in March 2014.

APPENDIX E PUBLISHED PAPER – IEEE ICC 2017

The wireless network traffic is expected to overload the existing licensed spectrum by 2020. One solution to deal with this traffic overload is to access opportunistically in LSA and unlicensed spectrum bands. LSA allows incumbent users to temporarily provide access to its resources. However, licensees must perform traffic steering to vacate the band without causing interference, whenever the incumbent requires. In this paper, a cognitive algorithm is proposed to take in advance decisions regarding to traffic steering routes in unscheduled evacuation scenarios. This solution aims at guaranteeing the QoS and seamless connectivity during traffic steering. A performance evaluation conducted in a scenario composed of one LTE-LSA and three Wi-Fi network operators, demonstrates that the proposed solution fulfills the time required by the unscheduled evacuation as well as guarantees the QoS and seamless connectivity of evacuees.

- **Title –**
  A Cognitive Algorithm for Traffic Steering in LTE-LSA/Wi-Fi Resource Sharing Scenarios
- **Conference –**
  IEEE International Conference on Communications
- **Date –**
  May, 2017
- **Held at –**
  Paris, France

A Cognitive Algorithm for Traffic Steering in LTE-LSA/Wi-Fi Resource Sharing Scenarios

Jean Cerrillo Fernandez, Rafael Kunst, Juergen Rochol
Federal University of Rio Grande do Sul, Brazil
{jecfernandez, rkunst, juergen}@inf.ufrgs.br

***Abstract*—The wireless network traffic is expected to overload the existing licensed spectrum by 2020. One method to deal with this traffic overload is to access unlicensed and shared spectrum bands using an opportunistic approach. Licensed Shared Access (LSA) allows incumbent users to provide temporary access to its spectrum resources. However, licensees must perform traffic steering to vacate the band without causing interference, whenever the incumbent requires. In this paper, a cognitive algorithm is proposed to take in advance decisions to promptly create a list of traffic steering routes whenever an unscheduled evacuation is demanded. This solution aims at guaranteeing the QoS and seamless connectivity during traffic steering. A performance evaluation conducted in a scenario composed of one LTE-LSA and three Wi-Fi network operators demonstrates that the proposed solution fulfills the time required by the unscheduled evacuation as well as guarantees the QoS and seamless connectivity of evacuees.**

I. Introduction

The traffic generated by mobile network operators is constantly growing and by 2020 it is expected to overload the existing licensed spectrum [1], leading to a resource scarcity problem. Licensed Shared Access (LSA) is an emerging solution to deal with this kind of problem, since it authorizes spectrum sharing by allowing the spectrum rights holder (*i.e.,* the incumbent user) to temporarily provide access to the LSA licensees [2]. However, the incumbent user is eligible to dynamically request the resources back at any time. Such request compels the LSA licensees to promptly evacuate the spectrum to avoid interfering with the incumbent services. In order to vacate the resources in a timely manner, the LSA licensees must implement fast handover strategies and consequently manage to steer the traffic of evacuees to available portions of the spectrum.

The main goal of this paper is to provide a cognitive mechanism to perform in advance decisions to allow traffic steering in unscheduled evacuation of LSA bands. Various solutions have been proposed to address this kind of evacuation ([3] [4] [5] [6] [7]). The main contribution of the proposed approach in comparison with these related works is to take decisions beforehand. In other words, the proposed approach enables the LSA licensee to create a list of potential traffic steering routes before an evacuation request is received, which considerably shortens the evacuation duration. An additional important contribution of the proposed solution is the in advance association of the Quality of Service (QoS) metrics considering different classes of service during the decision process. Thus, the traffic steering decision aims at maintaining the QoS of the evacuated users.

A novel traffic steering solution is proposed by extending an existing cognitive QoS-aware resources sharing architecture originally proposed by Kunst *et al.* ([8] [9]). The original architecture is used to gather updated information regarding resources usage of various operators in heterogeneous network scenarios. Since this architecture allows the implementation of different decision algorithms, in this paper, the original algorithm is replaced by one which is capable of taking in advance decisions. This kind of decision allows the selection of alternative routes for traffic steering in unscheduled spectrum evacuation scenarios. Specifically, a scenario composed of LTE-LSA and Wi-Fi network operators is considered to evaluate the performance of the proposed solution. Such evaluation is conducted via Matlab simulations based on an analytical system model. Results show that the proposed solution is able to allow fast spectrum evacuation and traffic steering, taking into account the QoS requirements of the evacuating users.

The main contributions of this paper are summarized as follows:

1) Proposal of a cognitive in advance decision algorithm to allow fast evacuation of LSA spectrum bands;
2) Fast traffic steering in unscheduled evacuation of LSA bands;
3) Performance and viability analysis (in terms of evacuation duration) of the proposed solution in heterogeneous network scenarios composed of LTE-LSA and Wi-Fi network operators.

The remainder of this paper is organized as follows. Current solutions for LSA spectrum evacuation are analyzed in Section II. The proposed solution is described in Section III. The performance evaluation is presented in Section IV. Finally, conclusions and directions for future work are presented in Section V.

II. Related work

Traffic steering is a current topic of research in LTE and LSA network scenarios. The goal of traffic steering is to find the most suitable evacuation route when vacating a frequency is necessary [10]. According to Mustonen *et al.* [11], the traffic steering is carried out on the basis of the capacity and load of heterogeneous networks. Nowadays, LTE features such as handover and traffic steering are oriented to be performed considering algorithms which provide cognitive decisions. This kind of decision brings intelligence to the allocation of radio and network resources, aiming at increasing the overall network QoS [7].

The cognitive engine designed by Martinmikko *et al.* [7] is the essential part of the cognitive radio trial environment to control different radio systems with the aim of guaranteeing QoS while carrying out handover and traffic offloading procedures. In fact, the cognitive engine analyzes alternative networks when high priority clients experience QoS degradation and when possible, carry out forced handover to deal with the problem. The cognitive decision making is an essential functionality to perform the forced handover of users and thus guarantee the QoS in accordance with their priority, regardless an evacuation of the LSA band takes place.

A Multilevel resource architecture was designed by Kunst *et al.* for the allocation of QoS-aware resources in heterogeneous wireless networks [8] [9]. This architecture relies on a broker which gathers together the updated information regarding the available network resources and them to be shared between the network operators.

Despite very relevant, related works are not concerned with time-sensitive traffic steering and handover procedures. Considering this limitation, in this paper is proposed a cognitive algorithm to carry out a fast traffic steering procedure. Our proposed solution takes into account both the QoS requirement of the evacuees and the time limit set by an LTE base station to release the LSA band ([6]) and thus avoid interference with the incumbent services. Details on the proposed approach are presented in next section.

III. Cognitive Traffic Steering Algorithm

An adaptation of Kunst *et al.* architecture ([9]) is presented in Fig. 1. The architecture allows communication among diverse network operators through a polling based mechanism. The left side of the figure illustrates the coexistence of LTE-LSA and Wi-Fi operators within the same geographical area. In the right side of the figure, the structure of the resources broker is represented. This Broker is responsible for coordinating resources sharing and is also adapted from Kunst *et al.* proposal.

The Broker plays the role of a centralized entity which keeps track of the network resources availability. Three levels are defined to provide independent and simultaneous control of different tasks of resources sharing management. These levels communicate with each other via Service Access Points (SAP) and are named accordingly to the function executed by each one: (I) Traffic Analysis Level, (II) Resources Knowledge Level, and (III) Cognition Level.

The first level of the Broker is responsible for controlling the polling mechanism used to gather updated information on the resources conditions of the Wi-Fi access cloud. The information received from each Wi-Fi operator contains a tuple composed of its identification, current average Delay, Jitter, and Throughput. This tuple is received and pre-processed by a Traffic Status analyzer and then relayed to the Traffic Profile Analysis block, which is responsible for keeping track of both current and historical values of the QoS parameters, which will feed the Resources Knowledge Level.

Databases are organized in the Resources Knowledge, which is the second level of the Broker. In the approach proposed in this paper, the Resources Knowledge level implements two databases to store information regarding the resources availability of LTE-LSA and Wi-Fi networks, respectively. This level plays a crucial role both on the traffic

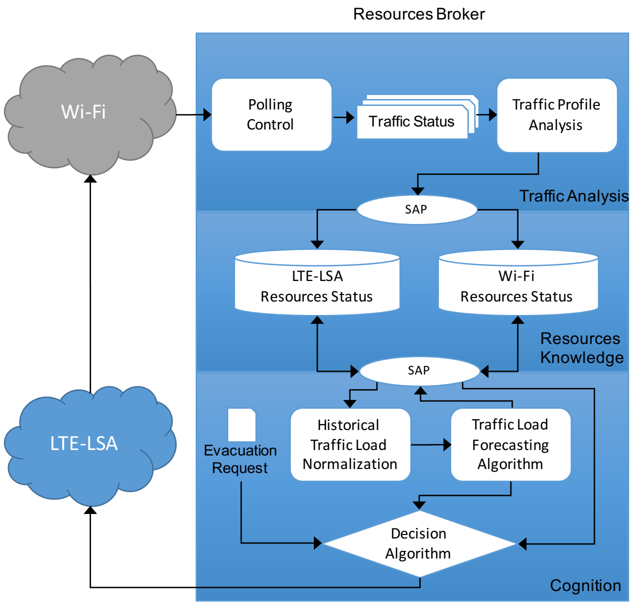


Fig. 1. Architecture Design

forecast and on the cognitive decision process which allows the unscheduled evacuation of the LSA band whenever necessary.

The Cognition Level accesses information on the second level of the Broker to take decisions when an evacuation is required. The evacuation request is composed of a struct which informs the CoS and the QoS requirements of the client. This level is constantly running, with the goal of taking in advance decisions regarding the traffic steering, which is used to promptly vacate the LSA band when required. This in advance decision demands the Cognition Level to forecast the traffic of the LTE-LSA and Wi-Fi operators in order to identify the best evacuation route. Such forecast requires knowledge about the historical traffic load, which is stored in the Resources Status Database of the LTE-LSA and Wi-Fi networks, respectively. Later, the historical traffic load is processed and normalized in the Cognition Layer. The resulting values serve as inputs to the Traffic Load Forecasting Algorithm.

In order to forecast the traffic behavior, a Multiple Linear Regression (MLR) model is implemented using Matlab. This model is based on a traffic measurement Y, which is related to a single predictor X for each observation. Therefore, the conditional mean function can be described as in (1), where α is the intercept and β is the coefficient.

$$E[Y \mid X] = \alpha + \beta X \tag{1}$$

Considering that multiple predictors (n) are available from the traffic traces, in this paper, the MLR modeled according to (2).

$$E[Y \mid X] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \tag{2}$$

The variability of the ith measurement Y around its mean value is specified in (3).

$$\mathrm{E}[Y \mid X_i] = \alpha + \beta_1 X_{i,1} + \beta_2 X_2 + \cdots + \beta_n X_{n,i} + \epsilon_i \quad (3)$$

In this case, the error assumptions for ϵ_i are: $\mathrm{E}[\epsilon_i] = 0$ and $\mathrm{var}(\epsilon_i) = \sigma^2$. The accuracy of the forecast can be measured by the mean absolute percent error (η), which is given by (4). In this equation, e_t represents the actual network occupation based on network traffic traces and y_t is the forecast occupation of the same network in a given instant of time.

$$\eta = \frac{1}{n} \left(\sum_{t=1}^{n} \left| \frac{e_{(t)}}{y_{(t)}} \right| \right) \quad (4)$$

The resulting forecast points compose a continuous traffic function, $f(x)$, which describes the occupied area of each analyzed network. In this context, let $f : D \rightarrow R$ be a function defined on a subset D of R and let $I = [a, b]$ be a close interval contained in D. In this paper, this closed interval represents the start and the end time of the forecast. Finally, let $P = \{[x_0, x_1], [x_1, x_2], \cdots, [x_{n-1}, x_n]\}$ be a partition of I such as $P = \{a = x_0, x_1, \cdots, x_n = b\}$. Thus, a Riemann sum (S) of f over I with partition P is defined in (5).

$$S = \sum_{i=1}^{n} f(x_i^*)(x_i - x_{i-1}) \quad (5)$$

When the number of points in P increase indefinitely, the equation (6) calculates the occupied area of each network, which can be related to the occupied network capacity.

$$A_{occupied} = \int_a^b f(x)dx = \lim_{x\rightarrow\infty} [s^*(P, f)] \quad (6)$$

This value is normalized considering the total capacity area (A_{total}) of each network operator. Its complement therefore represents the percentage of available resources of a given network. Let $\Theta = \{o_0, o_1, \cdots, o_{n-1}, o_n\}$ be a set of network operators. Thus, the free capacity percentage of the network operators is given by (7).

$$\forall o \in \Theta, A_{free}(o) = 1 - \left(\frac{A_{occupied}(o)}{A_{total}(o)} \right) \quad (7)$$

Three CoS are defined to accommodate different types of traffic regarding the QoS requirements. (I) Real-Time Services (RTS), to support delay and jitter sensitive real-time transmissions, (II) Multimedia Services (MS), comprehending real-time services with high throughput but no strict delay and jitter, and (III) Best Effort Services (BES), designed to support best effort transmissions without strict QoS requirements. Based on the CoS requirements and on the amount of free resource of each operator, calculated beforehand by the traffic forecasting algorithm, a decision algorithm is implemented, as defined in Algorithm 1.

In the proposed algorithm, the decision is based on information gathered from the Traffic Load Forecasting Algorithm. The outcomes of this algorithm are stored in the databases

Algorithm 1 Decision Algorithm

Require: r ▷ A struct containing a cognitive evacuation request
Require: $A_{total}(o)$ ▷ The total amount of resources of each operator
Require: $A_{occupied}(o) = \int_a^b f(x)dx = \lim_{x\rightarrow\infty}[s^*(P, f)]$ ▷ The amount of occupied resources of each operator

$c \leftarrow r.CoS$; $d \leftarrow r.Delay$; $t \leftarrow r.Throughput$
switch c **do**
  case RTS:
    for all $o \in \Theta$ **do**
      $A_{free}(o) = 1 - \left(\frac{A_{occupied}(o)}{A_{total}(o)}\right)$
      $delay(o) = get\_knowledge\_level(Wi - Fi, delay)$
      if $A_{free}(o) \leq t$ & $delay(o) \leq d$ **then**
        return o
      end if
    end for
  case MS:
    for all $o \in \Theta$ **do**
      $A_{free}(o) = 1 - \left(\frac{A_{occupied}(o)}{A_{total}(o)}\right)$
      if $A_{free}(o) \leq t$ & $delay(o) \leq d$ **then**
        return o
      end if
    end for
  case else:
    for all $o \in \Theta$ **do**
      $max\_operator = 0$
      $A_{free}(o) = 1 - \left(\frac{A_{occupied}(o)}{A_{total}(o)}\right)$
      if $A_{free}(o) > max\_operator$ **then**
        $max\_operator = o$
      end if
    end for
    return $max\_operator$

of the Resources Knowledge Level of the Resources Broker. Whenever an evacuation request is received, the decision algorithm queries the referred databases to obtain the updated forecast. This forecast is then considered along with the class of service of the request to search for a traffic steering route which is able to guarantee QoS of the evacuees.

IV. PERFORMANCE EVALUATION

In this section, an evaluation of the traffic steering is conducted that involves cognitive in advance decisions. This includes conducting an analysis of accurate decisions, processing time, and QoS requirements. The simulation scenario is discussed in Subsection IV-A, and the performance of the proposed solution is analyzed in Subsection IV-B with regard to three key factors: (I) traffic load forecasting, (II) cognitive decisions accuracy, and (III) cognitive traffic steering efficiency.

A. Simulation Scenario

Modeling the traffic demand of the LTE network operator is important to simulate the behavior of the proposed solution. The traffic models consider the arrival distribution and the traffic demanded per connection. This model is based on the WiMAX forum specification [12] and simulates three kinds of traffic: HTTP, Video, and VoIP. The remaining simulation parameters are summarized in Table I.

HTTP are used to model BES traffic. The transmissions comprise the main page, which has a given number of embedded objects, such as images, scripts, and other sorts of attached files. After requesting and receiving the files, the browser parses the page to make it readable to the user. The user then

TABLE I. TRAFFIC SIMULATION PARAMETERS

Parameter	Values for LTE-LSA Network
Channel Bandwidth	10 MHz
LTE Frame Length	10ms
Simulation Duration	1800s
% of HTTP Traffic	40%
% of VoIP Traffic	30%
% of Video Traffic	30%

reads the page before making a new request. The values of each phase of the HTTP statistical model are described in Table II.

TABLE II. HTTP TRAFFIC PARAMETERS

Component	Distribution	Parameters	PDF
Main Page Size	Truncated Lognormal	Mean = 10710 bytes SD = 25032 bytes Min = 100 bytes Max = 2 Mbytes	$\sigma = 1.37$ $\mu = 8.37$
Embedded Object Size	Truncated Lognormal	Mean = 7758 bytes SD = 126168 bytes Min = 50 bytes Max = 2 Mbytes	$\sigma = 2.36$ $\mu = 6.17$
Number of Embedded Objects	Truncated Pareto	Mean = 5.64 Max = 53	$\sigma = 1.1$ $\mu = 55$
Reading Time	Exponential	Mean = 30 s	$\mu = 0.033$
Parsing Time	Exponential	Mean = 0.13 s	$\mu = 7.69$

RTS are modeled to include VoIP transmissions, and Adaptive Multi-Rate (AMR) audio codec, which has ON/OFF behavior. This behavior is modeled to cover the activity of speech in conversations using this codec system. The duration of each period was modeled on the basis of an exponential distribution with an average of 1026 ms for ON period of (conversation) and 1171 ms for OFF period (silence). Finally, MS are modeled by video transmissions encoded using the MPEG-4 format.

The simulations are performed in Matlab considering the architectural model presented in Section III, the above traffic models, as well as the realistic traces obtained from CRAWDAD database to model Wi-Fi networks traffic [13]. The scenario consists of three Wi-Fi networks operating in no interfering channels and one LTE network operator using the LSA spectrum band.

B. Performance Evaluation

With regard to the performance of the proposed solution, the first factor to analyze is the accuracy of the traffic load forecasting model. The forecasting follows three key phases. The first is the time series extraction of traffic data from LTE-LSA and Wi-Fi networks. The second consists of fitting the polynomial curve of traffic data of both LTE-LSA and Wi-Fi networks. In the third phase, the forecasting is carried out by means of the MLR model as detailed in equation 2.

Considering that the time series is a sequence of data points, that generally consists of successive measurements made in a time interval [14]. These data points are divided into three data sets: training, validation, and testing. The training data set contains the traffic load measurement that corresponds to the first 15 minutes of the time series. The validation data set consists of 10 percent of the testing data set which is used to analyze the outcomes of the prediction, by taking account of metrics such as accuracy and processing time.

The MLR model processes the trained data set of simulated traffic demands for the LTE-LSA network, as well as the aggregate traffic of the Wi-Fi networks. The simulated traffic in the LTE-LSA network and the traffic traces of Wi-Fi are computed in units of seconds to improve the accuracy of the model. The first step of the analytical methodology involves calculating the polynomial curve fitting for smoothing out the peaks and noise of the network traffic. The polynomial was fixed at 10 degrees for curve fitting analysis of traffic of each network. The classification is then performed again and includes the new data points obtained from the ten degrees polynomial for the training, validation, and testing datasets. After this, the MLR model carries out the traffic load forecasting and the validation data set is used to evaluate its accuracy for each network.

The MLR accuracy is evaluated by the cross-validation method which involves the comparing the forecasted values with the current values. At this point, the MLR model can be adjusted to improve the accuracy of the upcoming predictions. The MAPE equation (4) is also used to calculate the accuracy of the MRL model. Fig. 2 shows the analysis of the accuracy of traffic load forecasting, which examines three Wi-Fi networks as possible traffic steering routes. As can be seen in the graph, the traffic load forecasting was very accurate and reached levels of 96.18%, 93.61%, and 94,20% degree of accuracy, for Wi-Fi networks 1, 2, and 3, respectively.

The traffic load forecasting is also correlated to the classes of service to analyze the QoS support feature of decision algorithm in terms of selecting the traffic steering route which presents the higher probability of preventing future network congestion. The outcomes of the simulation related to this scenario are depicted in Fig. 3 which shows the traffic load forecasting of VoIP, Video, and HTTP in the LTE-LSA network. In this case, the levels of accuracy are up to 95,72%, 98,47%, and 94,76%, respectively.

Every time the traffic load forecasting is performed, the values of the time series data points prediction are updated and input into the cognitive decision algorithm, which is responsible for selecting the traffic steering routes. The first step taken by the decision algorithm is to estimate the availability and occupation of bandwidth for each target network on the basis of the previous forecasting. The second step involves selecting the Wi-Fi networks which can guarantee the same level of QoS as that offered in LTE-LSA network. This kind of decision is made on the basis of the predicted availability of network resources. However, the same QoS level can only be ensured if the proposed solution is also able to include the delay metric. The third step performed by the decision algorithm is also related to the QoS and entails association of the CoS to the decision process.

The traffic load forecasting starts from the 15 minutes in Figs. 2 and 3 because it requires the historical traffic load measurements of the last 15 minutes to train the MLR model and predict the next 15 minutes traffic load trend with an accuracy close to 95% and to guarantee a fast response.

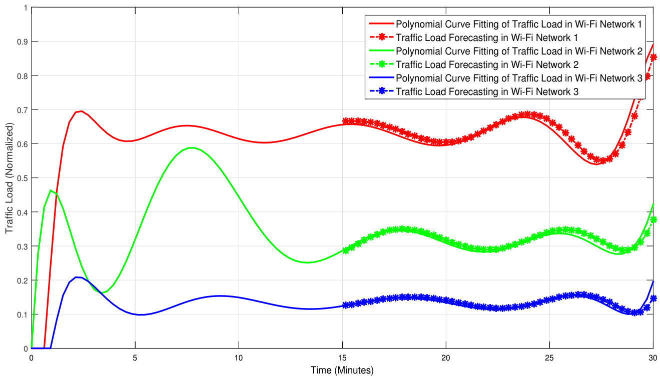


Fig. 2. Actual vs. Predicted Traffic Load for each Wi-Fi Network

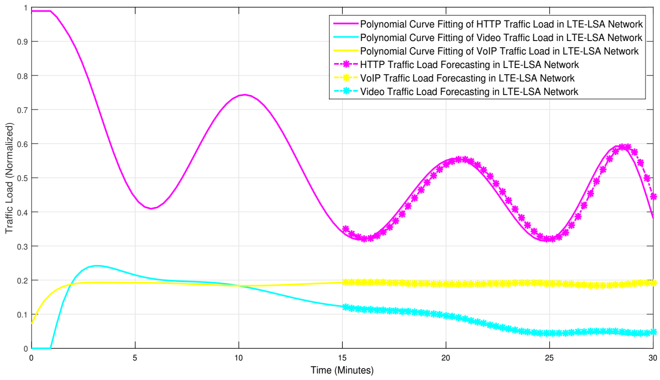


Fig. 3. Actual vs. Predicted Traffic Load for CoS in the LTE-LSA Network

The cognitive decision algorithm conducts the analysis of future bandwidth capacity of each overlapping Wi-Fi network by relying on the trapezoidal numerical integration to calculate the area under the curve of the MLR forecast. The area under the curve is equivalent to the percentage of occupied bandwidth resources for each Wi-Fi. As stated of the evaluated network scenario and an analysis of Fig. 2, the percentage of forecasted occupied bandwidth for Wi-Fi 1 is 65.8%, for Wi-Fi 2 is 31.6% and for Wi-Fi 3 is 15.4%. Based on these values, the in advance decision algorithm determines Wi-Fi 1 as a low priority route for traffic steering because of its very high traffic load. On the other hand, the cognitive decision defines Wi-Fi 2 and 3, as high-priority traffic steering routes. After this initial analysis, when an evacuation is required, the decision algorithm associates the class of services with the previous information to perform the traffic offloading while taking account of the QoS requirements of the evacuees.

The bandwidth occupation in Wi-Fi 1 oscillates close to 95% with its original users, making this network unavailable. Figs. 4 and 5 show the occupation of Wi-Fi 2 and 3 after the traffic steering. The bandwidth occupation in Wi-Fi 2 fluctuates between 40% and 70% after the offloading of Video and VoIP traffic demands from the LTE-LSA network, while Wi-Fi 3 network bandwidth occupation is around 80%. These results show that all the traffic was accommodated in the destination networks without overloading them. Thus, the QoS of the evacuees can be guaranteed without interfering with the original Wi-Fi users in terms of network capacity.

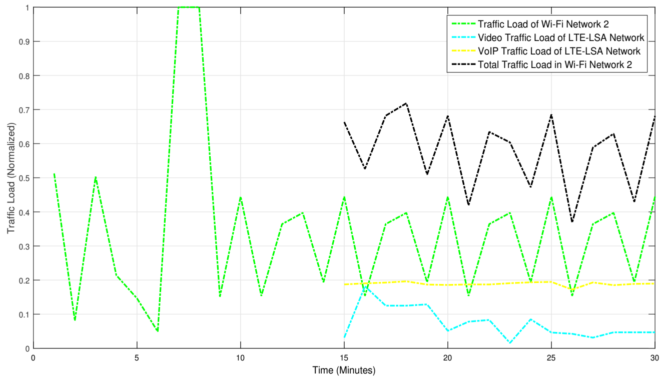


Fig. 4. Video and VoIP Traffic Steering from LTE-LSA to Wi-Fi 2

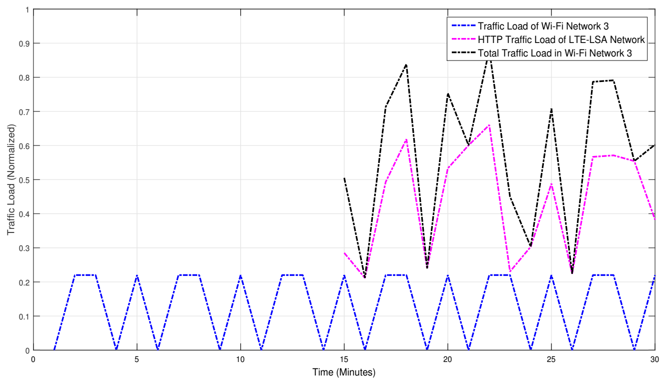


Fig. 5. HTTP Traffic Steering from LTE-LSA to Wi-Fi 3

Another important QoS metric is the delay. Fig. 6 shows the behavior of this metric considering a variable amount of connections accommodated by each Wi-Fi network. As can be seen in the graph, Wi-Fi 1 has the smallest delay value because it is a low-priority traffic steering route and thus the cognitive decision algorithm does not make it eligible to receive traffic from delay-sensitive applications. Wi-Fi networks 2 and 3, on the other hand, receive QoS sensitive traffic and are capable of keeping the average delay below 30ms. This value is sufficient to guarantee the QoS of multimedia traffic, which generally requires the delay to be between 100 and 200 ms.

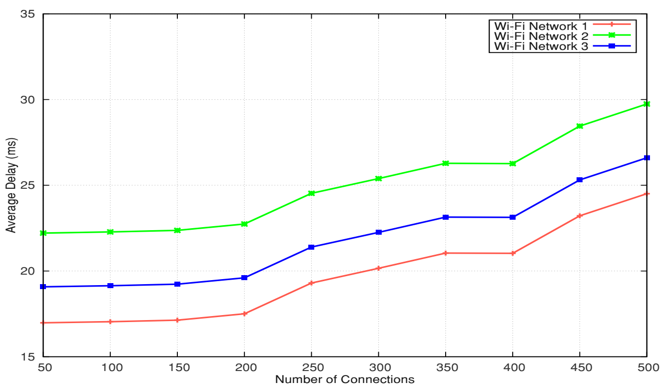


Fig. 6. Average Delay in Wi-Fi Networks

Another crucial factor that must be covered by the decision algorithm is to avoid interfering with the incumbent services in the event of an unscheduled evacuation. For this reason, the traffic steering must occur as fast as possible. The outcomes of this approach are similar to those of related work in the literature. Matinmikko *et al.* [7] was able to perform the decision in approximately 0.9 seconds on average, while Palola *et al.* [3] designed an algorithm which was able to carry out the decision in 0.624 seconds. Owing to the cognitive in advance decision mechanism, which is based on accurate forecasts, the proposed solution reduces the average decision time to values as low as 0.0371 of a second.

The processes related to the overall time required by the proposed solution to evacuate the LSA band and hence to offload the traffic to the selected Wi-Fi network, are outlined in Table III. Since the proposed approach involves making decisions in advance, the duration of both the decision process and the overall evacuation can be reduced. The CEPT Report 56 [6] stated that the duration for turning off an LTE base station with one sector delays 20.620 seconds on average. This limit of time constraints the ability of the traditional procedures to evacuate the UEs at a lower time to avoid interfering with the incumbent services in the LSA frequency and ensure the QoS of the evacuees. The results of the simulation show that the proposed solution allows the overall evacuation to be conducted in about 11.3 seconds, which represents a value that is around 46% below the specified time limit.

TABLE III. DURATION OF EVACUATION

Process	Average Duration [s]	Standard Deviation [s]
Traffic Load Forecasting	3.8267	0.2161
Cognitive Decision	0.0371	0.0051
Traffic Steering	7.3962	0.9477
Total Duration	11.2698	3.0163

V. CONCLUSIONS

This paper proposed a QoS-aware cognitive algorithm designed to take in advance decisions in the context of the unscheduled evacuation of LSA bands. This algorithm creates a list of candidate traffic steering routes taking into account the CoS and consequently the QoS requirements of the evacuating users. This kind of in advance decision allows a very fast evacuation to take place. The results show that the decision algorithm is faster than those in two related works and that the overall time consumed during the evacuation process is 46% faster than the maximum time allowed to avoid interfering with the incumbent user. Moreover, the outcomes of the simulations show that the proposed solution is able to guarantee QoS by including metrics such as throughput and delay.

Directions for future investigation include a deeper analysis of the performance of the proposed solution. This analysis can include the execution of the cognitive algorithm and the resources broker in realistic testbeds. Moreover, other QoS metrics, such as jitter and packet loss can be taken into account during the decision process. Finally, the proposed algorithm can be extended so that it can be executed in scenarios with a larger number of network operators which are able to implement different technologies, leading the application of the proposed solution to more heterogeneous scenarios.

ACKNOWLEDGMENT

The research leading to these results received funding from the European Commission H2020 programme under grant agreement no. 688941 (FUTEBOL), as well from the Brazilian Ministry of Science, Technology, Innovation, and Communication (MCTIC) through RNP and CTIC.

REFERENCES

[1] Cisco Visual Forecast, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015 - 2020," February 2016.

[2] A. Ponomarenko-Timofeev, A. Pyattaev, S. Andreev, Y. Koucheryavy, M. Mueck, and I. Karls, "Highly Dynamic Spectrum Management within Licensed Shared Access Regulatory Framework," *Communications Magazine, IEEE*, vol. 54, no. 9, pp. 100 – 109, March 2016.

[3] M. Palola, M. Matinmikko, J. Prokkola, M. Mustonen, M. Heikkilä, T. Kippola, S. Yrjola, V. Hartikainen, L. Tudose, A. Kivinen, J. Paavola, and K. Heiska, "Live field trial of Licensed Shared Access (LSA) concept using LTE network in 2.3 GHz band," *Dynamic Spectrum Access Networks (DYSPAN), 2014 IEEE International Symposium on*, no. 9, pp. 38 – 47, April 2014.

[4] M. Mustonen, M. Matinmikko, M. Palola, S. Yrjola, and K. Horneman, "An Evolution Toward Cognitive Cellular Systems: Licensed Shared Access for Network Optimization," *Communications Magazine, IEEE*, pp. 68 – 74, May 2015.

[5] ETSI TS 103 113, "Mobile Broadband Services in the 2300 MHz–2400 MHz Band under Licensed Shard Access Regime," July 2013.

[6] "Technological and Regulatory Options Facilitating Sharing between Wireless Broadband Applications (WBB) and the relevant Incumbent Services in the 2.3 GHz Band," *CEPT Report 56*, March 2015.

[7] M. Matinmikko, M. Palola, J. Prokkola, M. Heikkila, T. Kippola, T. Hanninen, S. Yrjola, and J. M., "Cognitive Radio Trial Environment: First Live Authorized Shared Access-Based Spectrum-Sharing Demonstration," *Vehicular Technology Magazine, IEEE*, vol. 8, no. 7, pp. 30 – 37, September 2013.

[8] R. Kunst, L. Avila, E. Pignaton, S. Bampi, and J. Rochol, "Improving QoS in Multi-operator Cellular Networks," *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, 2016.

[9] ——, "A Resources Sharing Architecture for Heterogeneous Wireless Cellular Networks," *The 41st IEEE Conference on Local Computer Networks (LCN)*, 2016.

[10] M. Dryjanski and M. Szydelko, "A unified traffic steering framework for LTE radio access network coordination," *IEEE Communications Magazine*, pp. 84 – 92, July 2016.

[11] M. Mustonen, M. Matinmikko, M. Palola, T. Rautio, and S. Yrjölä, "Analysis of requirements from standardization for Licensed Shared Access (LSA) system implementation," *Dynamic Spectrum Access Networks (DySPAN), 2015 IEEE International Symposium on*, pp. 71 – 81, 2015.

[12] WiMAX Forum, "WiMAX system evaluation methodology version 2.1," March 2014. [Online]. Available: http://www.wimaxforum.org/documents

[13] A. Schulman, D. Levin, and N. Spring, "CRAWDAD dataset umd/sigcomm2008 (v. 2009-03-02)," Downloaded from http://crawdad.org/umd/sigcomm2008/20090302/pcap, 2009.

[14] J. D. Hamilton, "Time series analysis," in *Time Series Analysis*. Princeton: Princeton University Press, 1994, ch. 4, p. 816.

REFERENCES

3GPP. *HTTP and FTP Traffic Modelfor 1xEV-DV Simulations*. 2008. 3GPP Forum Report TS GC5.

3GPP. *Service Aspects and Requirements for Network Sharing*. 2012. 3GPP Forum Report TS 22.951.

3GPP TS 36.300 v11.5.0, Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall Description.

4G AMERICAS. Spectrum sharing. p. 1–36, October 2014.

AKHTAR, A. M.; WANG, X.; HANZO, L. Synergistic spectrum sharing in 5g hetnets: A harmonized sdn-enabled approach. **IEEE Communications Magazine**, v. 54, n. 1, p. 40–47, January 2016. ISSN 0163-6804.

AKYILDIZ, I. et al. NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey. **Computer Networks**, v. 50, n. 13, p. 2127–2159, sep. 2006.

AKYILDIZ, I. et al. A survey on Spectrum Management in Cognitive Radio. **IEEE Communications Magazine**, v. 46, n. 4, p. 40–48, 2008.

Anatel. **Brazilian Spectrum of Frequencies Definition Plan**. 2015. Accessed november, 2015. Available from Internet: <http://sistemas.anatel.gov.br/pdff/>.

BARABAS, M. et al. Evaluation of network traffic prediction based on neural networks with multi-task learning and multiresolution decomposition. **Intelligent Computer Communication and Processing (ICCP), 2011 IEEE International Conference on**, p. 95–102, October 2011.

CAO, X.; CHEN, Y.; LIU, K. Cognitive radio networks with heterogeneous users: How to procure and price the spectrum? **Wireless Communications, IEEE Transactions on**, v. 14, n. 3, p. 1676–1688, March 2015. ISSN 1536-1276.

CEPT. Technical and operational requirements for the possible operation of cognitive radio systems in the white spaces of the frequency band 470 to 790 MHz. **ECC Report 159**, January 2011.

CHATZIKOKOLAKIS, K. et al. Toward spectrum sharing: opportunities and technical enablers. **Communications Magazine, IEEE**, v. 53, n. 7, p. 26–33, July 2015. ISSN 0163-6804.

CHEN, L. et al. Robustness, security and privacy in location-based services for future iot: A survey. **IEEE Access**, v. 5, p. 8956–8977, 2017. ISSN 2169-3536.

CISCO. **Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015 - 2020**. [S.l.], 2016. Available from Internet: <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html>.

COSTA-PEREZ, X. et al. Radio access network virtualization for future mobile carrier networks. **Communications Magazine, IEEE**, v. 51, n. 7, 2013.

DIXIT, S.; PERIYALWAR, S.; YANIKOMEROGLU, H. Secondary User Access in LTE Architecture Based on a Base-Station-Centric Framework With Dynamic Pricing. **Vehicular Technology, IEEE Transactions on**, v. 62, n. 1, p. 284–296, 2013.

ECC. **Licensed Shared Access (LSA)**. [S.l.], 2011.

ETSI. **Mobile broadband services in the 2300 MHz - 2400 MHz frequency band under Licensed Shared Access regime**. [S.l.], 2013.

FENG, H.; SHU, Y. Study on network traffic prediction techniques. **IEEE Wireless Communications, Networking and Mobile Computing, 2005. Proceedings. 2005 International Conference on**, September 2005.

FERNANDEZ, J.; KUNST, R.; ROCHOL, J. A cognitive algorithm for traffic steering in lte-lsa/wifi resources sharing scenarios. **IEEE International Conference on Communications**, 2017.

GARDELLIN, V.; DAS, S. K.; LENZINI, L. Self-coexistence in cellular cognitive radio networks based on the IEEE 802.22 standard. **Wireless Communications, IEEE**, v. 20, n. 2, 2013.

GUIZANI, M. et al. Large-scale cognitive cellular systems: resource management overview. **Communications Magazine, IEEE**, v. 53, n. 5, p. 44–51, May 2015. ISSN 0163-6804.

HAKIM, K. et al. Efficient Dynamic Spectrum Sharing in Cognitive Radio Networks: Centralized Dynamic Spectrum Leasing (C-DSL). **IEEE Transactions on Wireless Communications**, v. 9, n. 9, p. 2956–2967, sep. 2010. ISSN 1536-1276.

HAMILTON, J. D. **Time series analysis**. Princenton: Princeton University Press, 1943. 816 p.

HOSSAIN, E. et al. Evolution towards 5g multi-tier cellular wireless networks: An interference management perspective. **IEEE Wireless Communications**, v. 21, n. 9, p. 118–127, 2014. ISSN 1536-1284. Available from Internet: <http://arxiv.org/pdf/1401.5530.pdf>.

J., C.; W., L. Mobile broadband growth, spectrum scarcity, and sustainable competition. **The Research Conference Communication Information Internet Policy (TPRC'11)**, 2011. Available from Internet: <http://ssrn.com/abstract=1992423>.

KESHAVARZ, H. et al. Perturbation analysis for spectrum sharing in cognitive radio networks. **IEEE Transactions on Wireless Communications**, v. 9, n. 5, p. 1564–1570, may 2010. ISSN 1536-1276.

KHUN-JUSH, J. et al. Licensed shared access as complementary approach to meet spectrum demands: Benefits for next generation cellular systems. In: **ETSI Workshop on Reconfigurable Radio Systems**. [S.l.]: ETSI, 2012.

Kim, Bong Ho. **ApplicationTrafficMode**. 2007. WiMAX Forum AWG Contribution.

KITCHENHAM, B. et al. Systematic literature reviews in software engineering: A systematic literature review. **Information and Software Technology, Elsevier**, v. 51, n. 1, p. 7 –15, september 2009.

KUNST, R. et al. Improving qos in multi-operator cellular networks. **IEEE International Conference on Wireless and Mobile Computing, Networking and Communications**, 2016.

KUNST, R. et al. A resources sharing architecture for heterogeneous wireless cellular networks. **The 41st IEEE Conference on Local Computer Networks (LCN)**, 2016.

KUNST, R. et al. On the impact of hybrid errors on mobile WiMAX networks. **Computer Networks**, Elsevier, 2011. ISSN 0140-3664.

LIU, Y.; LEE, J. Y. B. An empirical study of throughput prediction in mobile data networks. **IEEE Global Communications Conference (GLOBECOM)**, December 2015.

LYE, K.; YUAN, M.; CAI, T. A spectrum comparison method for demand forecasting. **SIMTech technical reports.**, v. 10, n. 1, p. 32–35, 2009.

MALANCHINI, I.; VALENTIN, S.; AYDIN, O. Wireless resource sharing for multiple operators: Generalization, fairness, and the value of prediction. **Computer Networks**, v. 100, p. 110 – 123, 2016. ISSN 1389-1286. Available from Internet: <http://www.sciencedirect.com/science/article/pii/S1389128616300366>.

METIS. **Scenarios, Requirements and KPIs for 5G Mobile and Wireless System**. [S.l.], 2013. 1–5 p.

MITOLA, J.; MAGUIRE, G. Cognitive radio: making software radios more personal. **IEEE Personal Communications**, v. 6, n. 4, p. 13–18, 1999.

MOGENSEN, P. et al. LTE Capacity Compared to the Shannon Bound. In: **2007 IEEE 65th Vehicular Technology Conference**. [S.l.]: IEEE, 2007. p. 1234–1238.

MUSTONEN, M. et al. Cellular architecture enhancement for supporting the european licensed shared access concept. **Wireless Communications, IEEE**, v. 21, n. 3, p. 37–43, June 2014. ISSN 1536-1284.

NIAMI, S. et al. Anticipation of etx metric to manage mobility in ad hoc wireless networks. **International Conference on Ad-Hoc Networks and Wireless**, p. 29–42, 2014.

NOULAS, A. et al. Mining user mobility features for next place prediction in location-based services. In: **2012 IEEE 12th International Conference on Data Mining**. [S.l.: s.n.], 2012. p. 1038–1043. ISSN 1550-4786.

OSSEIRAN, A. et al. Scenarios for 5g mobile and wireless communications: The vision of metis project. **IEEE Communication Magazine**, v. 52, p. 26–35, May 2014. ISSN 0163-6804.

PALOLA, M. et al. Live field trial of Licensed Shared Access (LSA) concept using LTE network in 2.3 GHz band. **Dynamic Spectrum Access Networks (DYSPAN), 2014 IEEE International Symposium on**, n. 9, p. 38 – 47, April 2014. ISSN 14319481.

PAPADOPOULI, M.; RAFTOPOULOS, E.; SHEN, H. Evaluation of short-term traffic forecasting algorithms in wireless networks. In: **2006 2nd Conference on Next Generation Internet Design and Engineering, 2006. NGI '06.** [S.l.: s.n.], 2006. p. 8–109.

PONOMARENKO-TIMOFEEV, A. et al. Highly dynamic spectrum management in licensed shared access regulatory framework. **IEEE Communications Magazine**, v. 54, p. 100–109, 2016. ISSN 0163-6804.

PONOMARENKO-TIMOFEEV A. PYATTAEV, S. A. Y. K. M. M. A.; KARLS, I. Highly Dynamic Spectrum Management within Licensed Shared Access Regulatory Framework. **IEEE Communications Magazine**, v. 54, n. 9, p. 109–100, march 2016.

RODRIGUES, J.; NOGUEIRA, A.; SALVADOR, P. Improving the traffic prediction capability of neural networks using sliding window and multi-task learning mechanisms. **IEEE Evolving Internet (INTERNET), 2010 Second International Conference on**, November 2010.

SCHULMAN, A.; LEVIN, D.; SPRING, N. Crawdad dataset umd/sigcomm2008 (v. 2009-03-02). March 2009. Available from Internet: <http://crawdad.org/umd/sigcomm2008/20090302/pcap>.

SCHULMAN, D. L. A.; SPRING, N. **CRAWDAD dataset umd/sigcomm2008 (v. 2009-03-02)**. 2009. Downloaded from http://crawdad.org/umd/sigcomm2008/20090302/pcap.

SINGH, B. et al. Coordination protocol for inter-operator spectrum sharing in co-primary 5g small cell networks. **Communications Magazine, IEEE**, v. 53, n. 7, p. 34–40, July 2015. ISSN 0163-6804.

SIOMINA, I.; YUAN, D. Analysis of Cell Load Coupling for LTE Network Planning and Optimization. **Wireless Communications, IEEE Transactions on**, v. 11, n. 6, p. 2287–2297, 2012.

SOHUL, M. et al. Spectrum access system for the citizen broadband radio service. **Communications Magazine, IEEE**, v. 53, n. 7, p. 18–25, July 2015. ISSN 0163-6804.

SRINIVASA, S.; JAFAR, S. How much spectrum sharing is optimal in cognitive radio networks? **IEEE Transactions on Wireless Communications**, v. 7, n. 10, p. 4010–4018, oct. 2008. ISSN 1536-1276.

STRANG, G. Wavelets. **Sigma Xi, The Scientific Research Society**, v. 82, n. 3, p. 250–255, 1994. Available from Internet: <http://www.jstor.org/stable/29775194>.

STROBL, C.; MALLEY, J.; TUTZ, G. An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forests. In: . [S.l.: s.n.], 2009. v. 14, n. 4, p. 323–348.

UMTS Forum. **Mobile Traffic Forecast 2010-2020**. 2011. UMTS Forum Report 44.

UMTS Forum. **Report on Collective Use of Spectrum (CUS) and Other Spectrum Sharing Approaches**. 2011. RSPG11-392.

UMTS Forum. **RSPG Opinion on Licensed Shared Access**. 2013. SPG13-52.

WILKINSON, L. Classification and regression trees. **Systat**, v. 11, p. 35–56, 2004.

WiMAX Forum. **WiMAX System Evaluation Methodology version 2.1**. 2008. WiMAX Forum. in <http://www.wimaxforum.org/documents>. Accessed in March 2014.

XU, Q. et al. Proteus: Network performance forecast for real-time, interactive mobile applications. In: **Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services**. New York, NY, USA: ACM, 2013. (MobiSys '13), p. 347–360. ISBN 978-1-4503-1672-9. Available from Internet: <http://doi.acm.org/10.1145/2462456.2464453>.

YANG, C. et al. Advanced spectrum sharing in 5g cognitive heterogeneous networks. **IEEE Wireless Communications**, v. 23, n. 2, p. 94–101, April 2016. ISSN 1536-1284.

YI, C.; CAI, J. Multi-item spectrum auction for recall-based cognitive radio networks with multiple heterogeneous secondary users. **Vehicular Technology, IEEE Transactions on**, v. 64, n. 2, p. 781–792, Feb 2015. ISSN 0018-9545.

YUCEK, T.; ARSLAN, H. A survey of spectrum sensing algorithms for cognitive radio applications. **IEEE Communications Surveys & Tutorials**, v. 11, n. 1, p. 116–130, 2009.

ZHANG, N. et al. Beyond coexistence: Traffic steering in lte networks with unlicensed bands. **IEEE Wireless Communications**, v. 23, n. 6, p. 40–46, December 2016. ISSN 1536-1284.

ZOU, J. et al. Packet-based preamble design for random access in massive iot communication systems. **IEEE Access**, PP, n. 99, p. 1–1, 2017. ISSN 2169-3536.