Novel Machine Learning-Based Techniques for Efficient Resource Allocation in Next Generation Wireless Networks

Dissertation by
Ismail Ali AlQerm

In Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy

King Abdullah University of Science and Technology, Thuwal,
Kingdom of Saudi Arabia

February, 2018
The dissertation of Ismail AlQerm is approved by the examination committee.

Committee Chairperson: Prof. Basem Shihada
Committee Members: Prof. Mohamed-Slim Alouini, Prof. Xin Gao, Prof. Octavia Dobre
Cognitive radio has been found to be the paradigm that is able to satisfy the above requirements. It is a very interdisciplinary topic that incorporates flexible system architectures, machine learning, context awareness and cooperative networking. Mitola’s vision about cognitive radio intended to build context-sensitive smart radios that are able to adapt to the wireless environment conditions while maintaining quality of service support for different applications. Artificial intelligence techniques including heuristics algorithms and machine learning are the shining tools that are employed to serve the new vision of cognitive radio. In addition, these techniques show a potential to be utilized in an efficient resource allocation for the upcoming 5G networks’ structures such as heterogeneous multi-tier 5G networks and heterogeneous cloud radio access networks due to their capability to allocate resources according to real-time data analytics.

We study cognitive radio from a system point of view focusing closely on architectures; artificial intelligence techniques that can enable intelligent radio resource allocation and efficient radio parameters reconfiguration. We propose a modular cognitive resource management architecture, which facilitates a development of flexible control for resources management in diverse wireless networks. The core operation of the proposed architecture is decision-making for resource allocation and system’s parameters adaptation. Thus, we develop the decision-making mechanism using different artificial intelligence techniques, evaluate the performance achieved, and determine the tradeoff using one technique over the others.

In addition, we explore the use of enhanced online learning to perform efficient resource allocation in the upcoming 5G networks to maximize energy efficiency and
data rate. The considered 5G structures are heterogeneous multi-tier networks with device-to-device communication and heterogeneous cloud radio access networks. We propose power and resource blocks allocation schemes to maximize energy efficiency and data rate in heterogeneous 5G networks. Moreover, traffic offloading from large cells to small cells in 5G heterogeneous networks is investigated and an online learning based traffic offloading strategy is developed to enhance energy efficiency. Energy efficiency problem in heterogeneous cloud radio access networks is tackled using online learning in centralized and distributed fashions.
ACKNOWLEDGEMENT

My eternal cheerleader: I miss our interesting and long-lasting chats. My forever interested, encouraging and always enthusiastic my mother: Fatemah AlKhateeb, she was always keen to know what I was doing and how I was proceeding, although it is likely that she has never grasped what it was all about! I will miss your screams of joy whenever a significant momentous was reached and also just your general impudence. I am grateful to my siblings and father, who have provided me through moral and emotional support in my life. I am also grateful to my other family members and friends who have supported me along the way.

A very special gratitude goes out to all down at research fund and also rich foundation for helping and providing the funding for the work. I am also grateful to my PhD advisor: Prof. Basem Shihada for his unfailing support and assistance and the US-DDS for funding the PhD research. And last but by no means least, also to everyone in the impact hub, it was great sharing laboratory with all of you during last six years.

Thanks for all your encouragement!
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copyright</td>
<td>3</td>
</tr>
<tr>
<td>Abstract</td>
<td>4</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td>6</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>11</td>
</tr>
<tr>
<td>List of Figures</td>
<td>14</td>
</tr>
<tr>
<td>List of Tables</td>
<td>19</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>20</td>
</tr>
<tr>
<td>1.1 Problem Statement and Motivation</td>
<td>20</td>
</tr>
<tr>
<td>1.2 Thesis Objective and Contributions</td>
<td>23</td>
</tr>
<tr>
<td>1.3 Thesis Organization</td>
<td>26</td>
</tr>
<tr>
<td>2 Background and State of the Art</td>
<td>27</td>
</tr>
<tr>
<td>2.1 Cognitive Radio Resource Management</td>
<td>28</td>
</tr>
<tr>
<td>2.1.1 Software Defined Radios Evolution</td>
<td>28</td>
</tr>
<tr>
<td>2.1.2 Cognitive Resource Management Concept</td>
<td>30</td>
</tr>
<tr>
<td>2.1.3 Cognition Cycle</td>
<td>31</td>
</tr>
<tr>
<td>2.1.4 Cognitive Radio Resource Management Architectures</td>
<td>32</td>
</tr>
<tr>
<td>2.2 Cognitive Engine and Radio Parameters Adaptation</td>
<td>33</td>
</tr>
<tr>
<td>2.2.1 Cognitive Engine</td>
<td>33</td>
</tr>
<tr>
<td>2.2.2 Cognitive Radio Parameters Adaptation Techniques</td>
<td>34</td>
</tr>
<tr>
<td>2.3 Cognitive Resource Allocation in LTE Networks</td>
<td>37</td>
</tr>
<tr>
<td>2.3.1 Cognitive Radio Resource Management for OFDMA Systems</td>
<td>37</td>
</tr>
<tr>
<td>2.3.2 Cognitive Approaches for Interference Mitigation in LTE Heterogeneous Networks</td>
<td>39</td>
</tr>
<tr>
<td>2.4 Efficient Resource Allocation in 5G</td>
<td>40</td>
</tr>
<tr>
<td>2.4.1 Efficient Resource Allocation in 5G Heterogeneous Networks</td>
<td>40</td>
</tr>
</tbody>
</table>
### 2.4.2 Energy Efficient Base Stations Operation in 5G Heterogeneous Networks

### 2.4.3 Energy Efficient Resource Allocation in 5G Heterogeneous Cloud Radio Access Networks

### 2.4.4 Online Learning Background

### 3 CogWnet: A Resource Management Architecture for Cognitive Radio Networks

#### 3.1 Radio Resources Management in Wireless Networks

#### 3.2 CogWnet Architecture

- **3.2.1 CogWnet Design Principles**
- **3.2.2 CogWnet Resource Management Phases**
- **3.2.3 CogWnet Structure**
- **3.2.4 CogWnet Architecture Layers**

#### 3.3 CogWnet Resources Management Mechanism

- **3.3.1 Channel Allocation Mechanism**
- **3.3.2 Configuration Mechanism**

#### 3.4 CogWnet Testbed Implementation

- **3.4.1 Motivation**
- **3.4.2 Testbed Topology**
- **3.4.3 Evaluation Results and Discussion**

#### 3.5 CogWnet Decision-Making Evaluation

- **3.5.1 Resource Management for IEEE 802.22 Wireless Regional Area Networks (WRANs) Scenario**
- **3.5.2 Resource Management for an Efficient Spectrum Utilization**

### 4 Artificial Intelligence Approaches for Cognitive Engine Design and Radio Parameters Configuration

#### 4.1 Adaptive Multi-Objective Optimization Based Cognitive Engine Using Enhanced Genetic Algorithm

- **4.1.1 Constrained Multi-Objective Optimization in Cognitive Radio**
- **4.1.2 Adaptive Multi-Objective Optimization Cognitive Engine Mechanism**
- **4.1.3 Evaluation**

#### 4.2 Hybrid Cognitive Engine for Radio Parameters Adaptation

- **4.2.1 Hybrid Engine Design and Operation Mechanism**
- **4.2.2 Hybrid Engine Complexity Consideration**
4.2.3 Hybrid Engine Testbed Implementation ........................................ 106
4.2.4 Simulation Evaluation .......................................................... 108
4.3 Supervised Cognitive System: A New Vision for Cognitive Engine De-
sign in Wireless Networks ............................................................ 111
  4.3.1 Supervised Cognitive System Architecture ............................... 112
  4.3.2 SCS Operation ................................................................. 118
  4.3.3 SCS Performance Evaluation ................................................ 122

5 Cognitive Radio with Learning Capability for Efficient Resource
Management in LTE Networks ..................................................... 132
  5.1 Cognitive Based Resource Blocks Allocation and Radio Parameters
adaptation in LTE Systems .......................................................... 133
    5.1.1 Throughput Optimization and Interference Management Model 134
    5.1.2 Integration of LTE and CogWnet ......................................... 137
    5.1.3 Learning Mechanism to Enhance Cognitive Radio Resource Man-
agement (RRM) ............................................................ 138
    5.1.4 Cognitive RRM Scheme Functionality .................................. 139
    5.1.5 Cognitive RRM Scheme Performance Evaluation ................... 141
  5.2 Cognitive Aware Interference Mitigation Scheme for LTE Femtocells..... 148
    5.2.1 System Model ............................................................ 149
    5.2.2 The Interference Mitigation Scheme .................................... 152
    5.2.3 Performance Evaluation .................................................. 158

6 Machine Learning Approaches for Resource Allocation Toward Effi-
cient Next Generation 5G Networks .......................................... 163
  6.1 Cooperative Online Learning Scheme for Resource Allocation in 5G
Heterogeneous Networks .......................................................... 164
    6.1.1 System Model ............................................................ 165
    6.1.2 Problem Formulation ...................................................... 167
    6.1.3 Online Learning System Parameters .................................... 170
    6.1.4 Resource Allocation Mechanism ....................................... 171
    6.1.5 Performance Evaluation .................................................. 174
  6.2 Energy Efficient Power Allocation in Multi-tier 5G Networks Using
Enhanced Online Learning .......................................................... 179
    6.2.1 System Model ............................................................ 180
    6.2.2 Problem Formulation ...................................................... 183
    6.2.3 Power Allocation Learning Model ...................................... 185
6.2.4 Power Allocation with Approximated-Intuition Based Online Learning ........................................... 189
6.2.5 Power Allocation Algorithm Convergence ................................................................. 193
6.2.6 Numerical Results .......................................................................................................... 200

6.3 Energy Efficient Traffic Offloading in Multi-tier Heterogeneous 5G Networks Using Intuitive Online Learning ................................................................. 206
6.3.1 System Model ................................................................................................................. 208
6.3.2 Problem Formulation ....................................................................................................... 212
6.3.3 Stochastic Traffic Offloading Using Intuitive Online Learning ......................................... 214
6.3.4 Numerical Results .......................................................................................................... 225

6.4 Sophisticated Online Learning Scheme for Green Resource Allocation in 5G Heterogeneous Cloud Radio Access Networks ......................................................... 232
6.4.1 System Model and Problem Formulation ........................................................................... 233
6.4.2 Online Learning Model for Resource Allocation ............................................................. 239
6.4.3 Centralized Approximated Online Learning Resource Allocation Scheme ........................................... 242
6.4.4 Decentralized Multi-Agent Online Learning Resource Allocation Scheme ........................................... 247
6.4.5 Testbed Implementation ................................................................................................. 251
6.4.6 Numerical Results .......................................................................................................... 260

7 Concluding Remarks ........................................................................................................... 268

References ............................................................................................................................... 271

Publications ............................................................................................................................... 297
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3GPP</td>
<td>3rd Generation Partnership Project</td>
</tr>
<tr>
<td>A/D</td>
<td>Analog to Digital Converters</td>
</tr>
<tr>
<td>ABC</td>
<td>Artificial Bee Colony Algorithm</td>
</tr>
<tr>
<td>ACO</td>
<td>Ant Colony Optimization</td>
</tr>
<tr>
<td>AI</td>
<td>Artificial Intelligence</td>
</tr>
<tr>
<td>AMO</td>
<td>Adaptive Multi Objective Optimization</td>
</tr>
<tr>
<td>ANNs</td>
<td>Artificial Neural Networks</td>
</tr>
<tr>
<td>AWGN</td>
<td>Additive White Gaussian Noise</td>
</tr>
<tr>
<td>BBU</td>
<td>Baseband Unit</td>
</tr>
<tr>
<td>BEE2</td>
<td>Berkeley Emulation Engine</td>
</tr>
<tr>
<td>BER</td>
<td>Bit Error Rate</td>
</tr>
<tr>
<td>BSs</td>
<td>Base Stations</td>
</tr>
<tr>
<td>CAPRI</td>
<td>Common Application Requirement Interface</td>
</tr>
<tr>
<td>CBR</td>
<td>Case Based Reasoning</td>
</tr>
<tr>
<td>CCC</td>
<td>Common Control Channel</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CDMA</td>
<td>Code Division Multiple Access</td>
</tr>
<tr>
<td>CE</td>
<td>Cognitive Engine</td>
</tr>
<tr>
<td>CKL</td>
<td>Case Knowledge Learning</td>
</tr>
<tr>
<td>CR</td>
<td>Cognitive Radio</td>
</tr>
<tr>
<td>CRANs</td>
<td>Cloud Radio Access Networks</td>
</tr>
<tr>
<td>CSMA</td>
<td>Carrier Sense Multiple Access</td>
</tr>
<tr>
<td>D2D</td>
<td>Device to Device</td>
</tr>
<tr>
<td>DSA</td>
<td>Dynamic Spectrum Access</td>
</tr>
<tr>
<td>DSPs</td>
<td>Digital Signal Processors</td>
</tr>
<tr>
<td>DTMDP</td>
<td>Discrete-Time Markov decision process</td>
</tr>
<tr>
<td>DTs</td>
<td>Decision Trees</td>
</tr>
<tr>
<td>eNBs</td>
<td>E-UTRAN Node Bs</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>FBSs</td>
<td>Femto Base Stations</td>
</tr>
<tr>
<td>FCC</td>
<td>Federal Communications Commission</td>
</tr>
<tr>
<td>FER</td>
<td>Frame Error Rate</td>
</tr>
<tr>
<td>FFR</td>
<td>Fractional Frequency Reuse</td>
</tr>
<tr>
<td>FPGA</td>
<td>Field Programmable Gate Array</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithm</td>
</tr>
<tr>
<td>GENI</td>
<td>Generic Network Interface</td>
</tr>
<tr>
<td>GPPs</td>
<td>General Purpose Processors</td>
</tr>
<tr>
<td>GSM</td>
<td>Global System for Mobile Communications</td>
</tr>
<tr>
<td>GT</td>
<td>Game Theory</td>
</tr>
<tr>
<td>H-CRANs</td>
<td>Heterogeneous Cloud Radio Access Networks</td>
</tr>
<tr>
<td>HCS</td>
<td>Hill Climbing Search</td>
</tr>
<tr>
<td>Hetnets</td>
<td>Heterogeneous Networks</td>
</tr>
<tr>
<td>ICNIA</td>
<td>Integrated Communications Navigation and Identification Avionics</td>
</tr>
<tr>
<td>IDEN</td>
<td>Integrated Digital Enhanced Network</td>
</tr>
<tr>
<td>ISM</td>
<td>Industrial, Scientific and, Medical</td>
</tr>
<tr>
<td>JRRM</td>
<td>Joint Radio Resources Management</td>
</tr>
<tr>
<td>KBR</td>
<td>Knowledge Based Reasoning</td>
</tr>
<tr>
<td>KKT</td>
<td>Karush-Kuhn-Tucker</td>
</tr>
<tr>
<td>LTE</td>
<td>Long Term Evolution</td>
</tr>
<tr>
<td>MBSs</td>
<td>Macro Base Stations</td>
</tr>
<tr>
<td>MDPs</td>
<td>Markov Decision Processes</td>
</tr>
<tr>
<td>MRRM</td>
<td>Multi Radio Resources Management</td>
</tr>
<tr>
<td>MUEs</td>
<td>Macro Users</td>
</tr>
<tr>
<td>OFDMA</td>
<td>Orthogonal Frequency Division Multiple Access</td>
</tr>
<tr>
<td>PER</td>
<td>Packet Error Rate</td>
</tr>
<tr>
<td>PSO</td>
<td>Particle Swarm Optimization</td>
</tr>
<tr>
<td>PUs</td>
<td>Primary Users</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>RA</td>
<td>Resource Allocation</td>
</tr>
<tr>
<td>RATs</td>
<td>Radio Access Technologies</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>RBS</td>
<td>Rule Based Systems</td>
</tr>
<tr>
<td>RBs</td>
<td>Resource Blocks</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RL</td>
<td>Reinforcement Learning</td>
</tr>
<tr>
<td>RRRs</td>
<td>Remote Radio Heads</td>
</tr>
<tr>
<td>RRM</td>
<td>Radio Resource Management</td>
</tr>
<tr>
<td>RSRP</td>
<td>Reference Signal Received Power</td>
</tr>
<tr>
<td>RSS</td>
<td>Received Signal Strength</td>
</tr>
<tr>
<td>SA</td>
<td>Simulated Annealing</td>
</tr>
<tr>
<td>SAN</td>
<td>Software Adaptive Network</td>
</tr>
<tr>
<td>SB</td>
<td>Sub-Band</td>
</tr>
<tr>
<td>SCS</td>
<td>Supervised Cognitive System</td>
</tr>
<tr>
<td>SDRs</td>
<td>Software Defined Radios</td>
</tr>
<tr>
<td>SER</td>
<td>Symbol Error Rate</td>
</tr>
<tr>
<td>SINR</td>
<td>Signal to Interference and Noise Ratio</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>STs</td>
<td>Secondary Transmitters</td>
</tr>
<tr>
<td>SUs</td>
<td>Secondary Users</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>UE</td>
<td>User Equipment</td>
</tr>
<tr>
<td>ULLA</td>
<td>Universal Link Layer Application Interface</td>
</tr>
<tr>
<td>USRP</td>
<td>Universal Software Radio Peripherals</td>
</tr>
<tr>
<td>WCDMA</td>
<td>Wideband Code Division Multiple Access</td>
</tr>
<tr>
<td>WRANs</td>
<td>Wireless Regional Area Networks</td>
</tr>
<tr>
<td>WRAP</td>
<td>Wireless Research Open-Access Platform</td>
</tr>
<tr>
<td>XML</td>
<td>Extensible Markup Languages</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

2.1 Cognition Cycle ........................................... 31

3.1 CogWnet Architecture .................................... 52
3.2 The Repository ............................................ 57
3.3 The Parameter Mapper .................................... 59
3.4 The Policy Layer .......................................... 60
3.5 Radio Resource Management: CogWnet modules execute cognitive functions. Actions are scheduled back through the link configuration state ........................................... 62
3.6 Goodput and utility achieved in the home scenario .......... 69
3.7 Performance of the home network with and without CogWnet .... 70
3.8 Performance of CogWnet in Mitigating Interference ............ 71
3.9 Throughput comparison between CogWnet and 802.11 with interference from external packet injector and from 802.11 based source and external source respectively .......... 72
3.10 Throughput comparison between 802.11 and CogWnet based on the distance between sender and receiver .......... 72
3.11 Utility results for different cognitive engines .................. 75
3.12 Average adaptation time for different cognitive engines ........ 76
3.13 SNR vs system throughput for the three engines ............... 78

4.1 Adaptive Multi Objective Optimization (AMO) Cognitive Engine (CE) fitness score comparison in low power scenario .......... 95
4.2 AMO CE fitness score comparison in multimedia scenario .......... 95
4.3 AMO CE fitness score comparison in emergency scenario .......... 96
4.4 Throughput of different cognitive engines with physical layer optimization .......... 96
4.5 Throughput of different cognitive engines with physical and MAC layer optimization .......... 97
4.6 Throughput of different cognitive engines with cross layer optimization .......... 98
4.7 Hybrid cognitive engine adaptation stages ................. 100
4.8 Three objectives decision-tree ........................................ 103
4.9 Average Case Based Reasoning (CBR) access time with different database size ................................................................. 106
4.10 PER of different cognitive engines with narrow band interference in reliability scenario ......................................................... 107
4.11 SINR of different cognitive engines with narrow band interference in reliability scenario ......................................................... 107
4.12 Network throughput in variable traffic scenario ......................... 109
4.13 Average PER for different cognitive engines ................................. 110
4.14 Average user throughput for different cognitive engines ................. 110
4.15 SCS system components ...................................................... 113
4.16 SCS operation flow ............................................................. 119
4.17 Performance of different CE learning techniques ......................... 120
4.18 Office experiment layout ....................................................... 123
4.19 Average goodput with various network load in the office experiment . 124
4.20 Propagation experiment layout ............................................... 125
4.21 Average PER at different locations in the propagation experiment .. 126
4.22 Interference experiment layout ............................................... 126
4.23 Average SE against SNR in the interference experiment ................. 127
4.24 Performance of SCS in different network scenarios ..................... 128
4.25 System goodput as function of time step in SCS simulation .......... 129
4.26 System PER with variable SNR in SCS simulation ...................... 130
4.27 System SE with different SNR in SCS simulation ...................... 130

5.1 Integration of Long Term Evolution (LTE) and CogWnet Architecture .............................................................. 137
5.2 CogWnet with learning mechanism ........................................ 139
5.3 System throughput comparing cognitive scheme with non-cognitive LTE scheme ............................................................... 142
5.4 LTE System capacity comparison of cognitive RRM and non-cognitive scheme ............................................................... 143
5.5 LTE system normalized throughput comparison between cognitive RRM and random scheme ......................................................... 143
5.6 User satisfaction probability using different RRM schemes .............. 144
5.7 The evolution of the matching probability .................................. 145
5.8 Throughput of all UEs in LTE femtocells topology ....................... 147
5.9 Packet dropping rate of video UEs in LTE femtocells topology ........ 147
5.10 Femtocells Network .......................................................... 150
5.11 CDF of macrocell user’s SINR in LTE Heterogeneous Networks (Hetnets) .............................................. 160
5.12 CDF of femtocell user's SINR in LTE Hetnets .......................................................... 160
5.13 Spectral efficiency comparison for all schemes in LTE Hetnets ..................................................... 161
5.14 Femtocell user throughput as a function of the number of Femto Base Stations (FBSs) in LTE Hetnets .......................................................... 161

6.1 System Model of 5G Hetnets .......................................................... 166
6.2 Online resource allocation mechanism in 5G Hetnets .......................................................... 173
6.3 Aggregate system throughput achieved against number of active transmissions in 5G Hetnets .......................................................... 176
6.4 Spectral efficiency achieved by different resource allocation schemes in 5G Hetnets .......................................................... 176
6.5 Jain’s fairness index with different number of active transmissions in 5G Hetnets .......................................................... 177
6.6 Mean SINR at MUE receiver in 5G Hetnets .......................................................... 178
6.7 Average Outage ratio at the underlay tier in 5G Hetnets .......................................................... 178
6.8 Average outage ratio at the macro tier in 5G Hetnets .......................................................... 179
6.9 System Model of multi-tier 5G networks .......................................................... 180
6.10 Average system energy efficiency of different schemes in 5G multi-tier network .......................................................... 202
6.11 Impact of the number of STs on system energy efficiency in 5G multi-tier network .......................................................... 202
6.12 Average system energy efficiency under different CNR ratio in 5G multi-tier network .......................................................... 203
6.13 Average system spectral efficiency of different schemes in 5G multi-tier network .......................................................... 203
6.14 Average spectral efficiency per each femtocell with variable number of femtocells in 5G multi-tier network .......................................................... 204
6.15 Average system spectral efficiency as a function of maximum transmission power of pico BS in 5G multi-tier network .......................................................... 205
6.16 Average SINR for macro users of different schemes in 5G multi-tier network .......................................................... 205
6.17 Average SINR for pico users of different schemes in 5G multi-tier network .......................................................... 206
6.18 Average SINR for each macro UE with variable number of UEs in 5G multi-tier network .......................................................... 206
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.19</td>
<td>Average SINR for each pico UE with variable number of UEs in 5G multi-tier network</td>
</tr>
<tr>
<td>6.20</td>
<td>Average SINR for each D2D receiver with variable number of UEs in 5G multi-tier network</td>
</tr>
<tr>
<td>6.21</td>
<td>System model of multi-tier heterogeneous 5G networks</td>
</tr>
<tr>
<td>6.22</td>
<td>5G multi-tier heterogeneous networks simulation environment</td>
</tr>
<tr>
<td>6.23</td>
<td>Average system traffic Load in 5G Hetnets</td>
</tr>
<tr>
<td>6.24</td>
<td>Average system energy efficiency in 5G Hetnets</td>
</tr>
<tr>
<td>6.25</td>
<td>Average system energy efficiency under different network loads in 5G Hetnets</td>
</tr>
<tr>
<td>6.26</td>
<td>Average system energy efficiency against the number of small cells in 5G Hetnets</td>
</tr>
<tr>
<td>6.27</td>
<td>CDF of data rate for macro users associated with macro BS in 5G Hetnets</td>
</tr>
<tr>
<td>6.28</td>
<td>CDF of data rate for macro users offloaded to small BS in 5G Hetnets</td>
</tr>
<tr>
<td>6.29</td>
<td>Heterogeneous Cloud Radio Access Networks (H-CRANs) Architecture</td>
</tr>
<tr>
<td>6.30</td>
<td>Centralized testbed implementation topology of 5G H-CRANs</td>
</tr>
<tr>
<td>6.31</td>
<td>Decentralized testbed implementation topology of 5G H-CRANs</td>
</tr>
<tr>
<td>6.32</td>
<td>Conceptual model of the resource allocation system in 5G H-CRANs</td>
</tr>
<tr>
<td>6.33</td>
<td>Average high QoS pico BS capacity in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.34</td>
<td>Average low QoS pico BS capacity in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.35</td>
<td>Average macro BS capacity in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.36</td>
<td>Average BER experienced by high QoS RUE in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.37</td>
<td>Average BER experienced by low QoS RUE in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.38</td>
<td>Average BER experienced by MUE in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.39</td>
<td>Average energy efficiency for high QoS pico BS in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.40</td>
<td>Average energy efficiency for low QoS pico BS in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.41</td>
<td>Energy efficiency convergence in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.42</td>
<td>Energy efficiency with variable $P_{max}$ in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.43</td>
<td>Energy efficiency with variable SINR threshold for MUEs in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.44</td>
<td>Spectral efficiency convergence in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.45</td>
<td>Spectral efficiency with variable maximum RRH power in 5G H-CRANs system</td>
</tr>
<tr>
<td>6.46</td>
<td>Data rate CDF for RUEs with high QoS requirements in 5G H-CRANs system</td>
</tr>
<tr>
<td>Data rate CDF for RUEs with low QoS requirements in 5G H-CRANs system</td>
<td>265</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>Data rate CDF for MUEs in 5G H-CRANs system</td>
<td>266</td>
</tr>
</tbody>
</table>
LIST OF TABLES

3.1 Cognitive functions for radio resources management ........................................ 63
3.2 Adjustable parameters in the WRANs ............................................................. 75
3.3 Simulation settings ....................................................................................... 75
3.4 Simulation Parameters .................................................................................. 78
4.1 Parameters used in AMO CE simulation ....................................................... 94
4.2 Configuration parameters adaptation of hybrid CE ....................................... 107
4.3 System parameters and performance metrics for hybrid engine simulation .... 109
4.4 Time consumption of different cognitive engines ......................................... 111
4.5 Average performance metrics in the office experiment ............................... 124
4.6 Average performance metrics in the propagation experiment ................... 125
4.7 Average performance metrics in the interference experiment .................... 127
4.8 Transmission parameters adaptation ranges in SCS simulation ................. 128
4.9 Average adaptation time for different cognitive systems ............................ 131
5.1 LTE environment simulation parameters .................................................... 142
5.2 LTE femtocells environment simulation parameters .................................. 146
5.3 The QoS requirements for each traffic type in LTE femtocells ................. 146
5.4 Simulation environment parameters for cognitive interference mitiga-
   tion in LTE ................................................................................................. 159
6.1 5G Hetnets environment and online scheme simulation parameters .......... 175
6.2 5G multi-tier environment simulation parameters ...................................... 201
6.3 5G Hetnets environment simulation parameters ........................................ 226
6.4 Parameters of the 5G H-CRANs system in Testbed Implementation ........... 255
6.5 Parameters of the 5G H-CRANs system in Simulation ............................... 261
6.6 Testbed and numerical results comparison ............................................... 267
Chapter 1

Introduction

Modern wireless networks including the upcoming 5G are required to be robust, flexible and energy efficient, but provide high-quality and low-cost services to the users. The demand for mobile services and applications is rapidly saturating wireless spectrum capacities. In particular, license-exempt bands such as Industrial, Scientific and, Medical (ISM) bands are experiencing increased channel demand and contention, which lead to spectrum scarcity. 5G with its proposed trends is envisioned as a potential solution to accommodate this demand. Moreover, techniques such as deployment of multi-antenna, multi-carrier techniques and advanced power control are considered in the design of future network topologies. However, with all these attempts to handle this demand, still better spectrum and power efficiency remain as high priority goals for the future network designs.

1.1 Problem Statement and Motivation

The flexibility of future 5G mobile networks exploiting modern technologies such as cloud-optimized radio access and software-defined radio open the gateway for deploying dynamic strategies for better resource allocation. Dynamic Spectrum Access (DSA) was proposed to solve the spectrum scarcity problem by improving spectrum utilization of the unlicensed spectrum as in [1, 2]. This is achieved by allowing Secondary Users (SUs) to opportunistically access the licensed spectrum without interfering with Primary Users (PUs). Same approach is applied to heterogeneous networks,
where small cells share the spectrum with the macro cells without harming their communications. Nevertheless the main drive of Cognitive Radio (CR) was to improve spectrum utilization, it can go beyond that. According to Mitola’s vision, CR is a system, which is self-aware, can observe and plan according to the stimuli from the radio environment, learn from past actions and act accordingly \[3, 4, 5\]. This vision requires a design of a fully independent Cognitive Engine (CE) that is able to acquire environment conditions and applications’ requirements and utilize them to make decisions on transmission parameters adaptation. However, constrained optimization that is determined by application-layer requirements, and channel conditions acquisitions in heterogeneous topologies are challenging tasks. As Artificial Intelligence (AI) techniques are the core of the CE as they are employed for decision-making, it is necessary to address the tradeoff of different AI algorithms and the complexity issues. In addition, the fact that CE with certain AI technique can perform well at certain environment conditions such as high Signal to Noise Ratio (SNR), while another CE is more effective at low SNR. Therefore, the design of CE is a challenging task.

5G is a promising technology to satisfy the future demand for data services as it is expected to provide high data rates up to 10 Gbps with end to end latency of 2 to 5 milliseconds \[6\]. The vision of 5G networks is to have a global unified platform that provides seamless connectivity among existing standards (e.g., HSPA, LTE-A, and WiFi). One of the envisioned 5G structures is the multi-tier Heterogeneous Networks (Hetnets) with different sizes, transmission powers, and unprecedented numbers of smart and heterogeneous wireless devices \[7\]. The multi-tier structure consists of two tiers: primary tier and secondary tier. The primary tier includes high power macrocells that serve Macro Users (MUEs), while the secondary tier comprises picocells, femtocells and Device to Device (D2D) communications. However, small cells and D2D transmitters tend to increase their transmission power to maximize their performance, which causes severe interference to the primary tier and increases the
power consumption [8] [9]. Due to the conflict of interests among the Secondary Transmitters (STs), it is more suitable to address the power allocation problem in a non-cooperative fashion. This also reduces the overhead of either appointing central entity for information broadcasting and information exchange among the STs. However, the noncooperative approach may cause severe interference, increase in the power consumption, and degradation in Quality of Service (QoS) of the MUEs and the SUs due to the lack of the environment awareness [10]. Heterogeneous Cloud Radio Access Networks (H-CRANs) is another emerging trend of 5G that aims to overcome the capacity limitation of Cloud Radio Access Networks (CRANs) by decoupling data and control signals in order to alleviate the influence of fronthaul links on energy efficiency and dedicate small cells to provide high data rates without considering control functions [11]. On the other hand, high power Macro Base Stations (MBSs) in Hetnets support coverage and guarantee backward compatibility with traditional cellular networks since small cells focus only on boosting the data rate in special zones [12] [13]. Despite the fact that Hetnets are able to improve the coverage and the capacity, inter-tier interference and the cumulative power consumption of the small cells are critical challenges that must be considered [14] [15]. Comparing with CRANs and Hetnets, H-CRANs have been shown to exhibit significant performance gains though advanced collaborative signal processing and radio resource allocation are still challenging. Inter-tier interference between the MBSs and Remote Radio Heads (RRHs) has a severe impact on energy efficiency. Unlike traditional CRANs, inter-tier interference should be controlled by an advanced processing technique and the interference to the MUEs must be maintained at low levels with sophisticated power allocation techniques. The intra-tier interference between the RRHs is another factor to degrade energy efficiency in H-CRANs in addition to the inter-tier interference because of the fronthaul capacity constraints. In addition, the users usually prefer to associate with RRHs because lower transmission power is needed and more resources are allocated
compared to association with MBSs. This is reflected on energy efficiency of RRHs.

1.2 Thesis Objective and Contributions

This thesis aims to investigate radio Resource Allocation (RA) problem in various network structures including CR and the upcoming 5G. AI techniques with focus on machine learning are the tools toward achieving efficient RA in these networks driven by various performance objectives such as energy efficiency, interference mitigation and throughput maximization. The work in this thesis realizes CR as a practical technology to perform transmission parameters adaptation according to the environment conditions. This is demonstrated using multiple AI techniques in various networks topologies. In addition, machine learning is employed to allocate radio resources including power and frequency in potential 5G trends, which are Hetnets with D2D communications and H-CRANs.

The thesis also exploits customized AI techniques to design a cognitive resource management system (CogWnet) to perform radio parameters adaptations according to environment conditions. The core of this system is developed using single, hybrid and supervised AI technique approaches. The motivation and tradeoff for each approach is highlighted. The designed system is applied for resource management for several radio technologies including LTE systems and its impact on boosting performance and network efficiency is demonstrated. In addition, we develop an enhanced online learning approach with low complexity to tackle the RA problem including power, frequency and small cells operation mode in 5G Hetnets and 5G H-CRANs. The performance of the designated online learning approach is highlighted over other RA schemes in terms of convergence, energy efficiency and spectrum utilization with minimum complexity.

The main contributions of this thesis are summarized as follows:

- Background and State of Art. Survey and background about CR resource man-
agement starting from the development of Software Defined Radios (SDRs) to the development of cognition cycle are illustrated. The design of CR resource management architecture and the challenges that accompany the existing approaches to tackle them and the corresponding practical solutions implemented are explained. CE, several approaches from the literature to highlight its capability, and the application of CR resource management concept in real network technologies such as LTE networks are presented. RA problem with various network efficiency objectives in 5G Hetnets and 5G H-CRANs is presented along with the literature related work.

- **Cognitive radio resource management architecture (CogWnet).** CogWnet [16] is proposed to realize CR as per Mitola’s vision. CogWnet addresses the challenges raised in CR resource management and coordinates between the cognitive functions to resolve any conflicts. A cross-layer based optimization was adopted, which infers and extracts environment parameters from all layers of the network stack for use in the decision-making process. The designed architecture is component based and addresses adaptability, portability and modularity issues. We have demonstrated CogWnet functionality using SDRs testbed.

- **Artificial intelligence approaches for the decision-making function of cognitive architecture**

  This contribution comprises the developed approaches for CE using AI to perform radio system parameters adaptation. These approaches include single, hybrid and supervised cognitive engines. The single CE approach exploits enhanced version Genetic Algorithm (GA) with constrained and adaptive multi-objective optimization [17] [18]. The hybrid engine approach comprises CBR and Decision Trees (DTs) to achieve parameters adaptation with the goal of limiting the complexity and reduces the convergence time [19]. The supervised
approach tackles the tradeoff of using certain machine learning technique for parameters adaptation in different scenarios. Thus, it aims at selecting the most appropriate machine learning technique for the any encountered scenario [20].

- Cognitive radio resource management for LTE networks CR resource management is utilized to allocate Resource Blocks (RBs) and adapt related transmission parameters to maximize throughput [21]. The process incorporates integration between the LTE architecture and the designated cognitive system (CogWnet). In addition, a cognitive approach is proposed to mitigate interference in LTE Hetnets between femto and macro cells [22].

- Efficient RA in the next generation 5G networks using novel machine learning approaches. Two online learning approaches are proposed to perform resource allocation in 5G Hetnets. The first approach is cooperative online learning scheme to allocate power and frequency to maximize the users data rate in the downlink [23]. The second approach aims to maximize energy efficiency through efficient power allocation [24]. It develops a non-cooperative online learning with an intuition feature that allows each learning agent to conjecture other agent intended actions for power allocation, which reduces the complexity and enhance the achieved performance. In order to improve energy efficiency further in 5G Hetnets, a traffic offloading methodology is developed using online learning in which MUEs are offloaded to small cells to reduce power consumptions. Finally, a sophisticated online learning scheme is proposed for RA in H-CRANs to maximize energy efficiency. The proposed scheme targets the downlink communication and follows two approaches: centralized, where a controller is dedicated to make the RA decisions and decentralized where MBSs take the responsibility of resource allocation [25]. The developed online learn-
ing methodology approximates the Q-Value function to reduce the complexity of the proposed scheme [26] and expedite the learning convergence.

1.3 Thesis Organization

This thesis is divided into following chapters. The state-of-the-art technology is presented in Chapter 2 along with an overview of the related work. CogWnet system design, components and its decision-making functionality are presented in Chapter 3. In Chapter 4, different AI approaches for the design of CE for cognitive parameters adaptation are presented. In Chapter 5, resource allocation and interference mitigation techniques inspired by CR in LTE networks are proposed. Chapter 6 investigates the problem of resource allocation in 5G Hetnets and H-CRANs using enhanced machine learning techniques. In addition, traffic offloading using machine learning is proposed in the same chapter to address the problem of energy efficiency in 5G Hetnets. Finally, conclusions and future works are presented in Chapter 7.
Chapter 2

Background and State of the Art

The concept of CR was introduced by Mitola in [3]. Mitola aims to design a fully reconfigurable radio that is able to adapt its parameters according to the current environment state and users’ QoS requirements. This opens the path to exploit cognitive approach in radio resource management. AI algorithms are the tools used to develop efficient CR resource management schemes. In addition, these algorithms are extended to tackle RA problems in advanced network structures such as cellular and Hetnets.

In this chapter, we present background and relevant work that investigates radio RA in various network structures. First, an overview of CR resource management is given. Then, we elaborate on the concept of CE and how AI techniques are utilized to design cognitive engines. In addition, RA challenges in LTE and 5G networks, related interference issues, and the work done to overcome these challenges are presented. Then, we discuss the motivation for using of online learning to perform an efficient RA in the upcoming 5G with various trends including Hetnets and H-CRANs. Moreover, we give a brief background about online learning functionality.
2.1 Cognitive Radio Resource Management

2.1.1 Software Defined Radios Evolution

The concept of SDRs was introduced by Mitola in 1991 [27] referring to a radio, which can be easily reconfigured and reprogrammed. The ability to change the radio configuration gives the device the flexibility to perform variety of functions at different times. The technical definition for SDRs is that they are software implementation of the radio functionality, which allows the mobile terminal to adapt the radio environment accordingly.

The flexibility of communication offered by SDRs imposes an implementation difficulty. For example, designing an SDRs front end to handle wide range of frequencies and modulation types is extremely difficult [28]. Developing reconfigurable radio architectures was stimulated by the necessity of the U.S military to have multi-band multi-mode radio, which can operate across different frequency bands using variety of protocols. One of the first such radios built was the US Air Force’s Integrated Communications Navigation and Identification Avionics (ICNIA) system, developed in the late 1970’s [29]. The system used Digital Signal Processors (DSPs) based modem to operate multi-function multi-band airborne radios in the 30 MHz to 1.6 GHz band. ICNIA’s technology has been later the foundation for many other military radios.

In addition, several universities groups and companies have designed and built their own experimental SDRs platforms. For example, Spectrum Signal Processing Inc. is a company that is specialized in the development of SDRs. They provide a wide range of SDRs systems and baseband processing boards that utilize a combination of PowerPC, DSPs and Field Programmable Gate Array (FPGA) signal processing devices [30]. Motorola built their 4G base station using SDRs technology and has been one of the first manufacturers deploying SDRs in its products since 2001 [31]. Vanu
produces software for SDRs. They have designed their own software radio architecture which enables base stations to simultaneously operate as Global System for Mobile Communications (GSM), Code Division Multiple Access (CDMA), Integrated Digital Enhanced Network (IDEN), Wideband Code Division Multiple Access (WCDMA), and beyond.

One of the most popular and widely used SDRs platform in the research community is the Universal Software Radio Peripherals (USRP) designed by Ettus Research [32]. It is an integrated board, which incorporates Analog to Digital Converters (A/D) converters, a number of Radio Frequency (RF) front-ends covering different frequency bands, and an FPGA which does a part of computationally most expensive pre-processing of the input signal. The low-cost and relatively high speed made the USRP board the best choice for a GNU Radio user to implement some real time applications. It is also very often known simply as a GNU radio platform as it is mainly used in combination with the open-source GNU radio software [33]. The GNU radio software itself is built in a modular fashion. It provides a library of signal processing blocks and the framework to tie these blocks together. The programming environment is a combination of C++ and Python.

Many research groups have implemented their own SDRs platforms. The KUAR platform has been designed to be a low-cost experimental platform targeted at the frequency range 5.25 to 5.85 GHz with a tunable bandwidth of 30 MHz [34]. The platform includes an embedded 1.4 GHz General Purpose Processors (GPPs), Xilinx Virtex2 FPGA and supports Gigabit Ethernet and PCI-express connections back to a host computer. Another platform is Wireless Research Open-Access Platform (WRAP) developed at Rice University is one of the most used reconfigurable research platform to prototype high-speed wireless communication algorithms and systems. Due to its programmability and flexibility it allows for easy implementation of various physical and network layer protocols. Similarly but more powerful and expensive
reconfigurable platform called Berkeley Emulation Engine (BEE2) has been developed at Berkeley Wireless Research Center [35]. The engine consists of five Xilinx Vertex-II VP70 FPGAs connected with high speed internal links, which provide a possibility to execute in parallel intensive signal processing algorithms. One of the five FPGAs called a control FPGA executes Linux OS on its embedded PowePC to control the peripherals. links are available per platform to interface the BEE2 with radios a total of 8-10 Gbps full-duplex. Due to its modular design and scalability BEE2 is applicable to a wide range of high-performance applications such as real-time radio telescope signal processing, CR, computer architecture emulation and so on.

For Our implementation, we have used USRP-N210 [32] which is the latest version of USRP series SDRs platform.

2.1.2 Cognitive Resource Management Concept

Mitola’s vision was to have a radio that is self-aware, user-aware, RF aware and able to learn from the surrounding environment. This leads us to the concept of radio etiquette, where the radio is able to reason about a set of RF bands, air interfaces, protocols spatial, and temporal patterns of spectrum usage. Mitola has designed radio knowledge representation language to express knowledge that CR has about itself and the environment around it [1].

The potential view of CR helps to introduce future wireless networks that are flexible and exploit AI algorithms for radio reconfiguration optimization. The definition of CR was left to the research interest of wireless research groups. For example, for the groups, who work on Dynamic Spectrum Access (DSA), they consider it a definition for CR. However, it is only one part of CR, which is related to RF awareness and spectrum agility. Radio systems that can do context based routing, automated channel allocation, adaptive modulation or in general can exhibit any kind of self-configuration have been also defined as CRs, by system oriented researchers. Therefore, CR can
be classified into two types: the first one is specifically doing DSA and the second one follows the idea of having a fully configurable radio system, which can learn and decide based on the environmental stimuli and intelligently optimize its performance. As a formal definition for CR, The Federal Communications Commission (FCC) proposed in 2003 the following term for Cognitive Radio: CR is a radio that can change its transmitter parameters based on its interaction with the environment in which it operates. This interaction may involve active negotiation or communications with other spectrum users and/or passive sensing and decision making within the radio. The majority of CRs will probably be SDRs, but neither having software nor being field-reprogrammable are requirements of CR [36].

2.1.3 Cognition Cycle

Cognition cycle was introduced by Mitola in [3]. It illustrates the overall cognitive process as in Figure 2.1. It starts with observing the radio environment which is achieved via spectrum sensing. Spectrum sensing function aims to find spectrum opportunities, finding out n-hop neighbors, getting channel state and network state information, location information, etc. The observation also includes reading the current radio and protocol level settings as well as getting information from other sources.

![Figure 2.1: Cognition Cycle](image)
(cooperative users, databases, etc.) via a control channel. The information obtained in the observing state can be used in both learning and orientation. The orientation state comprises action scheduling and priority setting in case an immediate action is necessary. Transition can go from the orientation to the decision state in urgent scenarios. The last possible transition from the orientation state goes to the plan state for scenarios that are not time critical and need planning before taking actions. Learning state is the center of the cognition cycle as it is the process of extracting the useful information from the observations, passed decisions and actions and also being able to recognize the changes in the environment and adapt accordingly. The action state results in optimal RA and parameter settings such that the performance of the radio device or the network is optimized under certain conditions and preferences.

2.1.4 Cognitive Radio Resource Management Architectures

Several examples of cognitive architectures have been reported [37], [38]. They either focus on design optimization while ignoring issues that impact the overall system performance, or only consider DSA as the main objective. They use proprietary interfaces for communication between stack layers and the cognitive architecture. Moreover, no standard approach exists to configure the transmission parameters. There are a few high-level architectures that target resource management. However, often these lack a testbed evaluation that can determine their efficiency. For example, the design of a CE in [39] aims to learn and adapt to radio environment changes by extending legacy spectrum-sharing techniques. This engine lacks certain functions, such as the use of broadband waveform, reliability, and support for QoS requirements. The work in [40] uses GA for configuration of physical and MAC layer parameters, such as modulation, transmit power, number of sub-carriers, etc. The system consists of an engine that adjusts the parameters according to environmental constraints. However, this engine only considers physical and MAC layers constraints, thus limiting reconfigurability.
In addition, this architecture is very complicated as it uses GA without considering its complexity. Another design of a reconfigurable node was proposed in [41] to observe and reconfigure all radio parameters. The design is simple and flexible but does not involve the application layer requirements in the channel-selection process. In [42], the cognitive framework considers the interference caused by the coexistence of PUs and SUs. It is based on collecting channel information from the physical layer and using it to reconfigure the radio. The MAC layer includes the management and channel-sensing modules. However, this architecture lacks the involvement of the upper layers in the radio-management process, making the resultant decisions inconsistent with the QoS requirements. All of the above-mentioned architectures do not specify interfaces enabling the flow of information between the CE and the stack layers. In addition, there is no support for distributed cognitive systems that improve the overall performance of the network.

2.2 Cognitive Engine and Radio Parameters Adaptation

2.2.1 Cognitive Engine

CE is the intelligent agent that performs decisions driven by certain performance objectives to adapt system parameters, according to its observation and learning from the surrounding environment to achieve reliable communication and efficient resources utilization [17]. CE incorporates three attributes:

- Observation: Collect information about the operating environment, capability, and characteristics of the radio.

- Reconfiguration: Change the operation parameters of the radio.

- Cognition: Understand the environment and capability of the radio (awareness), make informed decisions on actions (reasoning), and learn the impact of these
actions on the performance of the radio, as well as the performance of the network in which the radio is embedded (learning).

Awareness refers to the process of extracting the information regarding environment and radio itself for a specific purpose. Reasoning is defined as the process of finding an appropriate action in response to a particular situation with a system target (e.g., maximum operating lifetime, maximum robustness, and lowest cost communications) based on the user application quality-of-service (QoS) requirement [e.g., latency and bit error rate (BER)] and willingness to share resources and collaborate with other devices in the network. Learning is defined as the process of accumulating knowledge based on the observed impact upon applying the action. Typically, these processes complement each other to improve the operation of the CR process as a whole [43].

The CE can be implemented as an independent entity interacting with the radio transceiver [e.g., reconfigurable radio transceiver implemented with SDRs or as a collection of interacting entities with each entity fulfilling a specific role. Given the input from its environment or user (observation), the CE analyzes and classifies the situation and determines the appropriate response to the stimulus (cognition) and carries out the decision (reconfiguration). As an example, this response can be adapting radio parameters such as the channel coding scheme, the modulation scheme, and the operation frequency, given user requirements, current environment conditions, and previous experiences at the CE [44]. How to implement the various aspects of a CE is an active area of research in which considerable attention is given to each of these three attributes.

2.2.2 Cognitive Radio Parameters Adaptation Techniques

Radio resource management and system parameters adaptation relies on application QoS requirements such as throughput, delay, Bit Error Rate (BER), and energy efficiency dictated by the application layer. The design of CE utilizes many AI tech-
niques to be able to establish automated wireless systems that is capable to configure radio transmission parameters according to environment conditions extracted from the TCP/IP layers including channel state information, received signal strength and Signal to Interference and Noise Ratio (SINR).

There are many AI techniques exploited for radio resource management and transmission parameters configuration schemes in the cognitive context, such as Rule Based Systems (RBS) [45], fuzzy logic [46], and Artificial Neural Networks (ANNs) [47]. These approaches can find solutions but may not produce the best ones as they have various limitations on parameters adaptation. For example, RBS are limited to predetermined capabilities by their own rule set. Fuzzy logic approximated solutions are not based on certain input, and hence, they are not accurate solutions. ANNs include extensive training to generate observed behavior, but become unstable when constraints are necessary to account for. Ant Colony Optimization (ACO) algorithm is used to solve the optimization problem by formulating it into single objective function [48]. Particle Swarm Optimization (PSO) is also being used as in [49] to determine the CR transmission parameters for a given set of objectives. Game Theory (GT) is in a nascent stage and used for interactive decision-making, but provides analytical tools to predict the outcome of complex interactions among the rational entities based on perceived results. In [50], Simulated Annealing (SA) cross-layer optimization is exploited to improve communication quality based on utility function. Transmission parameters are tuned based on the utility function calculation for certain communication quality metrics such as throughput, delay...etc. The aim of this optimization is to maximize these utility values. Although SA is fast in convergence, it does not improve with time, which is necessary as more search is required to find solutions in such dynamic environment with plenty of parameters and multiple objectives. The dynamic nature of real networks makes the implementation of these techniques in CR very complex. In addition, the heterogeneity of CR networks and the multiple layers
involved in radio configuration exacerbate difficulties and drawbacks in the above optimization techniques.

An adaptive multi-objective optimization schemes was proposed in [17] and [18] to improve the system efficiency with an enhanced version of GA. Despite the power of GA in multi-objective optimization, its complexity and the relatively slow convergence limit its performance. Simulation results in [51] have shown that the PSO can solve multi-objective optimization problems and dynamically select transmission parameters in CR applications. The work in [52] considered ACO in the CE design. However, long-term learning ability, which is essential function of CR to gather knowledge from its past running experiences is not investigated in these engines. Learning based algorithms such as CBR was exploited to create cognitive engines. The work in [53] exploited CBR in cognitive system adaptation in IEEE802.22 networks. CBR is fast in convergence with good scalability. However, it is not efficient if the encountered scenario does not match with any of the cases stored in its database. In order to overcome the limitations of each of the above techniques, hybrid CE design was introduced. This approach aims to combine learning and optimization algorithms to perform more efficient system adaptation. Authors in [54] exploits the advantage of quantum GA to design a hybrid engine with CBR. Another hybrid approach was investigated in [55] in which CBR and PSO are used as the core of the CE. Ashwin et al. [56] proposed a hybrid engine based on CBR and GA that has the capability to still adapt to new environments using GA. However, this engines relies on GA as an optimizer, which is slow in convergence and is stuck in local optima. In addition, none of proposed engines considered efficiency, complexity and configurability range.

In [57], Huang et al. designed a learning engine framework based on Support Vector Machine (SVM) to configure radio parameters using the estimation of BER and SNR. The work in [58] proposed a primitive architecture for meta-cognition, which is used to rate the solution achieved by CBR and from this point it can decide if it is
necessary to look for alternative adaptation algorithm such as GA. Y. Zhao et al. [59] looked into utility function selection for streaming video with a CE testbed. Our work in [17] considers a design of CE using multiple objective optimization for parameters adaptation with enhanced GA as an optimization tool. We in [19] proposed a hybrid CE to perform system adaptation using CBR and DTs to boost the performance and expedite the convergence. It is fact that CE with certain learning technique can perform well at certain environment conditions such as high SNR, while another CE is more effective at low SNR. This creates the need to redesign the CE structure from a new perspective, which is the capability to select the most appropriate CE technique for adaptation in the encountered network scenario. However, this requires establishing solid evaluation methodologies to be able to evaluate various CE designs and algorithms in different network conditions.

2.3 Cognitive Resource Allocation in LTE Networks

2.3.1 Cognitive Radio Resource Management for OFDMA Systems

As the aim of LTE networks is to provide higher data rate with lower latency, Orthogonal Frequency Division Multiple Access (OFDMA) is used as its access technology where radio resources are scheduled in the unit of RBs [60]. Each RB consists of successive sub-carriers over a certain number of OFDMA symbols. Most of RRM schemes are compatible with OFDMA [61]. Radio resource management is a challenging task because radio resources such as frequency and spectrum are limited and rapidly changing. Resource management in LTE networks involves different technologies to cope with the increasing demand for cellular data services, making it complicated and difficult to understand and develop. Thus, the designed RRM scheme must be flexible while keeping its complexity low. TCP/IP stack layers have strict boundaries to keep their data, thus requiring sophisticated cross-layer optimization to access the stack layers data and collect radio environmental information. In addition,
the co-existence of numerous radio access technologies necessitates the deployment of multiple RRM modules. These modules communicate with the radio environment and other terminals’ modules to share the selected configuration of transmission parameters. In such network, it is difficult to assign radio RBs to each user to satisfy QoS requirements such as packet loss and delay. Multi-path channel fading is another problem in LTE deployment. Therefore, a sophisticated RRM architecture is needed to manage spectrum allocation. Co-channel deployment is used in practice where all cells share the available spectrum with the ability to change frequency. Interference is another challenge when the spectrum is shared between macro and femto cells as in LTE-A networks. Joint Radio Resource Management (JRRM) is one of the proposed solutions for RRM in LTE. This solution allows co-existence of multiple Radio Access Technologies (RATs) and the integration of LTE and other wireless network standards. It is based on Reinforcement Learning (RL) for channel allocation. JRRM employs a dedicated agent in each cell to distribute users among different technologies. The focus of this solution is only on channel allocation without considering other transmission parameters optimization. Another solution is Multi Radio Resource Management (MRRM), which consists of radio and mobility managers. These managers allocate channels, manage handover, and optimize load-balancing over different networks. MRRM incorporates resource management at terminal and network levels. Dynamic Fractional Frequency Reuse (FFR) is exploited for RRM in LTE to overcome the interference problem. It uses different frequency reuse factors based on the distance between the terminal and the center of the cell. High reuse factors are assigned to the terminals close to the center as they experience less interference, and vice versa. This scheme adjusts the transmission power of the terminals based on the interference information received from the downlink. All of the above mentioned schemes are based on predetermined algorithms for channel allocation. However, less attention is given to QoS support or network performance
metrics, such as throughput.

### 2.3.2 Cognitive Approaches for Interference Mitigation in LTE Heterogeneous Networks

The coverage issues of the macro cells networks makes them unable to fulfill the data services demand in the indoor areas. An efficient solution is to deploy FBSs which are capable of communicating users over a broadband wire-line connection \[\text{[62]}\]. FBSs are short-range, low-cost/low power and can be easily installed by the users in addition to the fact that they reduce the load of the MBSs. However, interference is considered as a technical challenge that affects the femto cells deployment \[\text{[62]}\]. There are two types of interference: cross-tier and co-tier. Cross tier is the interference to the MUEs caused by the FBSs installed within the same Sub-Band (SB). Co-tier interference is the one among the deployed femto-cells contending for the same channel. These types of interference lead to service disruption, throughput degradation and connection droppings. There are several proposed schemes for resource allocation in femto cells deployments with interference consideration. For example, the schemes proposed in \[\text{[58]}\] and \[\text{[59]}\] aim to handle both types of interference using uncoordinated and coordinated resource assignment algorithms as in \[\text{[58]}\], and Q-learning based interference coordination as in \[\text{[67]}\]. However, the coordination between MBSs and FBSs is difficult due to the requirements of scalability, security, and the availability of backhaul bandwidth in addition to the fact that the number of the deployed FBSs is not fixed. The authors in \[\text{[68]}\] and \[\text{[69]}\] propose a scheme that assigns dedicated channels for the communication of FBSs over the uplink and the downlink. This goes against the idea of improving spectrum utilization by accessing the macro cells spectrum opportunistically. Femto cells resource allocation mechanisms are investigated in \[\text{[70]}\] and \[\text{[71]}\] to mitigate interference. These mechanisms use CR and GT to support their RA methodologies. However, both schemes are limited to channel
allocation without consideration of QoS requirements and the scheme in \cite{71} considers cross-tier interference only. The work in \cite{72} aims to maximize the weighted sum rate of the femto-macro network in a delay tolerant scenario. However, this requires high information overhead among MBSs and FBSs. FFR technique proposed in \cite{73} shows capability to mitigate interference in multiple cells deployments. However, the cell edge users suffer from lower data rates because of the increase in path-loss and interference \cite{74}. The FFR strategy has been widely used in multi-macrocell environments for suppressing interference as in \cite{75}. The use of CR \cite{76} in femto-macro deployment effectively contributes in solving the cross-tier interference problem by exploiting its spectrum sensing capability to allocate under utilized channels. In addition, it is considered for spectrum assignment in order to increase the flexibility and the autonomy of the network in addition to interference mitigation \cite{77}.

2.4 Efficient Resource Allocation in 5G

This section includes background and literature work toward efficient 5G networks including multi-tier Hetnets and H-CRANs.

2.4.1 Efficient Resource Allocation in 5G Heterogeneous Networks

One of the envisioned 5G structures is the multi-tier network with different sizes, transmission powers, and unprecedented numbers of smart and heterogeneous wireless devices \cite{78}. The multi-tier structure consists of two tiers: primary tier and secondary tier. The primary tier includes high power macro cells that serve MUEs, while the secondary tier comprises picocells, femtocells and D2D communications. The STs including pico and femto Base Stations (BSs), and D2D transmitters. They utilize the available resources (e.g., bandwidth and power) in an underlay mode as long as the interference caused to the macro tier remains below certain threshold. However, they tend to increase their transmission power to maximize their performance, which
causes severe interference to the primary tier and increases the power consumption

Although the multi-tier approach in 5G is promising to improve the performance of the network in order of magnitudes compared to the legacy cellular networks, it may lead to a significant interference between the primary tier and the secondary tier and also between the secondary tier devices [81]. This interference impacts energy efficiency and degrades the QoS experienced by all users. STs tend to increase their transmission power unnecessarily to overcome interference and this leads to significant sacrifice in energy efficiency. The interference is a crucial problem in 5G due to the following reasons: heterogeneity and dense deployment of wireless devices, various transmission powers of different transmitters, which may cause imbalance in the traffic load and coverage, public or private access restrictions in different tiers that lead to diverse interference levels and priorities in accessing different portions of spectrum plus the impact of carrier aggregation and D2D communications. Therefore, sophisticated power allocation mechanisms are necessary to account for the interference problem and enhance the system performance, which consequently reduces the power consumption and maintains QoS for different tier users [82] [83] [84] [85].

Power allocation problem in the multi-tier heterogeneous environment has become an interesting topic in the current research of wireless communication. Authors in [86] proposed a utility based power adaptation algorithm to mitigate the cross-tier interference at the macrocell from the femtocells. In [87], authors proposed game-theoretic framework in Hetnets, which enables both the small cells and the macrocells to strategically decide on their downlink power control policies. They formulate the power allocation problem as a stackelberg game to maximize the data rate of each cell. The work in [88] proposed a hierarchical game theoretical framework for optimal resource allocation on the uplink of Hetnets with femtocells overlaid on the edge of a macrocell. Authors in [89] summarized the challenges and opportunities to
improve energy efficiency while increasing the network capacity in both multi-RATs and single-RAT Hetnets. The downlink resource allocation problem in OFDMA Hetnets consisting of macrocells and small cells sharing the same frequency band was investigated in [90]. The authors aim to devise an energy efficient scheme that allows shared spectrum access to small cells, while ensuring a minimum level of QoS for the macro users. Hetnets based on large-scale user behavior was proposed in [91], where the heterogeneity of large-scale user behavior is quantitatively characterized and exploited to study the energy efficiency performance. The authors in [92] proposed a new resource allocation scheme presenting a low computational overhead and a low sub-band handoff rate in a dynamic ultra-dense Hetnets. The work in [93] addressed the energy efficiency optimization problem for downlink two-tier Hetnets, where the power allocation problem is decomposed into multiple optimization problems with single inequality constraint. Those optimization problems are solved using a sub-optimal solution based on the zero forcing precoding approach. An energy efficient radio resource allocation algorithm was proposed in [94] for interference management, maximization of throughput and energy efficiency to enhance the performance of a heterogeneous deployment of ultra-dense femto-cells overlaying the macrocells. The proposed scheme exploits CR technology and stochastic process to perform RA. The authors in [95] proposed a multiuser MIMO precoding scheme that is capable to reduce the negative impact of interference in Hetnets. The formulated optimization problem is solved using heuristics based techniques. Both works in [93] and [95] are cooperative and rely on information exchange among STs. The work in [96] aimed at network utility maximization via jointly optimizing user association, resource allocation and power control in a load-coupled Hetnets. The authors in [97] addressed a non-convex energy efficient optimization problem with resource assignment and power allocation for the OFDMA H-CRANs. They found closed-form expressions for the energy efficient resource allocation solution to jointly allocate the RB and trans-
mit power. RL \[64\] was considered as suitable solution for interference management and power allocation in such autonomous environment. An RL based algorithm was proposed to optimize the network performance by managing power allocation in femtocells in \[98\]. However, none of the above work consider 5G structure where D2D is involved and the network is dense with various types of small cells deployed i.e. (all of the previous work considers only one type of ST).

2.4.2 Energy Efficient Base Stations Operation in 5G Heterogeneous Networks

As dense Hetnets comprising macrocells and small cells are one of the main trends in the future 5G to enhance the throughput of cellular networks at relatively low operational costs \[62\] \[99\], the perceived increase in energy consumption is expected to be about 40% for wireless cellular networks from 2010 to 2020 \[100\]. Moreover, lack of coordination of small cells and high operational costs of macrocells reflected by the energy consumed for the operations are crucial challenges that may limit the density of future 5G. Traffic offloading with the aid of small cells is a promising solution to overcome the traffic congestion problem and improve the overall network energy efficiency. It aims at allocating more capacity for services while maintaining QoS for the users \[101\] and boosting the energy efficiency. However, traffic offloading in a multi-tier Hetnets, where small cells are activated for handling the offloaded traffic creates a new source of interference between the macrocells and the small cells. Without maintaining the BSs system load, offloading traffic from macrocells to small cells causes severe interference between macrocells and the activated small cells. Moreover, it increases the energy consumption across the whole network.

As there are increasingly more small cells deployed in the 5G cellular network, their power consumption is not ignorable. The authors in \[102\] showed that the typical power consumption of a small BS is 10W and that of a macro BS is 930W. Thus,
the power consumption of 100 small BSs is larger than one macro BS. Therefore, it is important to manage the power consumption of the macrocells and small cells for energy efficient communication [103]. Recently, power savings in heterogeneous 5G draw the attention of many researchers. The work in [104] proposed a cell zooming approach to reduce the macrocell power consumption by adjusting the cell size according to the covered traffic load, the QoS, and the channel conditions. The authors in [105] showed that 25 to 30% of the total power can be saved by reducing the number of active macrocells when traffic is low. Stochastic geometry was exploited to design an optimal ON-OFF BS adaptation scheme in [106]. The work in [107] proposed a multi-objective framework as an energy and cost-efficient solution for the resource allocation problem in Hetnets, and provided extensive analytical and experimental results to estimate the potential energy and cost savings that can be achieved. In [108], the authors considered small cells activation strategy but they did not shed the light on energy efficiency or interference caused by small cells activation. The work in [109] surveyed and compared the primary technical approaches to Hetnets load balancing: centralized optimization, GT, Markov Decision Processes (MDPs), and the newly popular cell range expansion. The studies in [110] and [111] further exploited small cell sleeping potential and proposed small cell control algorithms for power saving. In [112], authors studied the impact of implementing sleep/awake mechanism on energy efficiency. RL based scheme to intelligently offload traffic in a stochastic macrocells was presented in [113]. Interference and RA problem in heterogeneous cellular networks was tackled in [114]. To improve the energy efficiency for macro users, a traffic offloading algorithm based on the metropolitan advanced delivery network architecture was proposed in [115]. The data traffic is offloaded to a WiFi access point as long as transmitting the same volume of data consumes less energy in the WiFi transmission than using the cellular network. A prediction based traffic offloading protocol was presented in [116] for offloading large contents from cellular networks.
to save energy for macro users. Authors in [117] proposed an adaptive link selection algorithm to reduce the energy consumption by transferring large volume of data from the phone to the infrastructure. The authors of [118] proposed a reverse auction based offloading that targets achieving tremendous energy efficiency under the constraints of QoS requirements. The optimization problem with constraints is solved by dynamic programming method with Karush-Kuhn-Tucker (KKT) conditions. However, none of the above schemes considers the exploitation of the cell system load to determine the optimal offloading strategy in the 5G heterogeneous environment and they do not emphasize load maintenance of small cells. In addition, none of the RL based schemes considers the challenge caused by the growth in the number of small cells and how to benefits from other BSs offloading strategies.

2.4.3 Energy Efficient Resource Allocation in 5G Heterogeneous Cloud Radio Access Networks

CRANs are recognized to reduce operating expenditures, manage inter-cell interference, and provide high data rates with considerable energy efficiency performance [119] [120] [121]. It consists of rRRHs, which act as relays that forwards the users (UEs) data to the centralized baseband unit (BBU) pool for processing through wired/wireless fronthaul links. However, the fronthaul links limited capacity and long delays are limiting factors that degrade the performance of CRANs [122]. The energy efficiency of CRANs has been studied in the literature while designing resource allocation schemes. The authors in [123] formulated joint RRHs selection and power consumption minimization, subjected to user QoS requirements and RRHs power budget, as a group sparse beamforming problem. However, the proposed scheme ignored the fact that the fronthaul links have a limited capacity. The work in [124] aimed at optimizing the end to end TCP throughput performance of Mobile Cloud Computing users in a CRANs network through topology configuration and rate allocation. One
limitation of this work is that it did not constrain the capacity consumption of individual links. The authors in [125] minimized the total network power consumption in a CRANs subjected to users’ QoS for both secure communication and efficient wireless power transfer, limited backhaul capacity, and power budget constraints. However, this proposal considers single tier network only.

Although Hetnets are able to improve the coverage and the capacity, inter-tier interference and power consumption of the small cells are critical challenges that must be considered [14] [15]. The energy efficiency oriented resource allocation for Hetnets attracts the researchers’ attention in the literature. In [126], distributed power allocation for multi-cell OFDMA networks taking both energy efficiency and inter-cell interference mitigation into account was investigated, where bi-objective problem was formulated and solved using multi-objective optimization. The power allocation, RBs allocation and relay selection were optimized in [127] with the goal of maximizing energy efficiency. An optimal power allocation algorithm using equivalent conversion is proposed in [128] to maximize energy efficiency under interference constraints. The authors in [129] explore a system framework of cooperative green Hetnets for 5G wireless communication systems. To realize the 5G vision, Hetnets with massive densification of small cells and CRANs are combined in one network structure called heterogeneous cloud radio access networks to improve spectral efficiency, resource management, and energy efficiency [129] [130].

2.4.4 Online Learning Background

Online learning is a learning algorithm that uses reinforcement Q-learning [131] principles to determine a policy $\pi^*_s$ for decision-making without detailed modeling of the radio environment. This implies that Q-learning represents the performance metrics of interest and improves it as a whole. For instance, instead of tackling factors that affect network performance such as the wireless channel condition and mobility, Q-
learning monitors the feedback of its actions such as throughput monitoring. Online learning includes four parameters, which are state $s$, action $a$, probabilistic transition function from one state to the other $P_{s,s'}$, and reward function $r_{s,a}$. The state may describe internal phenomena, which are within the agent, such as instantaneous queue size, or external to the agent, such as the usage of the wireless medium. The reward function reflects the feedback for the quality of the action taken and consequently the system gains the experience. The interaction between the agent and the environment at time $t$ occurs as follows, the agent observes the environment state $s_t$. The action $a_t$ is selected based on the state $s_t$. According to $a_t$ and $P_{s,s'}$, the environment makes transition to the new state and the reward function $r_t = R(s_t, a_t)$ achieved as a result of this transition is recorded and fed back to the agent. The optimal Q-value is the metric defined for each state-action pair in the process to find the optimal policy $\pi_s^*$, and it is evaluated as follows,

$$Q^*(s, a) = E\{r(s, a)\} + \gamma \sum_{s' \in S} P_{s,s'}(a) \max_{b \in A} Q^*(s', b)$$  \hspace{1cm} (2.1)$$

where $S$, $A$ are the sets of the available states and actions respectively and $\gamma$ is the discount factor. The optimal policy can be determined by $\pi_s^* = \arg \max_{a \in A} Q^*(s, a)$. The online learning algorithm finds $Q^*(s, a)$ in a recursive manner using the following rule:

$$Q(s, a) = (1 - \alpha)Q(s, a) + \alpha (r(s, a) + \gamma \max_{b \in A} Q^*(s', b))$$ \hspace{1cm} (2.2)$$

Where $\alpha$ is the learning rate. An appropriate action is rewarded and its Q-value is increased. In contrast, an inappropriate action is punished and the Q-value is decreased. The Q-value is maintained in a two-dimensional lookup Q-table with size (state x action). It was proved in [131] that this update rule converges to the optimal Q-value when each state-action pair is visited infinitely often. Therefore, the learning comprises two processes that control its actions selections: exploitation
and exploration. Exploitation is the process of selecting the best action based on the found optimal policy $\pi^*_s$ while exploration is the process of selecting non-optimal actions randomly and discovering new ones. The exploration rate $\epsilon$ is the parameters that control the level of exploration against exploitation.
Chapter 3

CogWnet: A Resource Management Architecture for Cognitive Radio Networks

The current research aims to utilize CR as an enabling technology for intelligence, self-configuration and efficient resource allocation in the next generation of wireless networks. The ability to implement the radio configuration using software facilitates adaptive radio parameters tuning. However, the price of this facility is the complexity in finding an optimal radio configuration. In order to be able to introduce cognition and optimal self-management of resources in the future communications systems not only algorithmic solutions are needed but also new software architectures, which will support intelligent processing and modularity. The current RRM architectures such as the ones for cellular systems cannot cope with the future networks demands. Therefore, flexible, portable and easily implementable architectures are needed for RRM. In this chapter, we introduce a CR architecture called CogWnet, which has an objective of enabling efficient implementation of the cognition cycle and introducing more adaptive and flexible use of the network stack. In addition, it facilitates cross-layer optimization for radio parameters configuration. All CogWnet components are described along with its operation and performance evaluation in this chapter.

3.1 Radio Resources Management in Wireless Networks

RRM is an essential part of any recent radio system. RRM modules used in cellular networks are complicated and it is difficult to understand their dependencies
and interactions. Thus, it is unclear if these modules are able to provide fast adaptation and cross-layer optimization. Thus, the process of building a system that is adaptable to radio environment changes is complex and requires advanced control and management structures. One approach to design such system is to follow the SDRs design principles and provide modular, extensible and easier to implement framework instead of building highly integrated RRM that are technology and standard specific.

RRM principles exploited by the current radio systems such as cellular networks rely on a centralized entity like base stations to control all RRM tasks. Base stations are responsible for communication with mobile terminals within cells and it assigns certain amount of radio resources to each user to be able to communicate. However, these principles only focus on physical and MAC layers for a specific RATs and they use fixed optimization algorithms that are either proprietary technologies or often bounded by telecommunications standards. If the system performance is poor, system upgrade or human intervention is necessary.

Joint Radio Resources Management (JRRM) and Multi Radio Resources Management (MRRM) \cite{132} are invented as a multi-mode terminal RRM approaches to extend the above centralized RRM. However, these approaches have several drawbacks. First, they require almost complete redesign of RRM frameworks, and necessary extensions cannot be simply ”plugged in”. Substantially, more signaling traffic is needed to carry all the information required in the MRRM functions to the resources manager. Additional difficulties are highly non-linear dependencies between the radio resources of different technologies, especially those sharing the 2.4 GHz ISM band. The need for a priori knowledge about the interactions between the radio resources of different technologies also makes these types of monolithic systems unable to include new standards at run time, unless complete knowledge about the interactions between the legacy systems and new technology is available.

There are several cognitive architecture proposals, with various degrees of ma-
turity, emerged in the literature. In [133] [134], the authors present a three-layer software framework that can be used to implement cognitive networks. In addition, they introduced the concept of Software Adaptive Network (SAN), which provides the action space for cognitive processes. The work in [40] [135] introduced the concept of CE to exploit SDRs capabilities to optimize the configuration of radio parameters. The proposed CE uses GA as an optimization tool to configure physical and MAC layers parameters and it was tested on programmable radios [136]. Newman et al. at Kansas University [137] have proposed a CE implementation for OFDMA multi-carrier transceivers. This work is an extension of the earlier proposed flexible SDRs development platform called KUAR [34]. Another design of a reconfigurable node was proposed in [41] to observe and reconfigure all radio parameters. The design is simple and flexible but does not involve the application layer requirements in the channel-selection process. In [138], for example, an architectural framework called CogNet for integration of cognitive networks into the future Internet is proposed.

3.2 CogWnet Architecture

Since the CR paradigm is expected to drive the next generation radio devices and networks, new architectural approaches for control and management as well as new protocol and algorithm designs are necessary. The design of CogWnet [16] as in Figure 3.1 aims to realize full cognitive functionality as per Mitola’s vision. CogWnet addresses the challenges raised in RRM, spectrum allocation and learning phases. Cross-layer design approach [139] is used which infers and extracts network configuration parameters from all layers of the network stack for use in the adaptation process. CogWnet resource management tasks are explained in the following sections.
3.2.1 CogWnet Design Principles

CogWnet has been built based on the following design principles:

1. Adaptability and Flexibility: Achieved through optimization algorithms that evaluate and select suitable transmission parameters to meet QoS requirements under varying environmental conditions.

2. Portability: Allows CogWnet to act as a resource management architecture in various wireless environments.

3. Cooperative Decision-Making: The ability to negotiate the transmission parameters selected by CogWnet in one terminal with other terminals of the same network domain. This enhances the feasibility of the decision taken and avoids conflicts with other terminals.
4. Generic: CogWnet is not tied to any specific decision-making algorithm. It is a generic architecture that can accommodate various decision-making schemes.

5. Modularity: CogWnet is able to resolve conflicts between the challenge it may encounter and the function it performs.

6. Policy-Awareness: Ability to apply spectrum access policies relative to stakeholders, users, and authorities.

7. Comprehension: CogWnet includes input from multiple layers in the network stack. This optimizes the radio parameters according to the device requirements.

### 3.2.2 CogWnet Resource Management Phases

To achieve the mentioned design principles, CogWnet resources management process passes through the following phases:

1. Spectrum Sensing: Energy detection is employed \[140\] to sense the spectrum and detect channels occupancy. CogWnet implementation exploits USRP-N210 \[32\] as spectrum sniffers to check the spectrum occupancy. It is integrated with C++ data structures to implement energy detection for spectrum sensing.

2. Spectrum Allocation: This function is used to allocate channels for data transmission. CogWnet employs a spectrum allocation methodology \[141\] that is borrowed from spectrum agility principles. It has the ability to switch the operating frequency based on channels availability. Spectrum agility ensures that the allocated channels are set within an acceptable level of interference between the spectrum owners and the secondary users. Graph coloring is employed as spectrum allocation tool.
3. Parameters Reconfiguration: The objective of this phase is to comply with the available channels conditions and QoS requirements. DTs \[142\] are employed as the decision-making algorithm for parameters adaptation. The decisions aim to maximize utility functions. Utility function is a mathematical function that represents the network performance attributes such as throughput, delay, jitter, and packet loss rate.

4. Configuration Negotiation and Policy Check: In this phase, CogWnet exchanges channel conditions and the optimal transmission parameters selected during reconfiguration optimization with other nodes in the same domain. This cooperative function enhances the decision made for parameters configuration and avoids collision with the other nodes contending for the same channel. For example, CogWnet is able to determine through negotiation whether to assign the same, or two far apart channels to two contending pair of nodes. Consequently, communication can be established in a seamless manner while avoiding interference between the two nodes. Policy regulation are important to account in RRM. Therefore, CogWnet checks that the selected transmission parameters are within the confines of the policy rules before the configuration starts.

3.2.3 CogWnet Structure

CogWnet is a container-based architecture. Each layer consists of a group of components. These components are composed of containers, which exchange data through well-defined interfaces. A container is a data structure that contains a set of rules to be executed. These rules are defined as event, pre-condition, action, and post-condition. Rules statements are stored in the container as an ordered list representing their execution order. The event starts with a notification received from the communication layer upon availability of data in the container. The pre-condition is a Boolean function to be evaluated for the execution to occur and it provides a condition to
activate the containers. Post-condition is used to carry operation signals between containers (i.e. callback function). All the containers in the framework communicate using connectors. The connector is a data structure that uses read/write commands to send data between containers. The containers notify the connector about data availability for communication. In addition, CogWnet includes another set of data structure called modules. The module is a structure that consists of more than one container connected to process the data flows in the architecture.

3.2.4 CogWnet Architecture Layers

CogWnet architecture components are described in details in this section, which includes communication layer, decision-making layer, and policy layer.

Communication Layer

The communication layer abstracts information from the network stack layers and presents them to the decision-making layer. It consists of interfaces and channels that exchange control signals to collect application requirements and parameters that represent radio environment conditions. It uses three interfaces to collect data from the TCP/IP stack: Universal Link Layer Application Interface (ULLA) \[143\] or any equivalent C++ based spectrum sensing data structure, Generic Network Interface (GENI), and Common Application Requirement Interface (CAPRI). In addition, Common Control Channel (CCC) \[144\] is used to exchange control messages among nodes for transmission parameters negotiation. ULLA is a generic interface to support access to MAC and physical layers parameters such as SNR, Received Signal Strength (RSS) and frequency. In addition, it sends back selected transmission parameters to the corresponding stack layers such as the operating frequency range, adapt modulation scheme and adjust transmission power. All physical and MAC layers data extracted by ULLA is stored temporarily in the ULLA core before being fetched
by CogWnet modules. The purpose of GENI interface is to support communication between CogWnet and transport and network layers. GENI accesses TCP/UDP data flows and determines loss ratio according to the number of received acknowledgments at the transport layer and routing information at the network layer. GENI and ULLA use the same data structure, which is built using UQL query language. CAPRI is a generic interface for applications to impose their QoS requirements and preferences to CogWnet decision-making layer. Utility functions are used to express the QoS requirements and CogWnet aims to maximize this utility using optimization algorithms. The following equation illustrates how application requirements can be expressed using utility functions:

\[ U = 10 \times \log(t + 1000) + 3 \times d \text{ where } (d \leq 0.050) \]  

where \( U \) is the utility, \( t \) is throughput and \( d \) is the delay. This equation implies that the application cannot tolerate more than 50 ms delay and throughput has higher weight than delay with a weighting factor equal to 10. The utility function is registered in the core of CAPRI for later evaluation by CogWnet decision-making layer. The CCC is dedicated to exchange the selected transmission parameters between nodes. It can be utilized to negotiate the decisions made, improve spectrum usage, reduce network computation, discover neighbors and facilitate the detection of primary users in case of DSA.

**Decision-Making Layer**

The decision-making layer is the core of CogWnet as it accounts for receiving the sensory input from the communication layer and applying channel allocation to assign the suitable channel for communication and optimization algorithms to select the transmission parameters that improve the cognitive network performance. It consists
of two components: repository and parameter mapper.

1. Repository: The repository is an entity that stores the sensory data received from the generic interfaces for further processing. It also schedules various transmission parameter configuration decisions and sends it through the stack layers via the corresponding interfaces. The repository is composed of four containers and two modules as shown in Figure 3.2. Each container corresponds to different data collected from a distinct interface. The set of containers include:

   - **Utility Manager**: It stores and evaluates the utility function calculated by the CAPRI interface to extract the QoS requirements.
   - **Link/Flow Manager**: It stores the data collected from the network adapter about the traffic flows and the available links. This data is sent via GENI and ULLA interfaces.
   - **Physical Container**: It saves the data fetched from the spectrum sniffer through the C++ spectrum sensing data structure.
• **Negotiation Container**: It stores queries and commands to be sent to remote nodes to share the transmission parameters selections based on the demand of the parameter mapper.

The following repository’s modules reside in the user space to process the sensory data using cross-layer design, and schedule the decisions made.

• **Sensory Processing Module**: It is a filter-based on cross-layer design that combines all the data stored in the repository containers in the kernel space and sends them to the parameter mapper for decision-making. Moreover, it includes a database where the information about the existing scenario is saved for the purpose of learning.

• **Action Module**: It is a filter used to receive the selected transmission parameters from the parameter mapper, distribute them to the appropriate containers in the kernel level.

2. Parameter Mapper:

The parameter mapper realizes the decision-making process in CogWnet. It consists of two modules and one container as depicted in Figure 3.3. The container is basically a temporary storage for the data flow between the parameter mapper modules. The modules that reside in the parameter mapper structure are as follows.

• **Action Broker Module**: It acts as a coordinator for the tasks executed by the decision-making layer. It receives the sensory input from the repository and sends it to the decision module. In addition, the action broker sends policy queries to the policy layer to check if the transmission parameters selected are compatible with the policy regulations. After receiving the policy-layer response, the action broker conveys the decision made to the
action module in the repository.

- **Decision Module**: It receives requests from the action broker and take the following actions. First, check if there is any priority in case more than one channel requests are pending. Second, the channel allocation process starts to find the set of free channels according to the spectrum sensing information provided by spectrum sniffer. Finally, the decision module runs the optimization algorithm and outputs the optimal transmission parameters back to the action broker for policy check.

**Policy Layer**

Stakeholders, users, and operators might be interested in applying certain constraints on the spectrum operation. For example, the operator can fix the range of usable frequency bands, so the decision made for radio parameters should not exceed that range. Therefore, policy layer is required to enforce those constraints whether they are static or dynamic based on the geographical location. The policy rules are specified
using declarative languages such as CoRai [145]. The policy layer in Figure 3.4 consists of the following three components.

**Figure 3.4: The Policy Layer**

1. **Policy Server:** It stores all the spectrum policies and it is shared by all the nodes in the network to receive the set of rules that suit their queries. The server consists of a server manager and a database. The server manager is the gateway of the server. It registers the nodes in the same domain and accesses the database to fetch policy rules and sends them to remote policy engines. The database stores the registration of all policy engines with their corresponding policy regulations.

2. **Policy Engine:** It is the main component in the policy layer as it is responsible for checking whether the selected configuration parameters match the policy regulations. The policy engine is composed of a manager, a reasoner, and a database. The manager registers the policy engine in the server, receives all the policy regulations, and processes the policy queries. The reasoner checks the
policy query and replies whether the selected configuration complies with the policy rules. It replies with a boolean value of 1 (‘Yes’) if the configuration is compatible with policy regulations, or 0 (‘No’) otherwise. The reasoner suggests a reason and a solution in case of a ‘No’. All the policy rules that correspond to a certain spectrum class are stored in the database.

3. Trigger Manager: It is a ring that connects the policy layer with the decision-making layer. Trigger manager has two functions. First, it coordinates between the policy engine and the parameter mapper to check whether the selected configuration fits with the corresponding spectrum policy. This function occurs in the policy module of the trigger manager. Second, it tracks the wireless channels status and informs the decision-making layer in case the channel is occupied. This function is executed in the detection module.

3.3 CogWnet Resources Management Mechanism

This section describes the procedure followed by CogWnet to allocate suitable communication channels and configure transmission parameters to improve the quality of the wireless links. CogWnet constantly monitors the wireless environment and checks for channels availability using USRP-N210 sniffers. Channels are considered free if the degree of interference experienced by the spectrum owners at the communication layer is lower than a specified threshold. If a channel is free, the communication layer abstracts the channel information and sends it to the decision-making layer. Each interface in the communication layer accesses the corresponding layers in the stack to collect environmental parameters. At the same time, the sniffer keeps monitoring the interference level. It prompts the decision-making layer to switch to a different channel if the primary user accessed the channel again. The repository and the parameter mapper work closely to make environmental information available for the decision module. Figure 3.5 shows the state diagram for all functions executed by
CogWnet for RRM and spectrum allocation. The diagram consists of two terminals

Figure 3.5: Radio Resource Management: CogWnet modules execute cognitive functions. Actions are scheduled back through the link configuration state (transmitter and receiver) each equipped with USRP-N210 as a spectrum sniffer and GUN radio as a radio platform able to extract the physical parameters collected by the sniffer. In addition, CogWnet components are installed on these terminals for the decision-making process, reconfiguration and action negotiation. Most of the states occur within the repository as a main component for processing sensory information, and application requirements abstraction. The parameter mapper’s function is to allocate the suitable channel and run the decision-tree optimization algorithm to output the optimized transmission parameters. The trigger manager handles all policy issues. The cognition process functions are illustrated further in Table 3.1.

The decision module allocates the most suitable channel and configures the radio transmission parameters as in the following sections.
<table>
<thead>
<tr>
<th>Function</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application requirements</td>
<td>To collect the application requirements using CAPRI Interface</td>
</tr>
<tr>
<td>collection</td>
<td></td>
</tr>
<tr>
<td>Historical data collection</td>
<td>To consult the database at the repository for similar encountered scenarios</td>
</tr>
<tr>
<td>Spectrum data collection</td>
<td>To collect spectrum data using USRP-N210 sniffer</td>
</tr>
<tr>
<td>Device discovery</td>
<td>To discover the target device of the network flow and collect device capabilities</td>
</tr>
<tr>
<td>Optimal configuration</td>
<td>To determine the optimal network configuration based on the requirements and environmental information</td>
</tr>
<tr>
<td>determination</td>
<td></td>
</tr>
<tr>
<td>Possible link configurations</td>
<td>To compute multiple possible link configurations on a device</td>
</tr>
<tr>
<td>Find a new radio or channel</td>
<td>To search if there is another radio channel configuration that could meet application requirements</td>
</tr>
<tr>
<td>Monitor network behavior</td>
<td>To monitor the network behavior of different applications and try to identify possible bottlenecks during network transmission. In this case, the &quot;determine optimal configuration&quot; function is triggered</td>
</tr>
<tr>
<td>Configure new radio or channel</td>
<td>This is included in the link configuration function and is used switch into different radio access technology if it is necessary</td>
</tr>
<tr>
<td>Configure new link parameters</td>
<td>To change the current link parameters to those determined by CogWnet to accommodate new application requirements or network conditions</td>
</tr>
<tr>
<td>(in case of change to new channel)</td>
<td></td>
</tr>
<tr>
<td>Link Configuration</td>
<td>To push the parameters negotiated by a device to the underlying network card to create a link</td>
</tr>
<tr>
<td>Negotiation</td>
<td>To negotiate common channel and its parameters with other nodes</td>
</tr>
</tbody>
</table>

Table 3.1: Cognitive functions for radio resources management

3.3.1 Channel Allocation Mechanism

Graph coloring based algorithm is exploited as channel allocation technique in CogWnet. Graph coloring techniques have a long history in interference aware channel allocation, especially in cellular networks [146,147,148]. In graph coloring technique, each vertex corresponds to a transmitter. Edges connect two nearby interfering transmitters. The channel allocation can be modeled as Vertex-coloring in a way that adjacent vertices are not assigned the same color (channel) to avoid interference. Given a collection of base stations, an interference graph $G = (V,E)$ can be constructed where $V$ is the set of vertices and $E$ are the edges that correspond to the link between each two vertices. A set of colors $C$ also defined as a collection of channels available to certain base station. The channel allocation problem is simply the pro-
cess of finding a color for certain vertex that is not used by any adjacent (interfering) vertex. The size of the coloring set depends on the communication technology used. For more extensive background in graph theory and graph coloring, including proofs of the above statements, please see [149] [150]. The more base stations the network has, the more the coloring graph algorithm become inappropriate. However, efficient heuristics have been developed to build a number of different coloring algorithms that can optimally color an interference graph under different constraints.

A heuristic based algorithm suitable for interference graph coloring has been developed. This algorithm is inspired by ”degree of saturation” or DSATUR algorithm proposed by Brelaz [151]. DSATUR is a deterministic greedy algorithm that uses a heuristic method called ”degree of saturation” according to which the vertices of a given graph are colored. The algorithm aims to search and find the set of vertices that has the highest degree of saturation, which means the vertices with the largest number of differently colored neighbors. The coloring procedure is achieved through the following stages:

1. The set of vertices is sorted according to the vertex degree from the highest to the lowest.

2. Now, the vertex of the highest degree of saturation is selected for coloring with color X. If all the vertices have the same degree of saturation, the vertices degree is used as a tie breaking rule to start coloring. The last breaking rule will be to choose the vertex randomly if all the vertices have the same degree.

3. If one of the selected vertex’s neighbors is colored by the color X, another color is chosen for coloring Y. Otherwise, it is colored by X.

4. The same process repeated again until all the vertices are colored such that the neighbors have different colors.
There are two problems that encountered when DSATUR is applied for channel allocation. The first problem is when the base station must change its current channel to avoid connection loss. This is a typical scenario occurs when a new base station is added to the network. Therefore, it is the best to perform coloring again as a new base station is added to the interfering graph. Online coloring is exploited to tackle this problem, which colors a given graph $G$ in such that one vertex at a time is revealed. Once a vertex is added to the graph, the edges with the previous arriving vertices have to be established and then the color is determined. Detailed analysis of the on-line coloring in the channel assignment context, can be found in [152]. The second problem occurs in the dense graphs where the set of non-overlapping channels is not large enough to accommodate all the connections. For this case, T coloring is adopted, which imposes a distance-type of condition for the channels of adjacent vertices. It models the interference by non-negative set called $T$. The elements of $T$ determine the smallest allowed difference of the central frequencies used by interfering base stations. Separation of channels would mean operating in non-overlapping spectrum. T-coloring has been extensively studied in the literature (see for e.g., [153]).

This coloring algorithm has many advantages. First, channel allocation to be implemented in a distributed fashion without mandatory use of central coordinator to run the coloring algorithm and this is suitable for CR environment. In addition, it has low complexity as its running time is $O(m \log n)$ where $m$ and $n$ are the size and the order of the graph respectively.

### 3.3.2 Configuration Mechanism

The configuration process configures the transmission parameters in order to match the environmental conditions. DTs is the machine learning tool used in CogWnet for configuration and policy compatibility check. A decision tree is a directed graph consisting of a hierarchical set of nodes connected via arcs, where each node repre-
sents a choice or decision, and the arcs leading from that node to the next decision node represent the set of possible solutions for a given node. A decision tree can be visualized as a sequence of choices in which the path taken through the choices from the starting point to the endpoint is governed by the selection made at the starting node and each successive node. The root node provides the starting point and the leaf nodes contain the decisions or actions to be taken.

CogWnet DTs based configuration is performed through a set of predetermined conditions, and the system is trained to select the right configuration. DTs learn to associate a particular set of input stimuli with a learned response or output. The problem of finding the most proper transmission parameters can be modeled as a solution space search problem. That is, the system generates a set of possible alternative actions (possible configurations for certain parameters) for a given network scenario. Each of the generated alternatives is then explored and alternative actions are generated for each of these scenarios, and so on. As each alternative is generated, its weight is calculated, subsequent alternatives are generated, and a total weight or probability is calculated. Thus, for any given node in a decision tree, the total weight for the path leading to that point is maintained. The calculation of the weights can be dynamically adjusted based on an observed state compared to an anticipated state. Thus, the system can learn to respond differently to different scenarios. The learning aspect can be extended further by enabling the system to generate new actions for a given node. This enables the radio system to not only learn new behavioral responses based on performance observations, but also to extend the range of choices.

The output of the above optimization process comprises the set of parameters to be configured for all the stack layers. These parameters include flow rate for transport layer, transmission range, and number for hops for network layer, packet size and contention window size for MAC layer, modulation scheme, coding rate and transmission power for physical layer. The selected parameters are then checked for
policy compatibility using the policy layer. Subsequently, the negotiation process starts with CogWnets running on all the nodes of the network to avoid any conflict between terminals. Finally, the transmission parameters configuration is applied to the corresponding layers through the communication layer.

In addition, CogWnet architecture accounts for improving system capacity and signal quality variation. Adaptive modulation is used to match the modulation scheme with channel conditions. For channel allocation, the decision-making layer adjusts the modulation as a function of the distance of the spectrum owners in order to match the quality of the signal received. The distance between the cognitive user and the primary user is used to control the transmission power to avoid interference.

3.4 CogWnet Testbed Implementation

3.4.1 Motivation

Home network with IEEE802.11a/g has heterogeneous nature in terms of type of devices and connections characteristics. Access points are the only units responsible for resources management in the home networks. They use static channel allocation schemes and very often the access points are set up to a default channel by the manufacturers themselves. Consequently, this creates serious interference problems and service degradation if the number of users operating in the same spectrum increases. These practical problems with Wi-Fi deployments made the RRM problem an interesting research with direct technology transfer possibilities. In addition, ISM band where Wi-Fi networks operate, provide a useful model to quantitatively test CogWnet functionality.
3.4.2 Testbed Topology

The network topology used in this evaluation consists of two pairs of terminals with CogWnet installed on each. Between each pair, a different connection with different types of traffic is running. The wireless card of each terminal used in this evaluation is virtually divided into two cards, one for control exchange and negotiation and the other for data transfer. USRP-N210 hardware is exploited for spectrum sensing and the policy server is running on a separate computer. The evaluation is running on Linux/windows based machines. We employed two kinds of applications in this scenario: UDP based video streaming as a delay-sensitive application and normal TCP applications. The UDP video streaming application is running on the first pair of nodes. The second pair has more than one application running. With legacy IEEE 802.11, all the transmitted data has to go through the access point which uses one common channel. All the nodes have to compete for this channel even if it is congested or interfering with other sources, resulting in throughput degradation and network overloading. In addition, IEEE802.11 is not capable of supporting different applications if the bandwidth demand exceeds channel capacity. With CogWnet, however, the decision-making layer checks whether there is the possibility to establish a direct connection on a different channel between any of these pairs, based on the sensory information received from the communication layer. In this way, the throughput of the connection increases and does not compete with other connections that go through the same access point. CogWnet adjusts the channel width, center frequency, modulation order, symbol rate, transmission range for each hop and flow rate at transport layer to ensure that the capacity of the selected communication channel is sufficient to accommodate different QoS applications and switch to another channel if necessary. In this scenario, the sensory input from the communication layer includes ratio of successfully received packets, link load, RSS, BER, spectrum occupancy power, and application utility.
3.4.3 Evaluation Results and Discussion

Figures 3.6 (a) and 3.6 (b) show the goodput and the utility achieved by the home network with CogWnet enhancements. In the first period of 20 s, both pair of nodes run the same UDP video streaming with a data rate of 3 Mbps. Therefore, CogWnet decides to co-locate both connections on channel 6 with a bandwidth of 6 MHz which is sufficient to accommodate both streams. After 20 s, the second pair decide to send two more TCP flows, one for high definition video and another for file download. Consequently, the communication layer detects high link load in the driver and notifies the repository. The decision module is informed by the repository and runs the utility-based optimization and channel allocation algorithms that move the second pair of nodes to channel 11 and increases channel width to 20 MHz to satisfy the application requirements. Figure 3.6 (b) shows that CogWnet manages to keep the application utility high regardless of the presence of different traffic flows from different nodes. Another test was carried out to evaluate the advantage of setting up a direct connection between nodes versus communication through an access point. Figure 3.7 shows the throughput achieved by the direct connection established by

![Figure 3.6: Goodput and utility achieved in the home scenario](image)

CogWnet compared with IEEE802.11 connection through access point. It is clear
that the direct connection achieves better throughput as it is not affected by the distance from the access point or by contention with other nodes which send or receive their traffic through the access point.

![Figure 3.7: Performance of the home network with and without CogWnet](image)

Interference has been used as a performance metric to demonstrate CogWnet functionality. There are two possibilities for interference in the home network scenario: interference from another 802.11-Carrier Sense Multiple Access (CSMA) based node or interference from external non-802.11 based sources such as any kind of packet injector. The ULLA interface in the communication layer monitors the RSS and link error rate. If interference affects the quality of the current transmissions, a signal is sent through the repository to the parameter mapper to reconfigure the radio to mitigate the detected interference detected. At the parameter mapper, the type of interference is determined according to the current channel characteristics.

Figure 3.8 presents the performance of CogWnet in mitigating all kind of interference mentioned in the established link in terms of throughput. The red line shows the case where there is a packet injector using channel 6. CogWnet observes this event and decides to switch the traffic to channel 11 to avoid interference and exploit full channel capacity. However, in the case where both type of interference exist (external source on channel 6 and 802.11 CSMA based source on channel 11, black
line), the decision module co-locates the traffic on channel 11 with a bandwidth of 20 MHz as no other non-overlapping channel is available. IEEE802.11 device experiences throughput degradation because it cannot switch to a non-interfering channel automatically.

![Figure 3.8: Performance of CogWnet in Mitigating Interference](image)

Figures 3.9 (a) and 3.9 (b) present a comparison of throughput between 802.11 and CogWnet in the case of an external source interference and the case of both external and 802.11 based interference respectively. Initially, both frameworks obtain the same throughput as they use channel 6. Then, CogWnet starts the optimization process and switches all traffic to channel 11 as in Figure 3.9 (a). In Figure 3.9 (b), the decision-making layer colocates the traffic with the 802.11 based interference source on channel 11 and extends channel width to allow both connections to run with minimum interference. The 802.11 device stays in channel 6 with interference since it has no automated mechanism to change the channel. Finally, CogWnet is evaluated to quantify the reliability of transmission. RSS, propagation distance, and the number of successfully received packets are collected from the communication layer interfaces. Channel width, transmission range, and number of hops are tuned using the decision-tree optimization algorithm to reduce the packet loss ratio.
Figure 3.9: Throughput comparison between CogWnet and 802.11 with interference from external packet injector and from 802.11 based source and external source respectively.

Figure 3.10: Throughput comparison between 802.11 and CogWnet based on the distance between sender and receiver.

Figure 3.10 compares the throughput obtained by 802.11 and CogWnet for various cases: sender and receiver are in the same place, in adjacent offices, in isolated offices or far away. CogWnet achieves better throughput and utilization as it adapts transmission parameters to minimize packet losses. CogWnet also outperforms 802.11 since it is difficult for 802.11 devices to predict channel and environmental changes.
3.5 CogWnet Decision-Making Evaluation

In this section, the performance of the decision-making engine of CogWnet is evaluated to demonstrate the configuration capability of the system. The evaluation process incorporates two scenarios. The first scenario involves RRM scheme for WRANs IEEE 802.22 networks. The second scenario tackles the RRM in OFDMA primary/secondary users network. The valuation emphasizes the impact of learning on the network efficiency and spectrum utilization.

3.5.1 Resource Management for IEEE 802.22 WRANs Scenario

IEEE 802.22 WRANs is the first wireless standard to improve spectrum utilization by exploiting spectrum holes in the TV broadcast services [154]. Sharing spectrum resources is constrained by the condition that interference to the primary users which are TV receivers must be minimal. Secondary users must be aware of the primary users activities and adapt to them as well as other channels impairments such as channel fading. Spectrum sensing is utilized to provide this awareness. Radio adaptation is necessary to satisfy applications QoS requirements. All these processes are executed in the cognitive RRM engine and they occur with accordance to the IEEE 802.22 standard.

In this scenario, we compare the performance of CogWnet decision-making mechanism with other CEIs that were designed to tackle the spectrum utilization problem in WRANs. The comparison involves two search based engines which are GA [155], and Hill Climbing Search (HCS). In addition, we compare CogWnet to the CBR engine proposed in [156]. This engine is based on the emulation of human learning, understanding, and problem solving process using experience (cases) [157]. An optimizer is also designed to work in conjunction with the reasoning engine to check the solutions validity and improve them if necessary. This optimizer is based on HCS search. Another engine called Case Knowledge Learning (CKL) is used in this
comparison [158], which involves CBR and Knowledge Based Reasoning (KBR). The engine switches the learning mode from case based to knowledge based according to environment conditions. Two evaluation metrics were exploited in this comparison, which are application utility and average adaptation time. The utility function is a mathematical equation calculated to specify QoS metrics satisfaction. In this evaluation, we have three QoS metrics: BER, transmission power, and data rate. The generic utility function is calculated according to 3.2

\[ f(x, x') = \frac{1}{2} \tanh[\log(x/x') - \eta]\sigma + 1 \]  

where \( x \) and \( x' \) are the measured and target metrics respectively. \( \eta \) and \( \sigma \) are the threshold and spread factors that are used to maintain the utility value range between 0 and 1. The global utility function is defined in 3.3 as:

\[ U_{global} = \prod_k (u_k)^{w_k} \]  

where \( w_k \) is the weight assigned for certain metric utility \( u_k \).

The simulation environment is a scenario driven environment depicted in a series of Extensible Markup Languages (XML) files for different network conditions. The configuration parameters (Table 3.2) considered in the simulation are frequency (TV channel), transmission power, modulation scheme, coding rate and the number of the OFDMA sub-carriers allocated for uplink and downlink. The radio links between the terminals and the base station are randomly generated. The distance between the base station and the terminals is set to be 33 km as a maximum. Data rate and BER requirements were chosen randomly for each connection. The detailed simulation parameters are listed in Table 3.3.

The application utility is depicted in Figure 3.11 for all the tested cognitive engines. The average adaptation time for all engines is compared in Figure 3.12. We
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value or Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency channel</td>
<td>VHF / UHF (54 - 862 MHz)</td>
</tr>
<tr>
<td>Terminals and base station transmission power</td>
<td>Up to 4 Watts EIRP, subject to EIRP profile defined by 802.22 Standard</td>
</tr>
<tr>
<td>Modulation scheme</td>
<td>QPSK, 16QAM, 64QAM</td>
</tr>
<tr>
<td>Channel coding</td>
<td>1/2, 2/3, 3/4</td>
</tr>
<tr>
<td>Number of uplink and downlink subcarriers</td>
<td>Variable 4 - 256</td>
</tr>
</tbody>
</table>

Table 3.2: Adjustable parameters in the WRANs

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value or Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of base stations</td>
<td>1</td>
</tr>
<tr>
<td>Cell Radius</td>
<td>33 km</td>
</tr>
<tr>
<td>Number of Nodes</td>
<td>1-30</td>
</tr>
<tr>
<td>Distribution of Nodes</td>
<td>Uniformly distributed</td>
</tr>
</tbody>
</table>
| Type of service, Data Rate, BER          | Voice: 10kbps, 10\(^{-2}\)
                                             | Video: 100kbps, 10\(^{-3}\)
                                             | Low speed data: 250kbps, 10\(^{-6}\)
                                             | High speed data: 750kbps, 10\(^{-6}\) |
| Multiplexing/Duplex                      | OFDMA/TDD            |
| Number of subcarriers                    | 2048                 |

Table 3.3: Simulation settings

Figure 3.11: Utility results for different cognitive engines

can observe in Figure 3.11 and Figure 3.12 that all the algorithms manage to achieve good performance for 15 or less terminals connected to the base station. However, for larger number of terminals, we can observe that GA has the worst performance
as it is a stochastic search algorithm. The probability that the crossover or mutation affect certain radio parameters is low, especially for complex scenarios. Therefore, a significant portion of the search time is wasted on regions of the search space that do not affect the utility. Thus, the adaptation time of GA is the highest. HCS engine achieved considerable utility since it predictably adjusts each parameter and it tries to find the local maximum of the solution space. The results above suggest that the solution space for this particular simulation either has few if any local maximum, or has equivalent local maximums, thus allowing the HCS to reliably find a good solution. However, its adaptation time is not optimal. The CKL and CBR engines achieve good utility within short time when the number of nodes is low. However, they may simply not be able to find a viable solution under some extreme radio scenarios (e.g., when the required resources approach the capacity limit of WRANs systems). Therefore, they are only useful for closely matched situations with the similar utility functions. In addition, they require a dedicated memory to store the cases and complicated control to handle both KBR and CBR. The case library and knowledge base may also need to be updated accordingly when the utility function
of cognitive engine changes. On the other hand, CogWnet decision-making engine managed to achieve the highest utility within the shortest time as it employs decision tree based learning mechanism and cross-layer optimization fashion. The weights for the utility function is updated dynamically which makes the engine more suitable for dynamic environments, which is mostly the case in most of the wireless networks.

### 3.5.2 Resource Management for an Efficient Spectrum Utilization

The target of this scenario is to improve spectrum utilization and maximize the system throughput in an OFDMA radio environment. Here, we demonstrate the strength of Cogwnet capability for configuration of radio parameters by comparing it with two other cognitive engines. The first one as in [159], adopts cross-layer design approach to jointly consider the spectrum sensing, access decision, physical-layer modulation and coding scheme, and data-link layer frame size in CR networks to maximize the TCP throughput. The second engine [160] exploits Q-learning to learn the channel state information. A Q function is employed to organize and construct the search for optimal parameter configuration. Both engines uses Markov decision process to abstract the channel model.

The simulation was carried out using NS-3 software. The simulation scenario represents a typical spectrum sharing environment which incorporates two base stations: one primary with 5 terminals and one secondary with 5 associated terminals. The terminals are randomly distributed on 1000 X 1000 m field. It is assumed that the terminals always have TCP data to send. The rest of the simulation parameters are stated in Table 3.4

The simulation results in terms of the achieved system throughput as a function of SNR are depicted in Figure 3.13. We can observe that CogWnet learning capability achieved the highest throughput as it was learning faster than the Q-learning engine since the latter needs large computation time and complex processing. Moreover,
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value or Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP packet size</td>
<td>1500 bytes</td>
</tr>
<tr>
<td>Maximum number of frame re-transmissions</td>
<td>10</td>
</tr>
<tr>
<td>Initial timeout</td>
<td>2 sec</td>
</tr>
<tr>
<td>Frame header length</td>
<td>20 bits</td>
</tr>
<tr>
<td>Channel bandwidth</td>
<td>1 MHz</td>
</tr>
<tr>
<td>Modulation and coding schemes</td>
<td>3/4 BPSK, 3/4 QPSK, 3/4 8PSK and 3/4 16QAM</td>
</tr>
<tr>
<td>Data link protocol</td>
<td>ARQ based</td>
</tr>
<tr>
<td>number of simulation runs</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 3.4: Simulation Parameters

![SNR vs system throughput for the three engines](image)

Figure 3.13: SNR vs system throughput for the three engines

CogWnet is not limited to physical and MAC layer parameters in the radio reconfiguration process like the Q-learning and cross-layer engines which positively impacts the throughput. The cross-layer engine proposed in [159] has the lowest throughput as it does not have memory function and almost no learning ability.

In this chapter, we introduced an implementation of CogWNet. It added an awareness factor to the cognitive engine and provides an efficient radio resources for user applications. Numerical results proved that CogWNet can efficiently allocate the available spectrum and tune transmission parameters according to the user application. In the next chapter, we consider the cognitive engine in more details. We
show that using AI techniques for the decision making function can further enhance the accuracy of the decisions, better allocate the spectrum parameters, and address the tradeoff of different AI techniques in the design of CE.
Chapter 4

Artificial Intelligence Approaches for Cognitive Engine Design and Radio Parameters Configuration

Network efficiency and proper utilization of its resources are essential requirements to operate wireless networks in an optimal fashion. CR aims to fulfill these requirements by exploiting AI techniques to create the CE. CE exploits awareness about the surrounding radio environment to optimize the use of radio resources and adapt relevant transmission parameters. In this chapter, we explore different artificial approaches to realize the decision-making function of the CE, which aims at configuring the radio transmission parameters according to the environment conditions. This requires definition of set of parameters that represent the channel state information and other related environment conditions. The approaches considered in this chapter include GA, CBR, DTs, ANNs, and online RL. The design of the CE follows three different methodologies: Single heuristic algorithm. Hybrid CE that consists of two different decision-making algorithms, and supervised cognitive system, which incorporates multiple decision-making algorithms and one supervision engine to select the most appropriate CE for parameters configuration based on the encountered scenario. All the proposed CEs are implemented in IEEE 802.11 office environment using SDRs testbed.
4.1 Adaptive Multi-Objective Optimization Based Cognitive Engine Using Enhanced Genetic Algorithm

In this section, we propose an Adaptive AMO CE\[18\] that is capable to handle multiple competing objectives for radio parameters configuration with consideration of policy limitations. It exploits cross-layer optimization for radio environment awareness and GA as an optimization tool for selection of transmission parameters. glsga is well suited for multi-objectives optimization problems in CR networks as it can search for multiple sets of solutions over a large search space, and can enforce constraints \[161\]. Ease of implementation as well as optimization of discrete and continuous radio parameters render GA as an excellent optimization tool for making resource management decisions in CR networks \[162\]. AMO CE aims to meet QoS requirements, which include throughput, delay, and BER, and improves spectrum efficiency with minimum interference and power consumption. It is composed of six fitness functions that support minimization of power consumption, reduction of BER, minimization of delay and interference, and maximization of throughput and spectral efficiency. To ensure that GA runs effectively and exploit all solutions possibilities, AMO CE adjusts GA parameters (crossover and mutation) dynamically to fit with the available search space, which also helps GA not be stuck in local optima. In addition, the following salient features distinguished AMO CE from other cognitive resource management engines.

1. All TCP/IP stack layers in the radio environment are considered for environmental parameters input. This improves accuracy and adaptivity, and enables efficient transmission.

2. The objective functions used in AMO CE are comprehensive by incorporating throughput, delay, BER, spectral efficiency, power, and interference.

3. AMO CE is not limited to point-to-point link optimization and accounts for the
environment, where cognitive links are non-cooperative.

4. The weights and priorities of objective functions are adapted automatically according to the environment conditions and system capability.

4.1.1 Constrained Multi-Objective Optimization in Cognitive Radio

Assuming a multi-carrier dynamic wireless environment, the basic functions of CR are to sense the environmental parameters, and run optimization to adjust the value of transmission parameters to achieve the predefined QoS. In this section, we define the cognitive optimization problem and the associated parameters in the optimization process. In addition, our multi-objective optimization model is described with all related fitness functions.

Multi-Objective Optimization Problem

The multi-objective optimization problem is to determine the optimal value of a set of solutions \(x\) while optimizing a set of \(k\) conflicting objectives simultaneously. Mathematically, the multi-objective optimization problem is defined as

\[
\min/\max(Y) = f(x) = [f_1(x), f_2(x), \ldots, f_n(x)]
\]

subject to \(x = (x_1, x_2, \ldots, x_m) \in X\) and \(y = (y_1, y_2, \ldots, y_n) \in Y\) where there are \(n\) dimensional spaces, and \(f(x)\) is the fitness function for a certain dimension. \(X\) is the set of transmission parameters and \(Y\) the set of dimensions. The fitness function is a mathematical equation that is evaluated to reach certain optima according to the radio environmental parameters and QoS requirements. They are normalized to have score between 0 and 1. The CR parameters adaptation problem can be modeled as a multi-objective optimization problem with the goal to find \(x\) that minimize or maximize the fitness function. The objectives under consideration might conflict with each
other. For example, minimizing power and minimizing BER simultaneously creates a conflict due to the single parameter, transmit power, affecting each objective in a different way. The optimal set for multiple objective functions lie on what is known as the Pareto optimal front. This front represents the set of solutions that cannot be improved upon in any dimension. The solutions on the Pareto front are optimal and co-exist due to the trade-offs between the multiple objectives. These trade-offs represent the core of the multiple objective optimization problem. Therefore, CE must perform actions based on a single set of parameters, which should be selected from the Pareto front.

**Radio Parameters**

Radio parameters of CR are categorized into environmental parameters and transmission parameters. The former gives knowledge of environmental characteristics of the wireless environment and used as inputs to the CE. Environmental parameters include path loss ($PL$), noise power ($N$), signal-to-noise ratio ($SNR$), Acks of successfully received packets, spectrum occupancy information, fading statistics, and Frame Error Rate (FER). Transmission parameters are the output parameters of the CE. The engine adjusts the transmission knobs to corresponding values from the optimal parameter set. The transmission parameters are listed as: transmission power ($P$), type of modulation scheme used for the communication, modulation index ($m$), bandwidth ($B$), channel coding rate $R_c$, Time Division Duplex ($TDD$) in percentage, and symbol rate ($R_s$) for the physical layer, frame size ($L$), contention window size ($CW$), and source coding for the MAC layer, transmission range ($d$), and number of hops for the network layer, and congestion control mechanism for the transport layer. Modulation Index is defined as the total number of symbols in a modulation scheme and TDD represent the percentage of transmit time.
Objective Fitness Functions

In AMO CE, we adopt the weighted sum approach to find the unified optimization function because it captures the tradeoffs between multiple objectives. The weighted sum approach method suits the CR environment well since it provides a convenient process for applying weights to the objectives. Changing the objective direction of the fitness function requires only a simple change of the weighing factor. AMO CE adjusts the weighing factors automatically according to the application and environment awareness information. This information determines which of the objective functions is contributing more to meet the application QoS requirements. For example, a radio in default mode may be operating so to ensure the best throughput possible while not caring much about minimizing power. However, assuming this is a battery powered radio, the system may sense low battery power and modify the objective weights to emphasize minimizing power. In addition, the derived individual objective fitness functions are bounded by threshold values based on the applications QoS requirements. This narrows the search space and helps to enforce policy regulation by specifying certain range for each transmission parameter.

AMO CE optimization consists of the following fitness functions: power, BER, interference, and latency minimization, and throughput and spectral efficiency maximization. The total optimization function based on a weighted sum approach is calculated as:

\[
f_{TO} = w_1 f_P + w_2 f_T + w_3 f_{BER} + w_4 f_D + w_5 f_{SE} + w_6 f_I \tag{4.1}
\]

where \( w_1 \) to \( w_6 \) are the adaptive weights of the corresponding fitness functions. The fitness functions in (4.1) are: \( f_P \) is the power minimization function, \( f_T \) is the throughput maximization function, \( f_{BER} \) is the BER minimization function, \( f_D \) is the delay minimization function, \( f_{SE} \) is the spectral efficiency maximization function, and \( f_I \)
is the interference minimization function. Each of these fitness functions is detailed next.

1. Minimization of Power Consumption:

Battery life and power consumed are essential factors for reducing power consumption. The parameters contribute to the fitness for power minimization are bandwidth, modulation index, coding rate, time division duplexing symbol rate, contention window size, which impacts the number of re-transmissions, and transmission power.

The fitness function to minimize power consumption for $N_c$ subcarriers is calculated as:

\[
 f_P = \begin{cases} 
 1 - \left[ \sum_{i=1}^{N_c} (P_{\text{max}} - P_i) + (B_{\text{max}} - B_i) + 
  \log_2 m_{\text{max}} - \log_2 m_i + (R_{s\text{max}} - R_{si}) \right] \times 
 \frac{1}{N_c * (P_{\text{max}} + B_{\text{max}} + \log_2 m_{\text{max}} + R_{s\text{max}})} , & \text{if } P_i \leq P_{\text{max}} \\
 0, & \text{if } P_i > P_{\text{max}} , 
\end{cases}
\]  

(4.2)

The subscript $i$ refers to $i^{th}$ subcarrier of the multicarrier system, and $P_{\text{max}}$ is the threshold for maximum transmission power.

2. Maximization of Throughput:

Throughput is another important performance metric and considered as the main optimization objective in certain applications like multimedia. BER causes degradation of system throughput and should thus be maintained at an acceptable level. Throughput depends on the following parameters: bandwidth, coding rate, modulation index, frame size, probability of BER ($P_{be}$), contention
window size as it affects the frames transmission rate, congestion control mechanism, number of hops, which is inversely proportional to throughput, transmission range and percentage of transmit time.

The fitness function for throughput for $N_c$ subcarriers is

$$f_T = \begin{cases} \sum_{i=1}^{N_c} \frac{L_i}{L_i + O + H} \times (1 - P_{be,i})^{L_i + O} \times R_{ci} \times TDD_i \frac{1}{N_c}, & \text{if } T \geq T^* \\ 0, & \text{if } T < T^* \end{cases}$$ (4.3)

Where, $O$ is physical layer overhead, $H$ is MAC and IP layer overhead, and $T^*$ is the threshold for minimum throughput requirements.

3. Minimization of Bit Error Rate:

BER is used to measure the quality of each link in terms of number of errors each bit encountered. Generally BER depends on several parameters like transmit power, modulation type, modulation index, bandwidth, symbol rate, contention window size as it is proportional $P_{be}$, number of hops and noise power. The fitness function for minimizing BER is expressed as:

$$f_{BER} = \begin{cases} 1 - \left[\frac{\log_{10}(0.5)}{\log_{10}(P_{be})}\right], & \text{if } BER_t \leq BER \leq BER^* \\ 1, & \text{if } BER < BER_t \\ 0, & \text{if } BER_s > BER^* \end{cases}$$ (4.4)

where $BER$ is the calculated BER for a certain solution, $BER_t$ the target BER and $BER^*$ the threshold of maximum tolerable BER.

4. Minimization of Delay:

There are numerous applications sensitive to transmission latency, so the latency
becomes an essential metric in any wireless communication. The delay objective is represented by the following parameters: modulation index and frame size. The fitness function to optimize transmission latency is expressed as:

\[ f_D = 1 - L_i \times \left( \frac{\log_2 m_{\text{max}}}{L_i \text{min} \times \sum_{i=1}^{N_c} \log_2 m_i} \right) \]  

where \( L_i \) is frame size for \( i^{th} \) subcarrier.

5. Maximization of Spectral Efficiency:

Spectral efficiency can be defined as the amount of information that can be transmitted over a given bandwidth. The symbol rate and modulation index are used to determine the total amount of information being transmitted. Spectral efficiency is computed as

\[ S_e = \frac{R_s}{B} \] (4.6)

The fitness function to maximize spectral efficiency is expressed as:

\[ f_{SE} = \begin{cases} \sum_{i=1}^{N_c} \frac{m_i \times R_s + R_{\text{min}}}{B_i \times m_{\text{max}} + R_{\text{max}}}, & \text{if } S_e \geq SER^* \\ 0, & \text{if } S_e < SER^* \end{cases} \] (4.7)

where \( SER^* \) the threshold for minimum spectral efficiency imposed by QoS.

6. Minimization of Interference:

Interference is a fundamental problem in shared spectrum environments like cognitive networks. For example, minimizing it is given the highest priority in making spectrum allocation to SUs in a licensed band. minimization of interference is necessary to achieve high throughput with less error rate. Transmission parameters such as transmit power, bandwidth, and time division duplexing are used to determine the approximate spectral interference fitness value. The
following fitness function is used to minimize interference:

\[
 f_I = \begin{cases} 
 1 - \frac{\sum_{i=1}^{N_c} ((P_i \times B_i \times TDD_i) - (P_{min} \times B_{min} \times 1))}{N_c \times (P_{max} \times B_{max} \times 100)}, & \text{if } P_{min} \leq P_i \leq P_{max}, B_{min} \leq B_i \leq B_{max} \\ 
 0, & \text{otherwise} 
\end{cases} 
\]  

(4.8)

where \( P_{min}, P_{max} \) and \( B_{min}, B_{max} \) are the boundaries for the allowed transmission power and bandwidth respectively.

4.1.2 Adaptive Multi-Objective Optimization Cognitive Engine Mechanism

The optimization function specified in (4.1) is not linear since the objective functions cannot be treated independently. Therefore, we must consider trade-offs among the optimization goals; this is what makes AMO CE best fit for this multi-objective optimization problem. AMO CE is an evolutionary optimization engine for making resource management decisions, inspired by GA. It starts with a random population of solutions and evolves to reach an optimal solution. Each solution is represented by a trial, which consists of a string of binary bits that corresponds to transmission parameters. The parameters are encoded into bit strings trials using binary encoding and quantized \textit{a priori}, which has power-of-2 ranges. The trials are evaluated using fitness functions to determine the set of best solutions (trials) that meet the QoS requirements. The higher the obtained score of this function, the better the solution. AMO CE algorithm for transmission parameters adaptation is explained in Algorithm 1. First, the objective functions contribution and search spaces are specified based on the application type and QoS requirements. For example, if the application was multimedia, then the weight of the throughput fitness function will be the maximum. However, all the other five objective functions will participate
Algorithm 1 Transmission Parameters Adaptation

Require: Environment Parameters (SNR, PL, N, FER, BER)
Ensure: Transmission Parameters (P_i, m_i, B_i, R_{ci}, R_{si}, L_i, TDD_i)

BEGIN

Check application type and QoS requirements to initialize AMO CE algorithm
Define the objective functions
Assign the proper weight (w_1, w_2, w_3, w_4, w_5, w_6) for each objective function
\((f_P, f_T, f_{BER}, f_D, f_{SE}, f_I)\)

while \(generation \leq generation_{Max}\) do

Extract (SNR, PL, N, FER, BER)
Update the weight \((w_1, w_2, w_3, w_4, w_5, w_6)\) for each objective function according to the current channel conditions.
\(generation \leftarrow 1\)

while \(population_c \leq Population_{size}\) do

Generate random solution (trial) \((P_i, m_i, B_i, R_{ci}, R_{si}, L_i, TDD_i)\)
Send the frames
Evaluate the trial using the corresponding fitness function
Save the fitness score for this trial
\(population_c \leftarrow population_c + 1\)
end while

Create median population (Selection Process)
Elitism process (select trials with top 10% fitness score to duplicate)
Clear new population \(P'\)
while \(|P'| \leq Population_{size}\) do

Select 2 parents from population
Perform Crossover and Mutation
Update CF and MF
Insert offspring to \(P'\)
end while

Compute \(CF_{avg}\) and \(MF_{avg}\)
Adjust \(RC\) and \(RM\)
\(population \leftarrow P'\)
\(generation \leftarrow generation + 1\)
end while

Output the set of transmission parameters \((P_i, m_i, B_i, R_{ci}, R_{si}, L_i, TDD_i)\)

END
in the optimization process but with less weight. Then, AMO CE extracts radio environment parameters from the cross-layer interfaces of the designed cognitive architecture. The weights of the fitness functions are updated in real-time according to the current channel conditions. For example, if the channel quality degrades, the weight of BER is increased and the weight of throughput is reduced to compensate for the poor channel. The process of weights selection and update in real-time is the responsibility of the reasoning module of the cognitive architecture [16]. Next, an initial random trial \( x = P_1, P_2, \ldots, P_N, m_1, m_N, B_1, \ldots, B_N \) is generated and used to transmit the first frame. Once the frame is successfully received, the trial is evaluated using (4.1) and the transmission parameters are adjusted. Hence, a new trial is generated to transmit the next frame. The evaluation process is repeated again to generate a new better solution until the maximum number of desired solutions in the population is reached \((\text{Population}_{\text{size}})\). Then, the selection process starts to create the median population. Reminder Stochastic Sampling is exploited as a selection method according to [163]. During the selection process, the population is assumed to form a pie chart where each solution trial is assigned a space in the chart that is proportion to its fitness value. Next, an outer roulette wheel is placed around the pie with \((\text{Population}_{\text{size}})\) equally-spaced pointers. A single spin of the roulette wheel will now simultaneously pick all the members of a median population. After the selection is complete, the median population is created. An Elitism of 10% is considered to duplicate the best fit set of solutions to the next population to make the algorithm converges faster. Elitism is defined as the process of selecting the best trials from the given solutions set according to the total fitness measure and adding them to the new solutions set without performing crossover and mutation. However, the elitism percentage is maintained to be low to reduce the penalty of limiting the opportunities of exploring other solutions.

Crossover and mutation operations are then applied to the remaining trials. In the
crossover process, two trials are re-combined in a single point with each other to form two new offsprings [164]. Mutation is an operation of altering binary bit from zero to one, or vice verse, applied to trials after crossover [165]. The choice of the crossover and mutation rate is known to critically affect the behavior and performance of the adaptation scheme. The crossover rate controls the capability of exploiting trials to reach the local optima. The higher the crossover rate, the faster exploitation proceeds. A large crossover rate would disrupt solutions faster than they could be exploited. The mutation rate controls the speed of exploring a new set of solutions. Mutation rate is always chosen to be small to avoid damaging the structure of the good trials. Adaptive crossover and mutation rates can, however, traverse different search directions in the state space, thus affecting the performance of the applied algorithm. The overall performance of our algorithm depends on maintaining an acceptable level of productivity throughout the process of evolution. Thus, it is essential to choose the appropriate crossover and mutation rates. We have employed an adaptive crossover and mutation techniques that adjust crossover and mutation probability dynamically. The first step in our approach is to determine an initial crossover and mutation rate ($RC$) and ($RM$). We choose a large initial $RM$ to help to explore more solutions and locate a prospective area quickly. During this time, crossover should occur with a small probability to retain diversity. Thus, $RC$ and $RM$ are selected to be 0.6 and 0.4 respectively. Then, crossover and mutation factors are defined as $CF$ and $MF$ as in Equations (4.9) and (4.10) respectively.

$$CF = f_{parents} - f_{offsprings}$$ \quad (4.9) \nonumber

$$MF = f_{new} - f_{old}$$ \quad (4.10) \nonumber

where $f_{parents}$ is the sum of fitness function of the parents and $f_{offsprings}$ is the sum of fitness functions of offsprings. The fitness function of the offsprings after and before
mutation are denoted by \( f_{\text{new}} \) and \( f_{\text{old}} \) respectively. At the end of every crossover and mutation process, \( RC \) and \( RM \) are adjusted according to Equations (4.11) and (4.12) respectively.

\[
RC = 
\begin{cases} 
RC + \eta_1, & \text{if } CF_{\text{avg}} > MF_{\text{avg}} \\
RC - \eta_1, & \text{if } CF_{\text{avg}} < MF_{\text{avg}} 
\end{cases} \tag{4.11}
\]

and

\[
RM = 
\begin{cases} 
RM + \eta_2, & \text{if } MF_{\text{avg}} > CF_{\text{avg}} \\
RM - \eta_2, & \text{if } MF_{\text{avg}} < CF_{\text{avg}} 
\end{cases} \tag{4.12}
\]

where \( CF_{\text{avg}} = \frac{\sum CF}{n_c} \) and \( MF_{\text{avg}} = \frac{\sum MF}{n_c} \) which are the average of \( CF \) and \( MF \) respectively, and \( n_c \) is the number of times crossover and mutation processes are executed. The choice of \( \eta_1 \) and \( \eta_2 \) is adaptive according to the convergence of the algorithm. \( \eta_1 \) and \( \eta_2 \) are calculated as follows,

\[
\eta_1 = \eta_2 = 
\begin{cases} 
0.02 \frac{f_{\text{max}} - f_{\text{avg}}}{f_{\text{max}} - f_{\text{min}}}, & \text{if } f_{\text{max}} > f_{\text{min}} \\
0.02, & \text{if } f_{\text{max}} = f_{\text{min}} 
\end{cases} \tag{4.13}
\]

where \( f_{\text{max}}, f_{\text{avg}}, f_{\text{min}} \) are the maximum, average, and minimum fitness value of trials in the population respectively. After crossover and mutation, a new iteration of AMO CE will start and the new population will be re-evaluated during the transmission of the next frames and the process will continue until the stopping criteria is met. The execution of AMO CE stops when either of the following occurs: convergence to a stable solution or if the maximum number of generations is reached.

One thing to note is that AMO CE tackles the problem of getting stuck in local optima. The adaptive crossover and mutation rates increase the solution space diversity and introduce new possibilities. In addition, the selection method used to create
the median population tends to diversify the resulted solutions. All these factors help the algorithm does not get stuck in local optima. The AMO CE implementation also ensures that the final solution is compatible with policy regulations by setting limits in the fitness functions for the transmission parameters such as power and bandwidth, which make policy verification easier by the policy layer of the cognitive architecture.

4.1.3 Evaluation

In this section, the performance of AMO CE is evaluated using Matlab simulation. We simulate a multicarrier system with \( N_c = 64 \) subcarriers with sufficient cyclic prefix is assumed. Each subcarrier is assigned a random attenuation value \( |H_i|^2, i = 1, 2, \ldots N_c \) with chi-square distribution. Hence, the SNR varies independently from one subcarrier to another which induces a need for the power and rate adaptation for each individual subcarrier. The channel is assumed to be ”block-invariant”, implying that the transmission channel impulse response remains constant or undergoes only minor changes over several consecutive frame transmissions. The network topology, is assumed to be uniformly randomly distributed over a given area with consideration of the Rayleigh-distributed channel coefficients. Other simulation parameters are stated in Table 4.1.

The simulation targets three different scenarios to demonstrate the performance of AMO CE through the fitness score for all subcarriers and the speed of convergence. The first scenario is a low power mode scenario, where the most contributing fitness (objective) functions are minimization of power consumption and interference. The second scenario is the multimedia scenario, where delay minimization and throughput maximization are the main objectives. The last scenario is the emergency scenario where reliability is the transmission objective and this is represented by minimization of BER. Spectral efficiency is a generic objective that is essential in all transmission scenarios. In this simulation, we compare AMO CE with other meta-heuristic search-
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population size</td>
<td>50</td>
</tr>
<tr>
<td>Number of generations</td>
<td>1000</td>
</tr>
<tr>
<td>Congestion control</td>
<td>Depends on the buffer size</td>
</tr>
<tr>
<td>Number of hops</td>
<td>Multi-hop</td>
</tr>
<tr>
<td>Transmission range</td>
<td>d, d/2, d/4 and d/8 where d is the distance between communicating nodes</td>
</tr>
<tr>
<td>Channel coding rate ($R_c$)</td>
<td>$1/2, 1/3, 2/3, 3/4$</td>
</tr>
<tr>
<td>Contention window size ($CW$)</td>
<td>$CW_{min}$ 8 values between 4 and 64. $CW_{max}$ 8 values between 32 and 4096</td>
</tr>
<tr>
<td>Frame size ($L$)</td>
<td>24 bytes to 3072 bytes</td>
</tr>
<tr>
<td>Bandwidth ($B$)</td>
<td>2 to 128 MHz</td>
</tr>
<tr>
<td>Modulation type and order</td>
<td>M-PSK and M-QAM</td>
</tr>
<tr>
<td>Transmission power ($P$)</td>
<td>0.06 mw to 15 mw (-12dBm to 12dBm)</td>
</tr>
<tr>
<td>Symbol Rate ($Rs$)</td>
<td>125 Ksps to 4 Msps</td>
</tr>
<tr>
<td>Time Division Duplexing (TDD)</td>
<td>1% to 100%</td>
</tr>
<tr>
<td>Noise power</td>
<td>-30 dBm -5 dBm</td>
</tr>
</tbody>
</table>

Table 4.1: Parameters used in AMO CE simulation

...ing algorithms based engines including ACO [48], Artificial Bee Colony Algorithm (ABC) [166], PSO [49] and GA proposed by Newman in [137]. All these scheme aims to adapt radio parameters in a multi-carrier environment. Figures 4.1, 4.2, and 4.3 show the fitness score of the five schemes for low power, multimedia, and emergency scenarios respectively. AMOS achieves the highest score and has the fastest convergence compared to the other schemes. All the comparable schemes lacks adaptivity in decision-making optimization as their weights of the objective functions are fixed regardless of the continuous variation in the environment conditions. In addition, they focus only on physical layer parameters optimization. Both PSO and ACO are based on random decision and require more information for parameters selection. Newman GA uses fixed crossover and mutation rate and does not consider essential objectives such as delay and spectral efficiency.

In addition, we have conducted another set of simulations to demonstrate AMO CE adaptation with respect to TCP/IP stack layers. The first simulation focuses on physical layer parameters optimization. AMO CE is compared with the conven-
Figure 4.1: AMO CE fitness score comparison in low power scenario

Figure 4.2: AMO CE fitness score comparison in multimedia scenario

tional engine, which is based on predetermined transmission parameters in addition to searching engines compared before. Figure 4.4 presents the achieved throughput as a function of SNR for all the compared schemes with target BER equal $10^{-3}$. AMO CE outperforms other engines as it achieved the highest throughput. Low SNR represents the emergency and low power scenario, which favors reliability and power consumption on the account of throughput. High SNR values represent the multimedia scenario which targets throughput maximization on the expense of power
Figure 4.3: AMO CE fitness score comparison in emergency scenario

consumption and reliability. The conventional standard has the worst performance because it uses the same modulation order over all sub-carriers and has no optimization capabilities. The second simulation considers cross layer optimization of

Figure 4.4: Throughput of different cognitive engines with physical layer optimization

AMO CE in addition to physical layer optimization. Mainly, it aims to show the effect of adapting MAC layer parameters which include contention window and frame size. The choice of higher minimal contention window size ($CW_{\text{min}}$) reduces power consumption and increases the throughput. Figure 4.5 shows a comparison between
AMOS, the conventional scheme and "PHY then MAC" optimization engine. "PHY then MAC" adapts physical and MAC layer parameters separately. The physical layer parameters are optimized using the algorithm in [167] while the MAC-layer parameters are optimized using an exhaustive search. AMO CE achieves better throughput than both engines. The reason is that the conventional engine uses predetermined transmission parameters and "PHY then MAC" optimizes the contention window size and frame size based on target SNR and not the current BER value. This leads to non-optimal transmission parameters when the current BER is much smaller than the target BER at high SNR values. The third simulation focuses on network layer parameters adaptation, which include transmission range and number of hops as well as physical and MAC layer transmission parameters. In this simulation, we compare AMO CE, which has the capability to adapt transmission range and number of hops dynamically with conventional engine that uses fixed transmission range d/4, d (four hops and one hop respectively). Figure 4.6 shows that at low SNR, it is better to use more hops to ensure reliability. However, fewer hops are the best practice when SNR is high as the channel reliability is enough to transmit for longer distances. AMO

Figure 4.5: Throughput of different cognitive engines with physical and MAC layer optimization

Throughput (bits/symbol/user)

SNR (dB)

Conventional
PHY then MAC
AMOS

Figure 4.6: Throughput of different cognitive engines with physical and MAC layer optimization
CE is able to be consistent with the channel condition by modifying the transmission range and changing the number of hops.

![Figure 4.6: Throughput of different cognitive engines with cross layer optimization](image)

**4.2 Hybrid Cognitive Engine for Radio Parameters Adaptation**

In this section, we propose a hybrid CE that can fit with any cognitive resource management architecture from one side and addresses the trade-offs of different AI algorithms from the other side. The engine is capable to overcome the complexity problem and extends the range of configuration to include not only the physical layer. The hybrid engine comprises CBR for learning and DTs for parameters prediction. The design of the hybrid engine benefits from the rich experience of CBR to reduce the radio parameters adaptation time in domains with minimum knowledge [54]. DTs are useful for classifying situations with large amounts of data as they are simple and require minimum processing time. The performance of this hybrid engine is demonstrated through testbed implementation using SDRs USRP-N210 [168] in home network. In addition, simulation was conducted to validate the engine performance in multi-carrier environment. The home network was selected to be the
implementation environment as it is convenient and comprises multiple users with
different applications. Throughput, SINR, and Packet Error Rate (PER) are the
performance metrics used to quantify the performance of the hybrid engine.

4.2.1 Hybrid Engine Design and Operation Mechanism

This section presents the structure of the hybrid CE [19], which involves the radio
system parameters and the adaption process using CBR and DTs.

System Parameters

System parameters are classified into three groups: transmission parameters that are
tuned by the CE to reach the optimal state of operation, environment parameters,
which are the indicators that define the communication environment state, and per-
formance measuring parameters. The considered transmission parameters include
transmission power ($P$), modulation index ($m$), coding rate ($R_c$), bandwidth ($B$),
Frame size ($L$), contention window size ($CW$) and transmission range between each
two hops ($d$). The appropriate choice of these parameters guarantees efficient network
operation. Environment parameters are Interference power ($I$) and noise power ($N$)
which indicates the quality of the channel, path loss ($PL$), and ratio of successfully
received frames to all the transmitted frames ($SR$) which helps evaluate the current
transmission quality. Performance measuring parameters are the metrics exploited
to trigger the hybrid engine and quantify its performance. We selected throughput,
PER, and SINR to be the performance metrics as they are vital for any wireless
network performance evaluation.

Radio Adaptation Process

Radio adaptation is the process in which the CE decides on the best configuration
of radio transmission parameters that fits with the current environment state. The
engagement of the CE in parameters configuration is triggered by degradation in performance metrics (i.e. PER, throughput and SINR) and the priority of certain performance metric is determined according to the network scenario. The adaptation process consists of three stages as in Figure 4.7: observation, decision-making, and learning. Typically, observations include monitoring environment conditions and performance metrics that trigger the need for system adaptation. The control proceeds to the decision-making stage when the performance metrics are not within certain predefined thresholds. The decision-making exploits one of the following techniques: CBR decision-making, which is used in case the current environment conditions match with a decision-making case stored in the database. Otherwise, DTs decision-making is employed if CBR fails to find a case that matches the encountered scenario. The last stage is learning in which the decided configuration is recorded in CBR as a new configuration case if the feedback received from the environment is positive (i.e. the performance metrics are within acceptable ranges). The focus of this work is only on the decision-making and learning stages. Therefore, spectrum sensing and environ-
ment monitoring details are omitted. In the following sub-sections we illustrate the CBR, and DTs contribution in the decision-making process.

1. Case-Based Reasoning :

CBR is exploited to obtain the radio transmission parameters by matching the encountered scenario with list of cases for previous scenarios stored in the CBR database. Each case consists of three attributes: the first one identifies the environment, which is represented by (PL, SR, I) in our designed engine. The second attribute corresponds to the configuration of radio parameters obtained from DTs that was stored for the certain scenario. The last attribute records the quality feedback ($F_q$) received after the configuration is applied. At the initialization stage, CBR populates the case database to create the history and experience for the CE. The database is populated according to the implementation environment with different cases represented by different environment conditions. Case retrieval is the process that engages to fetch the cases that can be used to configure radio parameters for the current scenario. The similarity value is the factor that manage the retrieval process. The similarity value is a measure of how close the encountered scenario to the cases stored in the system database. This measurement relies on the environment parameters comparison. Let us assume that the case is stored as a vector of number where $E$ elements are used to describe the environment parameters and $C_i$ denotes the $i_{th}$ case. The similarity measurement between the two cases $C_i$ and $C_j$ can be found as follows,

$$\text{Sim}_{ij} = \sum_{z=1}^{E} \frac{TS_z}{|C_i(z) - C_j(z)| + 1}$$

(4.14)

where $TS_z$ is the similarity threshold for the z element and $C_i(z)$ is the $z_{th}$ element of $i_{th}$ case. As the considered elements for environment parameters are (PL, SR, I), if the absolute difference values of these elements between $C_i$ and
$C_j$ are zeros, then the similarity measurement of the two cases is 4 which is the maximum achievable value. We define $T_z$ as the threshold that determine if two cases are similar. If $Sim_{ij} > T_z$, then the two cases are similar enough to exploit the configuration that corresponds to that case. However, there can be more than one case similar enough to be exploited. Therefore, only the case with maximum ($F_q$) is selected. The CBR database is maintained by adding the configurations that have satisfactory performance.

2. Decision-Trees:

DTs are exploited as a machine learning technique [142] for prediction of transmission parameters. DTs consist of a set of hierarchical rules that divide the data into groups, where a decision is made for each group. We follow the Quinlan style of decision-tree algorithms, [169] which is based on utilizing observed data to create the decision-tree. The configuration of transmission parameters is linked to the estimation of performance metrics. PER is the indicator used for frame size adaptation. SINR measurement is exploited for bandwidth adjustment. For power and modulation adaptation, if the target PER is not achieved, then the modulation order is progressively reduced. If the minimum modulation order still does not lead the PER to reach its desired value, then power is adjusted. The engine decides how much to increase the power by estimating the interference power and SINR. The construction of the tree starts by checking, which one of these metrics has the highest priority and this is determined according to the radio environment. For example, if the objective is to maximize throughput, then, throughput is the metric that has the highest priority and it will be assigned as the root of the tree. After the root of the tree is specified, the branches are determined according to certain threshold and ranges. In each range, the action is taken to tune the transmission parameters or to consider
another performance objective. The sample decision-tree for the multimedia scenario is shown in Figure 4.8 provided that PL, I, and Noise are 10, -70, and -116 dBm respectively. We notice that the root of the tree is throughput per user with three different ranges: less than 300 kbps, between 300 and 600 kbps and above 600 kbps. Therefore, the typical range is between 300 and 600 kbps, which extends the tree to consider the second priority metric. If the throughput is above 600 kbps, the decision to be made is to decrease parameters including \((m_i, R_c, d, CW, L)\). If the throughput is less than 300 kbps, the action will be to increase these parameters. The second priority metric is PER, which involves three ranges including, above \(10^{-2}\), between \(10^{-2}\) and \(10^{-4}\), and below \(10^{-4}\). If PER is less than \(10^{-4}\), the engine increases parameters including \((m_i, R_c, d, CW, L)\) and reduces the bandwidth. On the other hand, when the PER is above \(10^{-2}\), all the parameters are decreased except the bandwidth. The last objective is SINR, which leads to the last three decisions shown in Figure 4.8. The decisions made are to increase, decrease or keeping the current configuration of bandwidth and power if the SINR typical range is not violated. The
length of the tree is determined according to number of the objectives involved in the decision-making process.

Algorithm 2 Hybrid Cognitive Engine Decision-making

Require: Environment Parameters (PL, SR, I), QoS metrics (Throughput, PER, SINR), Threshold \((th)\) for QoS metrics

Ensure: Transmission Parameters \((P, m, B, R_c, CW, d)\)

BEGIN

Case formulation: (Check environment parameters and monitor QoS metrics)

if \((PER \geq PER_{th} \text{ or } Throughput \leq Throughput_{th} \text{ or } SINR \leq SINR_{th})\) then

while \((Case_{size} \neq \text{max and out} = 0)\) do

Case Matching: check similarity value \((Sim)\)

if \((Sim \geq Sim_{th}) \text{ and } (F_q) \text{ is the maximum}\) then

Extract the corresponding case solution

Output \((P, m, B, R_c, CW, d), \text{out} = 1\)

else

\(Case_{size} \leftarrow Case_{size} + 1\)

end if

end while

if \((out = 0)\) then

Execute DTs decision-making

Output \((P, m, B, R_c, CW, d)^* \text{ out} = 1\)

end if

end if

Monitor performance metrics (Throughput, PER, SINR)

if \((PER \leq PER_{th} \text{ and } Throughput \geq Throughput_{th} \text{ and } SINR \geq SINR_{th})\) then

Update the CBR database,

end if

END

Algorithm 2 illustrates the adaptation process of the hybrid engine. The process engages only when it notices that the performance metrics are not in the typical ranges. If so the decision-making process advances to CBR to find a similar case with maximum quality feedback. If CBR cannot find a similar case, the DTs module engages for decision-making. The achieved performance is monitored to update the CBR with new cases that record adequate results. The \(Case_{size}\) in Algorithm 1 refers to the size of the tree branch that contains the cases that fall within the similarity predefined threshold.
4.2.2 Hybrid Engine Complexity Consideration

Studies conducted in [170] revealed that there are timing issues, when CBR is implemented using hardware testbed. A non-negligible latency is introduced in case searching, and certain timing deadlines may not be met, thus causing problems such as frames collisions. Therefore, latency and complexity in case searching and retrieval is an important issue to consider. In the efforts to solve these issues, we look for unique aspects of each case to reduce the overall searching time and this allows the CBR database to grow large for more complex networks, and maintains fast access time. Predefined thresholds for similarity relative to each radio parameter are used to index different cases within the database. The appropriate selection of the threshold for each radio parameter simplifies the determination of similarity. We take advantage of DTs module to split the search space into parts (nodes), which contain a number of similar cases according to the similarity calculated in (1). Thus, every node in the constructed tree represents a subset of the cases of the CBR database and the root leads to the whole cases. The tree design does not require any computation on deciding where to split the search space, as it is predefined before any case enters the database. The average achieved complexity for this tree-based indexing is \( O(\log(n)) \) where \( n \) is the number of cases in the CBR database. As our hybrid CE is mainly dependent on decision-trees, the average complexity of our DTs algorithm is \( O(\log(n)) \). The convergence speed investigated in the evaluation section shows how fast our hybrid engine in decision-making. We measure the access time achieved for different CBR base size and compare it to the traditional CBR access time. The measured access time presented in Figure 4.9 shows that our cases retrieval approach has minimum access time regardless of the CBR database size. This makes our engine capable to handle large scale networks with large database.
4.2.3 Hybrid Engine Testbed Implementation

We implemented two scenarios to demonstrate our hybrid engine radio adaptation capabilities. The first scenario focuses on reliability while the second scenario has various traffic where throughput is the main objective. Both scenarios are detailed in the next sub-sections.

Reliability Scenario:

This scenario incorporates a TCP file transfer between two nodes. Each node is connected to one USRP-N210 hardware for energy detection spectrum sensing and configuration. We evaluate the reliability of our hybrid engine under interference effect. The interference is formulated by using a third USRP-N210 that injects interference in a form of jamming signal (narrow-band interference). In this scenario, we compare our hybrid engine performance with non-cognitive (non-CE) radio that is incapable of changing its initial configuration parameters. In addition, we compare it with the cognitive system (CRM) implemented using USRP2 hardware in [171] and the hybrid engine (CBR+GA) proposed in [56]. Table 4.2 presents the parameters changes in reaction to the variation of PER and SINR to their threshold values. The measured values as function of time for PER and SINR are presented in Figure 4.10.
and Figure 4.11

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial Value</th>
<th>Optimized Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission power ($P$)</td>
<td>-22 dBm</td>
<td>-15 dBm</td>
</tr>
<tr>
<td>Frame size ($L$)</td>
<td>1024 Bytes</td>
<td>750 Bytes</td>
</tr>
<tr>
<td>Modulation type ($m$)</td>
<td>64-QAM</td>
<td>8-PSK</td>
</tr>
<tr>
<td>Coding Rate ($R_c$)</td>
<td>3/4</td>
<td>2/3</td>
</tr>
<tr>
<td>Bandwidth per sub-carrier</td>
<td>812.5 kHz</td>
<td>625 KHz</td>
</tr>
<tr>
<td>Contention Window</td>
<td>64 for min and 4096 for max</td>
<td>32 for min and 2048 for max</td>
</tr>
<tr>
<td>Transmission Range ($d$)</td>
<td>$d$</td>
<td>$d/4$</td>
</tr>
</tbody>
</table>

Table 4.2: Configuration parameters adaptation of hybrid CE

Figure 4.10: PER of different cognitive engines with narrow band interference in reliability scenario

Figure 4.11: SINR of different cognitive engines with narrow band interference in reliability scenario
As it can be seen that our hybrid engine outperforms the CRM and CBR+GA engines in terms of PER and SINR achieved. In addition, our engine is the fastest in convergence as it relies on CBR learning with DTs decision-making unlike the CRM engine which has no learning capability and CBR+GA, which only relies on complicated GA optimization. The average PER achieved by our engine is 0.006, which is less than the threshold 0.01, less than 0.025 achieved by CRM, and less than 0.011 recorded by CBR+GA. Moreover, Figure 4.11 shows that our hybrid engine is the only one that managed to keep SINR above threshold, which is 14 dB.

**Variable Traffic Scenario:**

In this scenario, we deploy two pair of nodes that have different types of traffic to send. The first one sends real time video traffic while the second one has best effort file sharing (FTP) traffic. The nodes were stationed in different rooms to avoid the high power leakage of WiFi cards that severely affects the expected performance. The traffic of the second pair of nodes was fixed to be on channel 11 while the first pair of nodes can select the best channel among the rest. Throughput measurement is the trigger of the hybrid engine in this scenario. Figure 4.12 shows the cumulative throughput of both pair of nodes achieved by our hybrid engine compared to CRM and CBR+GA engines. The measurements are plotted against the center frequency (channel number). The results in Figure 4.12 show that our hybrid engine achieved better throughput than the CRM and CBR+GA systems. One thing to note is that the throughput decreases as the transmission channel is closer to channel 11. Therefore, the best practice is to transmit on channel that is far from channel 11.

**4.2.4 Simulation Evaluation**

We simulated a multi-carrier system with 64 sub-carriers. Each sub-carrier was assigned a random attenuation value to simulate a dynamic channel. Hence, the SNR
Figure 4.12: Network throughput in variable traffic scenario

varied for each channel, inducing a need for the adaptation for each individual channel. The performance of the hybrid CE is compared to GA based CE proposed in [172], ANNs based engine proposed in [173], the hybrid engine (CBR+PSO) in [55], and with GA optimizer (CBR+GA) in [56]. The hybrid engine performance is demonstrated using two evaluation metrics, which are PER and throughput. Table 4.3 presents the ranges of the transmission parameters, CBR related parameters, and performance metrics. The average PER is measured and plotted in Figure 4.13 for all the engines. Our hybrid engine achieved the minimum PER with the highest convergence speed. The CBR+PSO achieved less PER than the regular GA and ANNs.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission range ($d$)</td>
<td>$d$, $d/2$, $d/4$ and $d/8$</td>
</tr>
<tr>
<td>Contention window size ($CW$)</td>
<td>$CW_{\text{min}}$ 8 values between 4 and 64. $CW_{\text{max}}$ 8 values between 32 and 4096</td>
</tr>
<tr>
<td>Frame size ($L$)</td>
<td>24 bytes to 1024 bytes</td>
</tr>
<tr>
<td>Bandwidth ($B$)</td>
<td>312.5 to 812.5 KHz</td>
</tr>
<tr>
<td>Channel coding rate ($R_c$)</td>
<td>1/2, 1/3, 2/3, 3/4</td>
</tr>
<tr>
<td>Modulation type and order</td>
<td>M-PSK and M-QAM</td>
</tr>
<tr>
<td>Transmission power ($P$)</td>
<td>-50 dBm to -10 dBm</td>
</tr>
<tr>
<td>Case Base size</td>
<td>400</td>
</tr>
<tr>
<td>PER</td>
<td>$5 \times 10^{-3}$ (threshold) to $10^{-2}$ (tolerable)</td>
</tr>
</tbody>
</table>

Table 4.3: System parameters and performance metrics for hybrid engine simulation
Although CBR+GA converges to low PER, the iteration consumes more time than CBR+PSO.

![Figure 4.13: Average PER for different cognitive engines](image)

Figure 4.13: Average PER for different cognitive engines

Figure 4.14 presents the average user throughput achieved by the CEs with variation of the number of nodes in the network in comparison with other CEs.

![Figure 4.14: Average user throughput for different cognitive engines](image)

Figure 4.14: Average user throughput for different cognitive engines

We notice that the advantage of the hybrid engine is clear at large number of nodes as all the algorithms except GA manage to achieve good performance with 15 or less nodes connected to the base station. Table 4.4 presents the results of average time consumed for each decision-making for all the engines. The minimum time
recorded by our hybrid engine is (0.052 sec) for 400 case CBR reflects the complexity consideration of the CBR case retrieval.

<table>
<thead>
<tr>
<th>Average Time</th>
<th>GA</th>
<th>ANN</th>
<th>CBR+PSOCBR+GA Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.162 sec</td>
<td>0.125 sec</td>
<td>0.22 sec</td>
</tr>
</tbody>
</table>

Table 4.4: Time consumption of different cognitive engines


Although CE is the structure known for deciding system parameters adaptation using optimization and machine learning techniques, these techniques have strengths and weaknesses depending on the experienced network scenario that make one is more appropriate than others. This creates the need to re-design the CE structure from a new perspective, which is the capability to select the most appropriate CE technique for the encountered network scenario. In this section, we develop a Supervised Cognitive System (SCS) that integrates typical CE with two additional components which are: evaluation component and control component. These two components distinguish SCS from typical CE designed for radio parameters adaptation in the literature. The evaluation component performs a comprehensive evaluation of the CE learning techniques in different environment scenarios. Our proposed technique for evaluation monitors the CE’s performance represented by certain metric at each decision step during its operation. Then, it estimates the relation between the learning of the CE and the achieved performance metric. The control component employs online learning to categorize the encountered scenarios and assigns the most appropriate learning technique to perform the system parameters adaptation accordingly. Online learning is chosen as it does not require detailed model or previous knowledge about the network and it can perform classification in the real-time. The performance of the
proposed SCS is evaluated using testbed and simulation and compared against other CEs proposed for system adaptation. The obtained results reflect the advantage of the addition of the evaluation and control components to the typical CE.

4.3.1 Supervised Cognitive System Architecture

Our new approach to re-design the cognitive system has the capability to predict the performance of the involved CEs adaptation algorithms in various network scenarios and it is able to automatically select the most appropriate learning algorithm with the dynamic environment [20]. This is achieved by introducing the evaluation and the control components to evaluate and control the learning process at a higher level. However, there are two critical challenges that must be tackled to realize the SCS system. The first challenge is that the evaluation of the CE algorithms is not trivial as it is necessary to find a way for monitoring the performance of individual CE techniques. Despite the fact that CEs should be able to sense their environment and autonomously adapt to the changing network conditions, different CE designs provide different levels of situation awareness and cognitive functionality. In order to tackle this challenge, we monitor the relation between learning evolution of the CE and the achieved performance. The second challenge is the ability to decide, which CE algorithm is the most appropriate one for parameters configuration according to the network scenario, operating objectives, learning technique, and radio’s capabilities. Thus, we employ model-free online learning to develop the CE selection algorithm that fulfill these requirements.

The SCS components including adaptation component, evaluation component, and control component. (see Figure 4.15) are described in the following sub-sections.
Adaptation Component

The adaptation component incorporates the CE learning algorithm exploited to configure radio parameters. This component is generic and has the capability to accommodate any learning or optimization algorithm. In this implementation, we exploit $\epsilon$-greedy, ANNs, and online learning as the learning algorithms of the adaptation component. A dedicated memory is exploited to store learning techniques related parameters. The $\epsilon$-greedy technique based on RL \cite{174} randomly explores different radio configuration parameters with probability $\epsilon$ and it exploits the radio configuration that achieves the maximum performance with probability $1 - \epsilon$. This technique guarantees that all the configuration options are explored before convergence. $\epsilon$ is the factor that determines the rate of exploration. Although high exploration rate helps to reach near optimal configuration quickly but this occurs on a penalty over the outcome because of the exploration cost. The Multilayer Feedforward Neural Networks (MFNNs) \cite{175} is exploited as the second learning technique and it provides
a non-linear input-output relation with function approximation capabilities. It can accommodate large number of inputs and outputs and estimates performance metrics based on certain observation. Thus, it is a good candidate for system adaptation design. Q-learning [59] is the third CE technique, which performs decision-making without detailed modeling of the environment and improves its decision quality by interacting with it. Experience is the only factor required by Q-learning to operate in an optimal fashion and it is acquired from online interaction with the environment. The learner benefits from interacting with the environment to select an action that maximizes a cumulative reward. In addition, this learning technique achieves context awareness and intelligence in CR networks without training complexity [176].

Goodput, which is the rate of the successfully received packets at the receiver is the performance objective considered in the configuration process. The environment conditions are represented by SNR $\gamma$ and eigen-spread $\kappa$ of the communication channel also known as the Demmel condition number [177]. $\kappa = \lambda_{\text{max}}/\lambda_{\text{min}}$ are the maximum and minimum eigen-values of the $HH^H$ matrix. It is assumed that the channel is constant during the transmission time of the packet. The configuration parameters considered in this paper are transmission power, frequency, frame size, and modulation order. Each of these configuration parameters is represented by a set of discrete values and the CE algorithm selects the most appropriate value of the parameter based on its learning methodology. The details of the learning techniques are omitted due to the space and they are out of the paper scope. Reader can refer to [178] [179] [180] for more information about using these techniques in cognitive resource management.

**Evaluation Component**

The evaluation component’s function is to evaluate the performance of the employed CE techniques in the adaptation component in order to make the control component
capable to classify the encountered environment scenario and selects the most appropriate CE technique to perform transmission parameters configuration. The output of the evaluation component is the feedback about the performance achieved as a result of using certain adaptation technique. The evaluation metric considered in the system design is goodput, which is calculated based on the BER and Symbol Error Rate (SER). The coding gain $G_c$ for a convolutional codec using soft decision decoder at $BER=10^{-3}$ is calculated as,

$$G_c = \frac{rd}{2}$$

(4.15)

where $r$ is the rate and $d$ is the free distance of the code. The hard code decision reduces the gain by factor of two. A hard decision decoder is used in the V-BLAST decoder because the received symbols need to be hard estimated as a part of the successive interference cancellation employed. Assuming Gray coding, the BER can be approximated from the SER as,

$$BER = \frac{1}{\log_2(M)}SER$$

(4.16)

where $M$ is the modulation index. Let’s assume that a packet with $N$ bits can tolerate a maximum $n_e$ error, then the goodput $GP$ can be calculated as follows,

$$GP = \sum_{i=0}^{n_e} \binom{N}{N-i} BER^i (1 - BER)^{N-i}$$

(4.17)

Using (6.34) and (6.35), $GP_u$ denoted the upper bound of the goodput can be estimated using the AWGN SER curves. Similarly, $GP_l$ (the lower band goodput) can be estimated using the fading channel SER curves. Both upper and lower bounds of the goodput are utilized for further evaluations and the resulting evaluations are used to initialize the experience database.
**Control Component**

This component is the core of SCS as it coordinates radio parameters configuration, identifies the encountered network scenario, and selects the appropriate technique in the adaptation component for parameters configuration. Identification is achieved by grouping data into sub-populations according to its similarities, on the basis of a training set of data containing instances whose category membership is known. To realize the identification process, we need to define various environment states and their impact on performance objectives. This definition is achieved by extracting some distinct features of the environment, which are defined by SNR, eigen spread and performance objectives such as goodput maximization. The estimated SNR and eigen-spread must be discretized for indexing. The range for eigen-spread is from 0 to 12, with 25 equally spaced value. Similarly, the range for SNR is from 0 to 50 dB, with 51 equally spaced values.

Enhanced online learning is the technique employed to perform identification of network scenario and selection of CE technique to perform system configuration. Online learning is a model-free RL technique. Specifically, it can be used to find an optimal action-selection policy for any given MDPs. It works by learning an action-value function that ultimately gives the expected utility of taking a given action in a given state and following the optimal policy thereafter. MDPs include a discrete set of environment states $S$ and a discrete set of actions $A$. At each time step $t$, the agent acquires network state information $s$ and selects certain action $a$ to perform. Consequently, the environment makes a transition to state $s'$ at time step $t + 1$ with probability $T(s, a, s')$ and receives certain reward $R(s, a)$. This process is iterative and repeated infinitely to converge to an optimal decision-making policy $\pi$ that maximizes the total received reward. This policy is a mapping from environment states to probability distributions over actions. The value function is defined as a
measurement of the quality of the action policy adopted and defined as follows,

\[ V^\pi(s) = E[R(s, \pi(s))] + \beta \sum_{s' \in S} T_{ss'}(\pi(s))V^\pi(s') \] (4.18)

where \( E \) is the expectation operator and \( \beta \in [0, 1) \) is the discount factor. As there is no prior knowledge about the state transition probability, the optimal policy Q-value for certain state/action pair is defined as,

\[ Q^*(s, a) = E[R(s, a)] + \beta \sum_{s' \in S} [T(s, a, s') \max_{b \in A} Q^*(s', b)] \] (4.19)

In this model free learning, the agent tries to find \( Q^*(s, a) \) recursively using the following update rule,

\[ Q^{t+1}(s, a) = (1 - \alpha^t)Q^t(s, a) + \alpha [R(s, a)^t + \beta \max_{b \in A} Q^t(s', b)] \] (4.20)

where \( \alpha \in [0, 1) \) is the learning rate. If each action is executed in each state infinitely and the learning rate decays properly, the \( Q^t(s, a) \) will converge to \( Q^*(s, a) \) with probability 1 as \( t \to \infty \). The online learning parameters including state, action, transition and reward functions are defined in our system as follows,

- **State:** the state \( s \) is defined as the network scenario denoted as \( s = \{\gamma, \kappa\} \).

- **Action:** the action \( a \) is the selection of the appropriate adaptation technique that maximizes the goodput. It is represented by a binary sequence to identify the adaptation technique as follows, \( a = (00) \) for \( \epsilon \)-greedy, \( a = (01) \) for ANN, and \( a = (10) \) for the Q-learning.

- **Transition function:** the transition function \( T(s, a, s') \) is the probability to
switch from one state to the other and is defined as follows,

\[ T(s, a, s') = Pr(s(t + 1) = s' | s(t) = s, a(t) = a) \] (4.21)

- Reward: the reward function \( R(s, a) \) is defined as the achieved goodput.

The procedure followed to select the most appropriate CE learning algorithm is explained in the next section.

4.3.2 SCS Operation

As the SCS aims to select the most appropriate adaptation technique to adapt radio parameters, this adds more complexity to the system as it requires evaluation the performance of different adaptation techniques in different scenarios. To overcome this issue, we design the system operation such that the evaluation process of the adaptation technique is activated on demand. This demand is triggered by monitoring the recorded PER, which represents the environment feedback of adaptation. With low PER, the system retains the current adaptation technique. Otherwise, the system activates the evaluation component to search for the best adaptation algorithm for the current scenario. In this way, the number of adaptation times required is reduced and this is demonstrated through the adaptation time recorded in the simulation section. The SCS operation flow goes as follows and illustrated in Figure 4.16.

First, the output (performance metric) of each CE learning algorithm is observed for certain time in different scenarios and all the observations are analyzed. Then, the evaluation process engages to check the appropriateness of each CE in the specified scenario and it is compared with other CEs. The observation of the CE is performed and output is monitored at each decision step in multiple runs in order to gain a good estimation about the performance of each CE technique. As a result, we can obtain the relationship between learning and the output achieved. Figure 4.17 presents a sample
of the average goodput achieved for all CEs learning techniques employed as a function of time steps consumed for learning. Figure 4.17 shows that Q-learning achieves minimal performance at the beginning as it relies on trial and error learning without any prior knowledge about the network environment. However, it converges to the maximum performance later compared to the other CE techniques as it has already obtained enough experience about the network environment. The $\epsilon$-greedy based CE with $\epsilon = 0.01$ does not assume any prior estimates about the performance metrics. Thus, the exploration is performed in a way that all the configuration parameters have the same probability to be selected. ANNs relies on offline training using a designated training set consists of network scenarios and their corresponding parameters configurations. Therefore, $\epsilon$-greedy and ANNs are not as aggressive as Q-learning and provide steady performance increase. After the evaluation stage, the SCS has enough insights about the appropriateness of each CE in different network scenarios. Thus, it can proceed to the classification of the current network scenario and selection of the CE technique that match with the experienced conditions. Online learning is
exploited to perform identification of network scenario and selection of CE technique as it has the capability to explore different network scenario and it is not limited to certain training data. Initially, we assume that there is no data about evaluation of CE algorithms. Therefore, the SCS selects CE techniques to perform parameters configuration randomly from the available CEs with probability $\tau$. After the SCS has gained experience about network scenarios and their corresponding evaluation with different CEs, it starts the exploitation process with probability $1 - \tau$. The exploitation process uses the obtained experience (old states) to handle the new experienced network scenarios (new state) based on the similarity calculated using the Euclidean distance between the current state and the old states $D_E = \sqrt{s(t) - s(x)}$. The SCS experience up to time step $t$ is given by the $\sigma$-algebra as follows,

$$\mathcal{F}(t) = \sigma\left(\{s(x), a(x)\}_{x=1}^{t}, \{R(s(x), a(x))\}_{x=1}^{t-1}\right)$$

(4.22)

where information about the network state $s(x)$, associated action $a(x)$ and the reward $R(s(x), a(x))$ can be extracted from the evaluation component. The system checks the distance between the current state $s(t)$ and $s(x)$ in $\mathcal{F}(t)$, and obtains a set $S'(s(t), \mathcal{F}(t))$, which includes $F$ different most recents states from $\mathcal{F}(t)$.
that minimize $\sum_{f=1}^{F} D_E(s(t), s(x_f))$. The reward function $R(s(t), a)$ is set to be 1 if $a = \arg \max_{a' \in A} Q^*(s(t), a')$ and 0 otherwise. The notation $A'(s(t))$ is used to represent the set of actions that achieve $R(s(t), a) = 1$. Let $z$ be an integer that satisfies $1 \leq z \leq F$. The SCS picks $z$ records $A^\tau(s'(t), F(t))$ with respect to $S'(s(t), F(t))$. The SCS performs selection action $a(x^*)$ if $(x^*) = \max_x \{x | a(x) \in A^\tau(s'(t), F(t)) \cap A'(s(t))\}$. Otherwise, SCS selects CEs algorithm randomly from $A^R(s(t)) = \{a | a = \arg \max_{a \in A} R'(s(t), a)\}$, where

$$R'(s(t), a) = R(s(t), a) \frac{\omega(s(t), a)}{z}$$  \hspace{1cm} (4.23)

This is calculated using $z$ records drawn randomly from the $F$ most recent performed actions where $\omega(s(t), a)$ is the number of times this action was selected before for $s(t)$. The procedure for CE algorithm selection using online learning is explained in Algorithm 3.

**Algorithm 3** Online learning for CE technique selection

**Require:** current state $s(t)$, old states $s(x), F(t)$

**Ensure:** selection of CE algorithm that output max $GP$

1. **Learning:** given network state $s(t)$;
2. Exploration with probability $\tau$;
3. Select action randomly;
4. Update $A'(s(t)) = \{a | R(s(t), a) = 1\}$ for $s(t)$;
5. Exploitation with probability $1 - \tau$;
6. Select $z$ records $A^\tau(s'(t), F(t))$ out of $F$ actions associated with $S'(s(t), F(t))$;
7. Calculate $R(s(t), a)$ according to (9) and populate $A^R(s(t))$
8. **if** (exist $x^* = \max_x \{x | a(x) \in A^\tau(s'(t), F(t)) \cap A'(s(t))\}$) **then**
9. Select action $a(x^*)$ ;
10. **else**
11. Select action from $A^R(s(t))$;
12. **end if**
13. Update $R(s(t), a)$;
14. $s = s^{t+1}$
15. $t = t + 1$
4.3.3 SCS Performance Evaluation

In this section, we evaluate the SCS performance through testbed experiments and simulation to demonstrate its capability of adaptation by selecting the most appropriate CE learning technique. The evaluation network structure consists of multiple ad hoc nodes each one has SCS system installed. The performance of SCS is compared with other systems in literature including [172] [179] [56] that performed testbed implementation to demonstrate the capability of their proposed CEs. The work in [40] proposed a GA based CE for parameters adaptation. The authors in [179] proposed CE that exploits ANNs to learn dynamically how to select the channel, which is expected to yield the best performance for the mobile users while the work in [56] proposed a hybrid engine design using CBR and GA. The hybrid engine relies on cases matching using CBR to expedite its convergence. In addition, we compare our simulation to the results obtained by the meta engine proposed in [58].

Testbed Evaluation Results

The considered testbed consists of one pair of nodes accessing the available network resources in IEEE 802.11 typical environment. The experimental setup comprises picture file transfer between each two nodes. The configuration parameters ranges are specified as follows: (-20 dBm to -10 dBm) for transmission power, (20 to 600 bytes) for the frame size, and (BPSK, QPSK, 8-PSK, 16-QAM, 32-QAM, and 64-QAM) for modulation. Each node in the testbed is equipped with USRP-N210, and RFX2400 daughter-board mounted on top of it. The USRP-N210 is used as the radio front-end to perform physical layer information acquisition and parameters adaptation in combination with liquid-DSP [181] and GNU radio [182] platform. To facilitate the ease of implementation, we used an IEEE802.11 feedback network, which is usually supported through an internal wireless network interface controller (NIC). SCS performs data processing for the radio front end, as well as message parsing and
takes the appropriate action. The implementation involves two radios, one for control exchange and the other for connection and data transfer between nodes. However, we use only one radio to perform both functions by using the concept of virtual WiFi as in [183]. We evaluate the performance of the SCS through three experiments.

1. Typical Office Environment Experiment:

The first experiment is conducted in an office environment whose layout is shown in Figure 4.18. The sender and receiver are located in the same room. We evaluate the SCS performance by the following metrics: the rate of correctly received packets (Goodput), PER, and spectral efficiency (SE). The average of the evaluation metrics during the experiment is recorded in Table 4.5 for the SCS and compared with the ANNs [179] based CE, GA [155] based engine and the hybrid engine CBR+GA proposed in [56]. The results show that our proposed SCS achieved the highest goodput, minimal PER and the highest SE compared to other CEs. Figure 4.19 plots the achieved goodput as function of the offered load. The offered load is the percentage of utilization of the system capacity. Figure 4.19 demonstrates the ability of the SCS to maximize the
goodput compared to superior cognitive systems proposed for radio parameters adaptation.

<table>
<thead>
<tr>
<th>Scheme Name</th>
<th>Office Experiment Results</th>
<th>Scheme Name</th>
<th>Office Experiment Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Goodput (Kbps)</td>
<td>PER</td>
<td>SE (bits/s/Hz)</td>
</tr>
<tr>
<td>ANN</td>
<td>420</td>
<td>0.0162</td>
<td>4</td>
</tr>
<tr>
<td>GA</td>
<td>480</td>
<td>0.0153</td>
<td>4.8</td>
</tr>
<tr>
<td>CBR+GA</td>
<td>600</td>
<td>0.0126</td>
<td>5.9</td>
</tr>
<tr>
<td>SCS</td>
<td>850</td>
<td>0.0088</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Table 4.5: Average performance metrics in the office experiment

Figure 4.19: Average goodput with various network load in the office experiment

2. Propagation Impact Experiment

The second experiment evaluates the performance of the SCS under signal propagation impact. Figure 4.20 shows the layout of the second experiment, where sender and receiver are located in far away isolated rooms. The goodput, PER, and SE for this experiment are recorded in Table 4.6. We notice that the SCS achieved 30 % improvement in the goodput, 65 % reduction in the PER, and 38 % raise in the SE compared to the best engine performance, which is the hybrid engine. Figure 4.21 compares the PER in two different locations: same office as in the first experiment and far away isolated rooms for all the cognitive adaptation engines. Figure 4.21 confirms the ability of the SCS to maintain the
Figure 4.20: Propagation experiment layout

PER at very low level compared to the other cognitive systems proposed for radio parameters adaptation. In addition, we notice that there is an increase in the PER and drop in the goodput and SE for all the schemes in the second experiment compared to the first. The reason for that is the increase in the distance between the sender and the receiver and environment signal propagation impact.

<table>
<thead>
<tr>
<th>Scheme Name</th>
<th>Goodput (Kbps)</th>
<th>PER</th>
<th>SE (bits/s/Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>380</td>
<td>0.073</td>
<td>2.8</td>
</tr>
<tr>
<td>GA</td>
<td>350</td>
<td>0.08</td>
<td>2.3</td>
</tr>
<tr>
<td>CBR+GA</td>
<td>490</td>
<td>0.043</td>
<td>3.9</td>
</tr>
<tr>
<td>SCS</td>
<td>700</td>
<td>0.0193</td>
<td>6.3</td>
</tr>
</tbody>
</table>

Table 4.6: Average performance metrics in the propagation experiment

3. Interference Impact Experiment

The interference impact is investigated in the third experiment, where an additional interference source is added as a USRP-N210 device that transmits signals by varying center frequency as in the layout presented in Figure 4.22.

The performance metrics measured in this experiment are recorded in Table 4.7.

The table shows that SCS achieved the highest goodput, minimum PER and
maximal SE compared to the ANNs, GA and CBR+GA CEs. It is also evident that the performance of all systems has dropped because of the interference effect compared to the previous experiments in terms of all performance metrics. Figure 4.23 plots the achieved SE against the SNR in the communication channel. The figure demonstrates the capability of the SCS system compared to other engines. It also shows that the SE remains stable after certain SNR value as there is a limit for transmission power that can enhance the SE.

The results in Figures ??, 4.21 and 4.23 demonstrate the capability of the SCS
<table>
<thead>
<tr>
<th>Scheme Name</th>
<th>Interference Experiment Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Goodput (Kbps)</td>
</tr>
<tr>
<td>ANN</td>
<td>340</td>
</tr>
<tr>
<td>GA</td>
<td>335</td>
</tr>
<tr>
<td>CBR+GA</td>
<td>460</td>
</tr>
<tr>
<td>SCS</td>
<td>670</td>
</tr>
</tbody>
</table>

Table 4.7: Average performance metrics in the interference experiment

![Graph showing SE against SNR](image)

Figure 4.23: Average SE against SNR in the interference experiment

in system adaptation. This highlights the advantage of making the CE supervised and able to select the most appropriate adaptation technique to perform system parameters adaptation. In addition, the figures points to the fact that each one of the engines exploited in the comparison has various performance at different network conditions. For example, GA achieves better spectral efficiency than ANNs at low SNR while ANNs outperforms GA at high SNR. Moreover, GA records better goodput than ANNs at the first experiment but worse goodput in the second one.

**SCS Simulation Results**

We simulated a multi-carrier system with 64 sub-carriers. Each sub-carrier was assigned a random attenuation value to simulate a dynamic channel. Hence, the SNR
varied for each channel, inducing a need for the adaptation for each individual channel. The performance of the SCS is compared to the performance achieved by the ANNs based scheme [179], the hybrid engine (CBR+GA) [56] and the meta engine [58]. The SCS performance is demonstrated using three evaluation metrics which are goodput, PER, and SE. Table 4.8 presents the ranges of parameters for a system deployed under multiple users environment.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Iterations</td>
<td>1000</td>
</tr>
<tr>
<td>Exploration rate $\epsilon$</td>
<td>0.3</td>
</tr>
<tr>
<td>Number of users</td>
<td>Variable</td>
</tr>
<tr>
<td>Frame size ($N_b$)</td>
<td>24 bytes to 1024 bytes</td>
</tr>
<tr>
<td>Bandwidth ($B$)</td>
<td>312.5 to 812.5 KHz</td>
</tr>
<tr>
<td>Channel coding rate ($R_c$)</td>
<td>1/2, 1/3, 2/3, 3/4</td>
</tr>
<tr>
<td>Modulation type and order</td>
<td>M-PSK and M-QAM</td>
</tr>
<tr>
<td>Transmission power ($P$)</td>
<td>-30 dBm to -10 dBm</td>
</tr>
</tbody>
</table>

Table 4.8: Transmission parameters adaptation ranges in SCS simulation

Figure ?? presents the goodput achieved by each regular CE ($\epsilon$-greedy, ANNs, Q-learning) and the one achieved by SCS. SCS selects one of the three learning techniques that is the most appropriate according to network scenario variation. The scenario variation is conducted by altering SNR to have the range of 0 to 50 dB and eigen-spread varies between 0 and 12. The initial parameters configuration for

Figure 4.24: Performance of SCS in different network scenarios

power, modulation, coding rate, frame size and bandwidth follows a naive style where
power is configured at the minimum while modulation, frame size and bandwidth are at the maximum level. The configuration changes as the SCS becomes aware of the environment and the involved adaptation techniques. The performance metric considered for this evaluation is the total goodput achieved in a session with network scenario that varies every 100 time step. Figure 4.24 clearly shows how the SCS switches between learning techniques to keep goodput at the maximum level. It also shows that the aggressive learning algorithms like Q-learning perform better than ANNs and $\epsilon$-greedy at high SNR levels (i.e., between 100 to 200 and 400 to 500 time steps).

The rest of the simulations compare the performance achieved by SCS with the ANNs based scheme [179], the hybrid engine (CBR+GA) [184] and the meta engine [58]. We evaluate the convergence speed for the competing schemes for parameters adaptation. The normalized goodput is measured as a function of the time epoch in Figure 4.25. We observe that our proposed SCS achieved the highest goodput with the fastest convergence as shown in Figure 4.25 after 200 epoch. GA based engine is the slowest in convergence due to large computation required to perform adaptation. In addition, we evaluate the reliability and spectral efficiency of transmission by plotting the average PER and SE versus SNR in Figure 4.26 and Figure 4.27 respectively.
We notice that SCS outperforms all other engines. The reason is that the hybrid engine assumes that CBR is the main technique for adaptation and GA (slow in convergence) can be used to enrich the experience of CBR. The meta engine does not consider functionality of fully implemented system for resource management and does not account for convergence time. However, the SCS performs extensive evaluations to rank the CE algorithms in different scenarios and employs enhanced online learning to determine the most appropriate adaptation technique. Finally, we present the average time consumed to perform parameters adaptation in all the systems in Table[4.9]. The table shows that SCS achieved a comparable results, which is little higher than the
Table 4.9: Average adaptation time for different cognitive systems

<table>
<thead>
<tr>
<th></th>
<th>ANN</th>
<th>GA</th>
<th>CBR+GA</th>
<th>SCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>0.523 sec</td>
<td>1.123 sec</td>
<td>0.185 sec</td>
<td>0.202 sec</td>
</tr>
</tbody>
</table>

hybrid CBR+GA system thanks to the on-demand approach for activating the SCS evaluation component and altering the adaptation technique. However, the time consumption pitfall is acceptable as the SCS manages to boost the CE performance according to the performance metrics investigated in the testbed implementation and simulation.

In this chapter, we introduced and examined several machine learning techniques for the CE engine in a wireless cognitive system. Performance results showed that machine learning can further efficiently allocate the available spectrum according to the user demands with maximum QoS. In the next chapter, we consider CR contribution with embedded machine learning in the LTE networks to handle resource allocation and interference problems. We embed CogWNet in LTE networks and attempt to further optimize the user application QoS.
Chapter 5

Cognitive Radio with Learning Capability for Efficient Resource Management in LTE Networks

RRM is a challenging problem for future wireless networks specifically cellular networks as the demand for data services is increasing dramatically. The use of CR in RRM can enrich the performance of LTE networks by providing environment awareness and real-time decision-making for tuning transmission parameters. CR resource management can be implemented at the terminal or network level. In LTE networks, it can be exploited for frequency allocation, power control, and modulation adaptation in order to improve throughput and limit interference. These objectives are subjected to QoS requirements, user location, network operators, and network policies. In addition, The lack of resources of the macrocells networks makes them unable to fulfill the data services demand in the indoor areas in LTE networks. An efficient solution is to deploy FBSs, which are capable of communicating users over a broadband wire-line connection. FBSs are short-range, low-cost/low power and can be easily installed by the users in addition to the fact that they reduce the load of the MBSs. However, interference is considered as a technical challenge that affects the femtocells deployment.

In this chapter, CogWnet is embedded in the LTE infrastructure for RRM. The designed optimization architecture provides support for QoS requirements by maximizing throughput and mitigating interference using intelligent frequency assignment. A dedicated communication interface is utilized to gather spectrum information from
physical and MAC layers to be used later for decision-making. In addition, we propose an interference mitigation scheme that aims to mitigate cross-tier interference caused by FBSs to the MUEs and the co-tier interference that affects the FBSs that contend to access the free sub-channels in the Downlink.

### 5.1 Cognitive Based Resource Blocks Allocation and Radio Parameters adaptation in LTE Systems

The use of CR in resource management can enrich the performance of LTE networks by providing environment-awareness and real-time decision-making for tuning transmission parameters [5]. In LTE networks, it can be exploited for RBs allocation, power allocation, modulation adaptation in order to improve throughput and limit interference. These objectives are subjected to QoS requirements, user location, network operators and network policies. The nature of LTE technology makes it easy to integrate with cognitive solutions for RRM. In this section, we propose CR resource management scheme that is integrated with our cognitive based architecture (CogWnet) proposed in [34] for RRM in LTE networks. It also provides support for QoS requirements by maximizing throughput and interference mitigation using intelligent RBs assignment. Cross-layer optimization is used in the architecture to facilitate environment awareness and transmission parameters configuration. A dedicated communication interface is utilized to gather spectrum information from physical and MAC layers to be used later for decision-making. Bandwidth, modulation, and transmission power are adjusted to satisfy the QoS requirements and adapt to environment conditions. Learning mechanism is introduced to reduce complexity, expedite the adaptation process and improve the decision-making quality.
5.1.1 Throughput Optimization and Interference Management Model

Throughput is a fundamental performance metric for any spectrum allocation framework to keep at an ultimate level. Throughput threshold is used to determine the minimum acceptable throughput value that can meet QoS requirements. In order to increase throughput, it is necessary to increase other parameters such as modulation index, which will increase the number of bits/symbols. However, this increase has to be tied to the environment changes reflected by BER levels and application demands. The following network model is considered to formulate throughput maximization. The model employs utility based optimization, where a utility function is specified for the total throughput for users assigned to different RBs. The utility function maps the network resources a user utilizes into a real number. Utility based optimization balances the efficiency and fairness. A set of users $U = \{ x : x = 1, ..., M - 1 \}$ are considered to be served by the network. Another set of frequency sub-carriers $FC = \{ y : y = 1, ...K - 1 \}$ represents the available sub-carriers in the network. The throughput that can be achieved by a user $x$ assigned to sub-carrier $y$ is,

$$T_{(x,y)} = (1 - P_e)N_cM_iR_c \tag{5.1}$$

where $P_e$ is the BER, $N_c$ is the number of sub-carriers, $M_i$ is the modulation index and $R_c$ is the coding rate. The total throughput in which the user obtaining a service is,

$$T_{all} = \sum_y [a_{(x,y)} * T_{(x,y)}] \tag{5.2}$$

where $a_{(x,y)}$ is an assignment indicator for the RBs. If $a_{(x,y)} = 1$ then the RBs is assigned to the user and $a_{(x,y)} = 0$ otherwise. The utility function selected to capture the user satisfaction about the assigned sub-carriers is,

$$U_i = 0.16 + 0.8ln(T_{all} - 0.3) \tag{5.3}$$
The target of the system configuration is to maximize this utility while maintaining the following constraint: $T_{all} \geq T_{min}$ where $T_{min}$ represents the threshold of minimum throughput to satisfy application needs [21].

Interference is another essential performance metric to consider especially in such heterogeneous network. In LTE context, mutual interference occurs between BSs, which are called E-UTRAN Node Bs (eNBs). These eNBs are contending to utilize the available radio RBs. Therefore, the RBs assignment must take place in a way that limit this interference effect and maximize the throughput of the eNBs. The considered LTE network model to mitigate interference is based on SINR measurement. This model has multiple eNBs that cover various Hexagonal cells. The available spectrum is divided into a set of RBs. Each mobile terminal measures the Reference Signal Received Power (RSRP) for itself, associated eNB, and neighboring eNBs for all the eNBs. RSRP is an LTE specific metric that averages the RF power in all of the reference signals in the passband. The SINR is calculated by taking the average of these RSRP measurements. The SINR is recorded for all the RBs used by each eNB. As the network conditions such as network load, noise, and BER are changing, periodic updates between adjacent eNBs are exchanged. These updates include RSRP measurements related to the eNB itself and its neighbors as not all the RBs experience the same network conditions. The eNBs that the terminal is associated with is denoted by eNB(w). SINR over each RB is estimated according to the frequency used by the eNBs, the power received by the terminal, and the cell load factor measurements at each eNBs. The cell load factor is defined as the amount of resources consumption in relation to that is available in the cell. The load value grows with the cells’ traffic demand and the amount of inter-cell interference. The
SINR of eNB\((w)\) link is,

\[
SINR_{eNB(w)} = \frac{P_w f_w}{\sum_{a \in S^{(w)}_{srb}} P_{a} f_{a} L_{a} L_{w} + n_0}
\] (5.4)

where \(P_w\) and \(P_a\) are the powers received by the terminal from eNB\((w)\) and other eNBs respectively. \(f_w\) and \(f_a\) are the frequencies of the RBs used by eNB\((w)\) and other eNBs respectively. \(L_w\) and \(L_a\) are the cell load factors for eNB\((w)\) and neighboring eNBs respectively and \(n_0\) is the noise. \(S^{(w)}_{srb}\) is the set of neighboring eNBs that have used the resource block \(srb\). The eNBs of others eNBs’ links is,

\[
SINR_{eNB(a)} = \frac{P_a f_a}{P_w f_w L_w + \sum_{b \in S^{(w)}_{srb}} P_{b} f_{b} L_{a} L_{b} + n_0}
\] (5.5)

where \(S^{(w)}_{srb}\) is the set of the rest eNBs that use the resource block \(srb\).

For RBs allocation, \(SINR_{target}\) is selected based on the network status to be the reference for \(SINR_{eNB(w)}\) and \(SINR_{eNB(a)}\) to compare with. \(SINR_{target}\) is updated periodically when the system acquires the environmental information based on the application type and the network conditions. This enhances the decision made for spectrum assignment. The decision about the RBs assignment is determined by the following,

\[
srb_{dec1} = SINR_{eNB(w)} - SINR_{target}
\] (5.6)

\[
srb_{dec2} = SINR_{eNB(a)} - SINR_{target}
\] (5.7)

If \(srb_{dec1} \geq 0\) and all the results of (5.7) are greater than or equal zero as (5.7) is applied for multiple eNBs. Then, the RB is selected. Note that SINR is checked for both eNB\((w)\) and other neighbors to make the selection for the RB that was not used before by other eNBs, or used by the furthest and least number of eNBs.
5.1.2 Integration of LTE and CogWnet

Figure 5.1 presents the integration of LTE components and the RRM scheme. The LTE components comprise one LTE User Equipment (UE) and one LTE eNB. LTE UE is the terminal equipment that includes all user activities and is compatible with the 3rd Generation Partnership Project (3GPP) standard. LTE eNB represents the network base station that communicates with LTE UE. In addition, it has CogWnet architecture installed, which executes RRM functions and controls the network activities. The communication interface between CogWnet and the LTE platform consists of two components: an interpreter and physical/MAC layer adapter. The interpreter is implemented in the eNBs to interpret CogWnet messages without changing LTE code or specifications. The PHY/MAC-layer adapter receives input from LTE PHY/MAC-layer and modifies the parameters of these layers according to the decisions made by CogWnet. The eNB is connected to an application server on the Internet to retrieve the data requested by the LTE user. The system always monitors
SINR, throughput, BER, and traffic load and reacts to changes in these parameters.

5.1.3 Learning Mechanism to Enhance Cognitive RRM

The proposed cognitive RRM scheme adopts RL\cite{64} to improve the quality of decision made to configure the transmission parameters. User satisfaction with the obtained service is used as a metric to represent the quality of the decision made by the CogWnet and this should reflect the goals and the needs of the system. Throughput and interference are evaluated through the time and compared with thresholds as discussed in their presented model. If the measured interference is below the threshold and the throughput is above the threshold, then the user is satisfied. In addition, a database is considered in the design of the cognitive RRM scheme to store instant interactions and decisions affecting the radio environment. These interactions are exploited to make the system identifies situations encountered in the past and react. As a result, the optimization becomes faster as some complex optimization procedure will not be repeated again. The system continues to evolve gradually until it becomes aware of the best spectrum configuration. The database in the repository of CogWnet has a table that records environmental parameters, which represent the status of the network and the corresponding solution represented by the transmission parameters. This table is updated after each optimization process. A new module is added to the repository in CogWnet as shown in Figure \ref{5.2} to perform the matching between the instant environment conditions and the table reference conditions. The flow of control in the repository takes place according to the following steps:

1. The Initial Phase: Repository receives the environmental parameters from TCP/IP stack layers through access interfaces.

2. The Matching Phase: Matching module in the repository is triggered to check if there is a match for the current environment situation. The module will consult the database table for matching.
Figure 5.2: CogWnet with learning mechanism

3. If there is a match, the corresponding configuration parameters will be passed to the action module in the repository.

4. If there is no match, the sensory information will be processed to the decision module to run the normal optimization procedure.

5. When the optimization procedure ends, the new parameters are used to configure the radio.

6. The Update Phase: The new configuration is sent to the repository database and a new entry for this solution is added to the table.

### 5.1.4 Cognitive RRM Scheme Functionality

The cognitive RRM scheme accounts for RBs allocation and adaptation of transmission parameters based on the environment input to achieve its overall goals. This adaptation increases LTE network efficiency by increasing throughput, improves spec-
The cognitive RRM scheme works as follows. When the system starts, CogWnet discovers the existing LTE platform. Then, it conducts periodic sensing to identify unoccupied channels. Unoccupied channels are the channels whose received RSRP for certain sub-carrier is less than a given threshold. When a free channel is found, CogWnet requests LTE eNB or FBSs (femtocells deployment) to collect information about the LTE environment such as BER, cell load factor, and SINR values. During the environmental information acquisition, data frames cannot be sent. Finally, CogWnet architecture acquired all the necessary information for decision-making. It runs its optimization for the proper adaptation of transmission parameters. All policies are loaded dynamically to the core of CogWnet. In case the scheme detects a new service that requires a higher bit rate within a cell that is overloaded, the scheme considers allocation of more RBs for the associated eNB. These RBs are freed when the service is terminated. The measurement of SINR is used to manage RBs allocation and interference control. If SINR for the RB of the associated eNB is low, the scheme looks for another RB with higher SINR. BER is exploited to adapt the modulation index for throughput maximization purpose. The modulation index is increased for better throughput if the detected BER is low. On the other hand, the modulation index is decreased if the LTE link quality is poor. Transmission power is adjusted according to the measured SINR for certain RB. If the link experiences high interference, the transmission power is decreased, and vice versa. In addition, the transmission is switched to a different frequency to eliminate interference if there exists other free channels, which can meet the applications’ requirements.
5.1.5 Cognitive RRM Scheme Performance Evaluation

Scenario 1: Typical Multiple eNode-B LTE Network

RRM is challenging task in LTE networks as there is a high demand from eNBs for the spectrum to satisfy various application QoS requirements. Consequently, interference is a fundamental issue in the RRM process in LTE networks. In this scenario, extensive simulations have been performed to demonstrate the advantage of using this cognitive scheme for LTE RRM. The feasibility of cognitive RRM is represented by the optimization of network resource allocation and decision-making improvement. This is achieved by exploiting CogWnet capabilities to reconfigure LTE system parameters. These parameters include bandwidth, coding rate, modulation index, and transmission power. The performance of the system is evaluated using a topology, which has 20 eNBs and a total bandwidth of 100 MHz. The transmission power for each eNB is uniformly divided over the RBs. The throughput of the downlink is used to evaluate the performance of this model. Table 5.1 presents a list of simulation parameters for the LTE network modeling and their default values. Figure 5.3 presents the throughput achieved by our model compared with typical LTE interference management scheme without cognitive capability (random scheme). The throughput is evaluated against the number of active eNBs. The LTE interference management scheme without cognitive capability allocates radio RBs for each frame based on predetermined measurements and without being aware of radio resources information. It is similar in concept to interleaved resource blocks allocation to combat the block fading channel [187]. The non-cognitive scheme has the reuse factor of 1 due to the high network load in the tested scenario. Figure 5.3 shows that our cognitive scheme is superior compared to the non-cognitive scheme especially when the network has more active eNBs. In addition, the achieved capacity of the LTE system using the cognitive RRM scheme is compared with the one obtained by the random scheme for
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Bandwidth</td>
<td>100 MHz</td>
</tr>
<tr>
<td>RB Bandwidth</td>
<td>180 KHz</td>
</tr>
<tr>
<td>Number of available RBs</td>
<td>500</td>
</tr>
<tr>
<td>User distribution</td>
<td>Uniform</td>
</tr>
<tr>
<td>Background noise in downlink (PN)</td>
<td>-102 dBm (One RB PN /24)</td>
</tr>
<tr>
<td>Max eNB Tx power</td>
<td>20 W</td>
</tr>
<tr>
<td>Multipath Fading</td>
<td>3GPP-Case 3</td>
</tr>
<tr>
<td>Sub-carrier spacing ($\Delta f$)</td>
<td>15KHz</td>
</tr>
<tr>
<td>Frequency reuse factor</td>
<td>3</td>
</tr>
<tr>
<td>Modulation and coding scheme</td>
<td>16-QAM: 1/2, 2/3, 3/4</td>
</tr>
<tr>
<td></td>
<td>64-QAM: 1/2, 2/3, 3/4, 4/5</td>
</tr>
<tr>
<td>Cell layout</td>
<td>Hexagonal grid</td>
</tr>
<tr>
<td>Cell radius</td>
<td>167 m</td>
</tr>
<tr>
<td>Scheduler</td>
<td>Proportional fair queuing</td>
</tr>
</tbody>
</table>

Table 5.1: LTE environment simulation parameters

Figure 5.3: System throughput comparing cognitive scheme with non-cognitive LTE scheme

RRM. LTE system capacity is specified as the maximum constant arrival rate that can be supported by the system subjected to a given required QoS. Figure 5.4 shows that the cognitive RRM scheme outperforms the non-cognitive scheme in terms of the achieved system capacity as a function of the traffic load. Figure 5.5 shows a normalized throughput comparison between our approach and the random scheme against
the number of active users. The Figure shows that our approach performs much bet-

er than the random scheme as it achieves a 24% higher throughput for 30 active users per cell. The reason is that the random scheme has limited interference mitigation capability and has no throughput optimization. Overall, the cognitive solution has shown great potential for overcoming the RRM problem in LTE networks.

Another evaluation is conducted to demonstrate the quality of decision-making of
the proposed scheme for LTE system configuration. This quality is represented by user satisfaction probability, which is the ratio between the achieved throughput to the maximum possible throughput. Figure 5.6 shows the achieved user satisfaction probability as a function of the network load. Results shows that the cognitive system achieves better performance than the legacy LTE resource management, the load-balancing resource management [188] that provide access to the least loaded base station and the JRRM scheme proposed in [189]. Complexity decrease and how fast is the decision-making process for RRM is tested by the measure of the probability of success in matching between the current scenario and previously experienced ones. Figure 5.7 depicts the evolution of successful matching probability. The evolution is tested over 5 different environment conditions. It starts from 0 as no solution is recorded and it keeps increasing as more and more solutions are recorded. Figure 5.7 shows the advantage of the learning mechanism, which includes saving resources, faster configuration and less complexity.

**Scenario 2: Femtocells LTE Environment**

FBSs share the same spectrum with the MBSs, which leads to cross tier interference. Interference also can occur between FBSs themselves as their coverage areas are highly
overlapped with each other. In addition, FBSs are user-installed base stations. This makes the operator so not know the locations of FBSs when MBSs connections are established, and cannot efficiently mitigate the interference from FBSs. Therefore, there are three main requirements in femtocells LTE deployment: autonomous cross-tier and intra-tier interference mitigation, (ii) QoS guarantees provisioning for femtocells users, and (iii) fully radio resources exploitation.

In this evaluation scenario, we aim to utilize our cognitive scheme interference mitigation and throughput maximization model to control interference and maintaining QoS requirements for femto users. In addition, we compare the performance of our cognitive scheme with three other schemes that aims to manage interference problem in the femtocells environment. The first scheme, which is called Cognitive Radio Resource Management (CRRM) [190] treats the femtocells environment as a DSA environment where femtocells are secondary and macrocells are primary. It conducts periodic sensing to identify the radio resources usage by the MBSs and only allocate the unoccupied channels (overlay mode). The second scheme, which is called Decomposition based Resource Management (DBRM) [191] exploits cognitive capability with dual decomposition to manage spectrum sharing in overlay mode. Moreover, it involves joint power control scheme for downlink transmission. The
third scheme is a Cognitive Priority Based Resource Management (CPRM) scheme [192] for LTE systems. It dynamically adjusts the sensing period to achieve certain false alarm probability. Moreover, the CPRM scheme assigns priority values and minimum numbers of transmission bits to serving terminals to guarantee their QoS requirements and utilize spectrum efficiently. The simulation environment is compatible to 3GPP standard. It consists of one Macrocell divided into three sectors. Each sector has three femtocell blocks and each block contains 15 FBSs. The LTE network parameters are stated in Table 5.2. The LTE system supports four traffic types: voice and video traffics of real-time service, HTTP traffic of non-realtime service, and FTP traffic of best effort service. Each traffic type has the same number of terminals and the number of the Macro users is double the one for femto users. The QoS requirements for each traffic type are shown in Table 5.3.

![Figure 5.8](https://via.placeholder.com/150)

Table 5.2: LTE femtocells environment simulation parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macrocell radius</td>
<td>1000 m</td>
</tr>
<tr>
<td>Carrier Frequency</td>
<td>2 GHz</td>
</tr>
<tr>
<td>System Bandwidth</td>
<td>5 MHz</td>
</tr>
<tr>
<td>MBSs transmission power</td>
<td>43 dBm</td>
</tr>
<tr>
<td>FBSs transmission power</td>
<td>20 dBm</td>
</tr>
<tr>
<td>Thermal noise density</td>
<td>-174 dBm/Hz</td>
</tr>
<tr>
<td>Multipath channel</td>
<td>six-taps Rayleigh-faded path</td>
</tr>
<tr>
<td>Modulation and coding scheme</td>
<td>QPSK</td>
</tr>
<tr>
<td></td>
<td>16-QAM: 1/2, 2/3, 3/4</td>
</tr>
<tr>
<td></td>
<td>64-QAM: 1/2, 2/3, 3/4, 4/5</td>
</tr>
</tbody>
</table>

Table 5.3: The QoS requirements for each traffic type in LTE femtocells

<table>
<thead>
<tr>
<th>BER</th>
<th>Voice</th>
<th>Video</th>
<th>HTTP</th>
<th>FTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delay</td>
<td>40 ms</td>
<td>100 ms</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Packet Rate Dropping</td>
<td>1%</td>
<td>1%</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Data Rate</td>
<td>N/A</td>
<td>N/A</td>
<td>100 kbps</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Figure 5.8 depicts the throughput of all LTE UEs (including macro and femto
users) achieved by our scheme compared to CRRM, DBRM and CPRM schemes. The throughput is plotted as a function of traffic intensity \( p \) which is the ratio of the total arrival rates of all traffics over the maximum transmission rate per macro sector. Note that it exceeds 1 as FBSs are included.

![Throughput of UEs in LTE femtocells topology](image)

**Figure 5.8**: Throughput of all UEs in LTE femtocells topology

Figure 5.9 presents the packet dropping rate for video UEs for all the schemes.

![Packet dropping rate of video UEs in LTE femtocells topology](image)

**Figure 5.9**: Packet dropping rate of video UEs in LTE femtocells topology

From Figure 5.8 and Figure 5.9, we notice that our cognitive scheme (CogWnet) achieved the highest throughput and the minimum dropping rate. This shows the
feasibility of the throughput maximization model and the dynamic power adjustment
based interference control model. In addition, the awareness approach and cross layer
optimization of cogWnet make the FBSs able to select the most suitable channel
with proper radio configuration that guarantee minimum interference to the macro
UEs and keeps their performance at acceptable levels. Non of the proposed schemes
has any cross layer or learning capability. Both CRRM and DBRM follow overlay
channel allocation approach, which reduce the utilization of radio resources. The
CRRM has no power control mechanism and it randomly assigns the available sub-
channels to selected UEs, Therefore, it has the lowest performance. The DBRM uses
fixed and predetermined power assignment for the FBSs, which degrades the system
performance because of interference. Moreover, it does not consider QoS requirements
so it is just spectrum allocation scheme. The CPRM scheme targets multimedia
networks, which makes it not comprehensive for heterogeneous UEs applications.

5.2 Cognitive Aware Interference Mitigation Scheme for LTE
Femtocells

In this section, we propose an interference mitigation scheme that aims to mitigate
cross-tier interference caused by FBSs to the MUEs and the co-tier interference that
affects the FBSs that contend to access the free sub-channels in the downlink [22].
The scheme enhances spectrum sensing and improves detection capability to find free
sub-channels for FBSs to access in which the cross-tier interference is minimal. An
adaptive power graph coloring spectrum assignment algorithm is used in conjunction
with environment awareness to allocate sub-channels for FBSs that mitigates the
co-tier interference. In addition to interference control, the scheme ensures that the
selected sub-channel satisfies QoS requirements and matches with the traffic load.
Other advantages of the proposed scheme include considering traffic priorities, and
maintaining efficient utilization of the available sub-channels, which enhances the
spectrum efficiency.

5.2.1 System Model

We consider the interference problem of the downlink in a network that consists of macrocells and femtocells where the priority of sub-channel access is for the macrocell users. FBSs can access the sub-channel but with minimal interference to the MUEs. Spectrum sensing is employed to detect the available sub-channels that FBSs can access. Spectrum occupancy information is used to allocate the free sub-channels according to the awareness based algorithm to be discussed in the next section. The considered deployment employs OFDMA as a channel access technique for the downlink with \( M \) hexagonal grid macrocells and \( F \) femtocells in range of each macrocell. The bandwidth allocated for each MBSs is divided into 6 SBs using FFR. Each SB is composed of \( N_c \) sub-channels. The macro users can access any of these sub-channels at any time instant. However, the sub-channels are not utilized most of the time.

The femtocells considered deployment is depicted in Figure 5.10 where they are distributed randomly and uniformly in each SB. Each femtocell is assumed to have variable number of users active at any time instant. The sub-channels are assumed to be almost static with minor variations and follow Rayleigh multi-path fading distribution. Femtocells deal with two types of connections, which are the link between femtocells and macrocells and the link between FBSs and their associated users. There are three types of gains considered in the signal propagation model and they contribute to the total channel gain calculated in (5.8). These gains include antenna’s gain \( A \), shadowing gain \( S \) and path loss gain \( G \).

\[
H = A + S + G
\] (5.8)
The SINR of a macrocell user $k$ over sub-channel $n$ is calculated as,

$$SINR_{k,n} = \frac{P_{k,n}H_{k,n}}{I_1 + I_2 + N_{n,k}} \quad (5.9)$$

where $P_{k,n}$ is the received power of the macro user $k$ over sub-channel $n$, $I_1$ and $I_2$ are the two interference imposed by the other MBSs and FBSs respectively and $N_{n,k}$ is the Additive White Gaussian Noise (AWGN) power. The two types of experienced interference by the macro user $I_1$ and $I_2$ are calculated according to (5.10) and (5.11) respectively.

$$I_1 = \sum_{l=1}^{M} P_{l,n}H_{k,l,n} \quad (5.10)$$

$$I_2 = \sum_{j=1}^{F} z^* P_{j,n}H_{k,j,n} \quad (5.11)$$

where $P_{l,n}$ and $P_{j,n}$ are the transmission powers of the other MBSs and FBSs over the $n$th sub-channel respectively and $z^*$ is the factor that indicates if the sub-channel is assigned to a certain femtocell. It takes a value of 1 if the sub-channel is assigned.
and 0 otherwise. \( l \) and \( j \) are the indexes of the MBSs and FBSs respectively. The achievable throughput by the macro user \( k \) over sub-channel \( n \) is given by,

\[
T_{k,n} = B \log(1 + SINR_{k,n})
\]

(5.12)

where \( B \) is the sub-channel bandwidth.

Following a similar process, the SINR for a femto user \( i \) served over sub-channel \( n \) is calculated as,

\[
SINR_{i,n} = \frac{P_{i,n}H_{i,n}}{I_1 + I_2 + N_{n,i}}
\]

(5.13)

The interference imposed by MBS \( I_1 \) and the other FBSs interference \( I_2 \) are calculated according to (5.14) and (5.15) respectively.

\[
I_1 = \sum_{l=1}^{M} P_{i,n}H_{i,l,n}
\]

(5.14)

\[
I_2 = \sum_{j=1, j \neq i}^{F} z^*P_{j,n}H_{i,j,n}
\]

(5.15)

The achievable throughput of the femto user \( i \) over sub-channel \( n \) is given by,

\[
T_{i,n} = B \log(1 + SINR_{i,n})
\]

(5.16)

The cross-tier interference that impacts the performance of the macrocell users is caused by either miss detection during spectrum sensing or hidden macro users problem. Due to the limitation on the software and/or the hardware sensing capability, interference is caused to the macro users as a result of an incorrect detection. Note that the probability of miss detection depends on the sensing methods, (e.g., the energy detector, the cyclostationarity-feature sensing, and the matched-filtering sensing). Matched-filtering is known as the best approach for spectrum sensing as
it maximizes the received SINR \[^{193}\]. However, it is difficult since it requires dedicated receiver for each signal. The performance of energy detector is limited by the energy threshold and the types of signals. Besides, it fails when the noise becomes non-stationary because of the presence of the cross-tier interference. However, energy detector is the easiest to implement in actual systems. The hidden macrocell users problem is similar to the hidden node problem in CSMA. It is caused by many factors including severe multi-path fading, shadowing, and high penetration loss in the areas sensed by femto-cells.

Contending between FBSs for channel access especially in the dense femtocells deployment is the main reason to encounter co-tier interference that affects the performance of the femto users. Other problems like hidden terminal and exposed terminal problems also contribute to this interference as the femtocells network is similar to other wireless network once free channels are detected. Adjacent channel interference is another type of interference that affects the macro users on the edge of cell. It is caused if different but adjacent channels are occupied by the macro users and FBSs respectively. However, this interference can be leased by reasonable layout of base stations deployment.

5.2.2 The Interference Mitigation Scheme

In this section, we describe the interference mitigation scheme to control both cross-tier and co-tier interference that impacts femtocells operate under the coverage of macrocells. The scheme also aims to maximize throughput, allocate sub-channels that satisfy QoS requirements by assigning priorities for different types of traffic, and ensure efficient spectrum utilization. Spectrum sensing is exploited to support this scheme in order to mitigate cross-tier interference. Both interference mitigation mechanisms are detailed in the following sections.
Cognitive Based Cross-tier Interference Mitigation

Cognitive spectrum sensing is employed by the FBSs to determine whether certain SB includes free sub-channels. This forces the FBSs to cease their channel access if the sub-channel is busy with macro user transmission. If all SBs are busy, the FBSs tries to access the SB with the minimal interference to macro users. The presence of MBSs transmissions is detected in the Downlink signal. Our scheme implements an enhanced energy detection based spectrum sensing that effectively explores the interference range and maximizes the detection sensitivity. According to the energy detection approach, the signal observed by the FBSs is expressed as,

\[ y(x) = h(x)s(x) + w(x) \]  \hspace{1cm} (5.17)

where \( s(x) \) is the signal transmitted by the MBS, \( h(x) \) is the channels gain from the MBS to FBS, \( w(x) \) is the AWGN sample, and \( x \) is the sample index. The average received energy is given by,

\[ Y(X) = \frac{1}{X} \sum_{x=0}^{X-1} |y(x)|^2 \]  \hspace{1cm} (5.18)

where \( X \) is the total number of samples. Spectrum sensing aims to distinguish between the following two hypotheses,

\[ H_0 : y(x) = w(x) \]  \hspace{1cm} (5.19)

\[ H_1 : y(x) = h(x)s(x) + w(x) \]  \hspace{1cm} (5.20)

The hypothesis \( H_0 \) is for miss-detection and \( H_1 \) is for correct detection. Energy detection is defined by two probabilities, the probability of detection \( P_D \) and the probability of false alarm \( P_F \). The occupancy of sub-channels by macrocell users can
be determined by comparing the metric $Y$ against a threshold $\lambda$. Therefore, the $P_D$ is calculated as follows,

$$P_D = Pr(Y > \lambda|H_1) = Q_m(\sqrt{2 \ast SINR}, \sqrt{\lambda})$$  \hspace{1cm} (5.21)

where $m$ is the product of time and bandwidth and $Q_m(.,.)$ is the generalized Marcum Q-function [194]. The $P_F$ is calculated as follows,

$$P_F = Pr(Y > \lambda|H_0) = \frac{\Gamma(m, \lambda/2)}{\Gamma(m)}$$  \hspace{1cm} (5.22)

where $\Gamma(.)$ and $\Gamma(.,.)$ are the complete and incomplete gamma functions, respectively [34]. Both probabilities are calculated for each SB as the product function of all sub-channels contained in each SB.

We enhance the normal energy detection procedure to improve the detection capability. The decision regrading channel access by FBS is determined by quantifying how harmful is the interference caused to the macro receivers if the FBS uses the sub-channel. The interference to the MUEs is deemed to be harmful if it causes the signal-to-interference ratio (SIR) at the macro receiver to fall below a threshold $SIR^*$. This threshold depends on the macro receivers robustness toward interference and varies from one service to another. In addition, it may depend on the characteristics of the interfering signal (e.g., signal waveform, continuous versus intermittent interference, etc.) [195]. From the above definitions, we define the interference range of the FBS as the maximum distance from a macro receiver at which the incurred interference is still considered harmful. Consequently, the interference range depends on the macro user interference tolerance not just the FBS transmission power. Let $P_m$ and $P_f$ denote the transmission power of the MBS and the FBS respectively. The distance between the macro-cell transmitter and receiver is denoted as $R$. The
interference range of the FBS ($D$) is determined according to,

\[ \frac{P_m R^{-\alpha}}{P_f D^{-\alpha}} = SIR^* \]  \hspace{1cm} (5.23)

where $\alpha$ is the path loss factor. We deduce from (16) that the macro receiver can tolerate the interference caused by FBS as long as the distance between them is greater than $D$. As a result, we can define the detection sensitivity ($DS$) as the minimum SNR of the MBS at which an FBS should be still capable of detecting the macro signal. The FBS should be able to detect active macro transmission within a radius of $R + D$ according to 5.24. Therefore, the sensitivity is calculated as follows,

\[ DS = \frac{P_m (D + R)^{-\alpha}}{N_0} \]  \hspace{1cm} (5.24)

The spectrum sensing is conducted periodically to stay aware of any MBS starts to transmit. During the sensing period, the QoS degradation incurred by the MUEs in accessing the band is determined. The choice of the sensing period depends on the type of the service running on the MUE terminal and has to be set for each SB. For example, the sensing period is less for services that vary over a much larger time scale.

**Awareness Based Co-tier Interference Mitigation**

In this part, we develop an awareness based algorithm to mitigate co-tier interference between femtocells that are using the same SB. The algorithm aims to maximize system throughput, improve spectrum efficiency, and adapt transmission power according to FBS SINR requirements. It is assumed that each group of FBSs are assigned to certain SB and able to access its sub-channels. Each FBS is aware of its interference profile which is characterized by certain interference weight $W_i$. This weight is exploited to label the interference over the link between any two FBSs and
is calculated as,

\[ W_i = \sum_{e_i} 10^{I_i/10} \]  

(5.25)

where \( I_1 \) is \( I_2 \) that is calculated using (5.14) and (5.15) respectively. Due to the relatively small distance between the FBS and its associated user, the throughput maximization problem for a femtocell can be written as follows,

\[
\max \sum_{i \in V} \sum_{j \in V} \sum_{n \in N_c} z_i^* z_j^* \log(1 + \frac{P_i}{\sum_{j \neq i} P_j H_{ji} + N}) 
\]  

(5.26)

where \( V \) is the group of FBSs that share the same SB and \( z^* \) is the assignment indicator of the sub-channel, \( P_i \) is the transmission power of the FBS, \( H_{ji} \) is the total gain between FBS \( j \) and FBS \( i \), and \( N \) is the corresponding noise power. Note that \( i \) and \( j \) are the indexes of the interfering FBSs.

The FBSs channel allocation problem can be modeled as a graph coloring problem with support of an awareness based mechanism between the FBSs sharing the same SB. The awareness mechanism aims to share the interference weight, data rate requirement, traffic type and traffic load information for each FBS among other FBSs sharing the same SB. Consequently, each FBS is aware of its network environment. Interference weight is the basic metric for the graph coloring channel assignment while data rate requirement, traffic type and traffic load are exploited to ensure QoS, assign traffic priority, and improve spectrum utilization. The sub-channel assignment process as in Algorithm 4 starts by mapping the network into a bidirectional graph \( GR = (V, E, W) \) with a group of vertices \( V = \{v1, v2\ldots\} \) where each vertex \( v \) represents an FBS and a group of edges \( E = \{e1, e2\ldots\} \) where an edge \( e \) is the link connecting two vertices with interference weight \( w \). A larger weight implies having a larger sum of the path loss and shadowing values. The problem is equivalent to coloring each vertex with one color from \( C = \{c1, c2\ldots\} \) and assign proper power level to the respective vertex in order to maximize the throughput and mitigate interference.
The color of the vertices represents the available sub-channel, and a color pool of the interference graph relies on which particular SBs can be used by that graph. Once the graph is established, we label each FBS with the total interference weight calculated in 5.25. At this moment, the vertex with the highest interference weight is colored with an appropriate color with condition that the selected sub-channel satisfies application QoS. If there are more than one FBS with the same interference weight, we check the traffic priority. For example, if one FBS has real-time traffic or it is loaded more than others, it will have the priority to access the sub-channel. Traffic load and application data rate are used to quantify the traffic priority. An indicator $T_p$ is assigned to the FBS to refer to the priority of its traffic. The indicator is ordered

**Algorithm 4 Co-tier Channel Assignment**

**Require:** $GR = (V, E, W)$, $C = \{c_1, c_2, \ldots\}$, FBS data rate, traffic load and traffic type

**Ensure:** Sub-channel assignment with maximum throughput $T$, maximum spectral efficiency and minimum interference $I$

1: BEGIN
2: Define $V_0$ as the number of vertices in $GR$
3: Define $i$ as the index for the FBS
4: for ($a = 1$ to $a = V_0$) do
5: Calculate $W_i$, $T_p$, $\forall V \in GR$
6: if ($W_i = W_{\text{max}}$) then
7: Select FBSs with maximum $W$
8: end if
9: if FBS with $W_{\text{max}}$ is not unique then
10: find FBS with ($T_p = T_{\text{p}_{\text{max}}}$)
11: end if
12: Color the vertex $v$ with color $c$ (Assign the sub-channel to the FBS)
13: Adapt transmission power as in 5.28
14: end for
15: while (1) do
16: Check $U$ for each FBS
17: if ($U < T_{h_{\text{min}}}$) then
18: Switch the users associated the FBS to other FBSs with condition that ($U \leq T_{h_{\text{max}}}$) for the target FBSs
19: end if
20: end while
21: END
in ascending order as the traffic priority increases. The coloring process to maximize throughput can be characterized as follows,

\[ \arg\max \sum_{i \in V} \sum_{n \in N_c} z^* \log(1 + SINR_{i,n}) \]  

(subject to \( \sum_{n \in N_c} z^* \leq 1, \forall i \in V \) )

where \( N_c \) denotes the specific set of sub-channels, which are used by the vertices involved, \( SINR_{i,n} \) is the SINR of femtocell \( i \) over the sub-channel \( n \) and \( V \) is a set of vertices in the graph. Then, the power is adapted for the FBS according to \( SINR_{target} \) and the current transmission power as follows,

\[ P_i^* = P_i \frac{SINR_{target} I_2 + N(i)}{H_{i,n}} \]

where \( P_i \) is the current transmission power, \( H_{i,n} \) is the channel gain experienced by FBS \( i \) while accessing sub-channel \( n \) and \( N(i) \) is the corresponding noise power. Finally, the FBS is removed from the graph \( GR \) and the process repeated again until the set \( V \) is empty. In addition, the awareness mechanism manages to improve spectrum utilization by sharing the number of users associated with each FBS (\( U \)). If this number falls below certain threshold \( Th_{min} \), all the users associated with this FBS are switched into another FBSs in the same domain with a condition that \( U \) for these FBSs is not exceeding certain threshold \( Th_{max} \) and QoS of the switched users is guaranteed. Consequently, the sub-channel is released for other FBSs and this improves spectrum utilization and reduces contention of other FBSs.

5.2.3 Performance Evaluation

In this section, the performance of our proposed cognitive aware interference mitigation scheme is evaluated through simulation. The simulation environment parameters
are presented in Table 5.4. Note that $d$ is the distance between the user and the base station.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier frequency</td>
<td>2 GHz</td>
</tr>
<tr>
<td>Cellular layout</td>
<td>Hexagonal grid</td>
</tr>
<tr>
<td>Macro-cell radius</td>
<td>500 m</td>
</tr>
<tr>
<td>Femto-cell radius</td>
<td>10 m</td>
</tr>
<tr>
<td>Path loss MBS user</td>
<td>$L = 15.3 + 37.6 \log(d)$</td>
</tr>
<tr>
<td>Path loss FBS user</td>
<td>$L = 38.46 + 20 \log(d) + 0.7(d)$</td>
</tr>
<tr>
<td>Lognormal shadowing</td>
<td>0 mean, 8 dB standard deviation</td>
</tr>
<tr>
<td>MBS transmission power</td>
<td>45 dBm</td>
</tr>
<tr>
<td>FBS transmission power</td>
<td>20 dBm</td>
</tr>
<tr>
<td>White noise power density</td>
<td>174 (dBm Hz$^{-1}$)</td>
</tr>
<tr>
<td>Number of SBs</td>
<td>6</td>
</tr>
<tr>
<td>Number of sub-channels per SB</td>
<td>10</td>
</tr>
<tr>
<td>Macro-user per macro-cell</td>
<td>30</td>
</tr>
<tr>
<td>The penetration loss of walls</td>
<td>$L_w$ 10 dB</td>
</tr>
</tbody>
</table>

Table 5.4: Simulation environment parameters for cognitive interference mitigation in LTE

We compare our scheme with the two schemes proposed in [196] and [197] for interference mitigation in femtocell macro-cell deployment. In addition, we compare it with the standard scheme that does not implement any interference mitigation mechanism. The scheme proposed in [196] is cognitive based (CR-based) and aims to allocate resources in femtocell networks to maximize the throughput while minimizing interference to macrocell users nearby only. However, the scheme (femto-macro) proposed in [197] considers mitigating the interference among femtocells in addition to MUEs interference control. The Cumulative Distribution Function (CDF) of the MUEs’ SINR is presented in Figure 5.11. The proposed scheme achieves the highest MUEs’ SINR in contrast to the standard random resource allocation, the CR-based and the femto-macro schemes. The reason is that both the CR scheme and the femto-macro scheme employ simple energy detectors to detect the vacant sub-channel. However, our scheme considers accurate detection by better evaluating the interference range and improving the detection sensitivity. The CDF of the femto-user’s SINR is shown in Figure 5.12. The proposed scheme achieves significant performance...
improvement in comparison with all the other schemes owing to the interference coordination between FBSs with the awareness based spectrum allocation mechanism. The CR-based scheme has no capability to mitigate interference between FBSs while the femto-macro scheme is based on clustering, which limits each cluster of femtocells to access only one sub-channel. This increases the probability of collisions between the contending FBSs. On the other hand, the proposed scheme is not limited to one sub-channel and it comprises awareness and information exchange between the FBSs for channels allocation which does not only mitigate interference but also improves
spectrum utilization and maximizes throughput.

Figure 5.13 presents the spectrum efficiency achieved by all the schemes. It can be noticed that the proposed scheme recorded the highest spectrum efficiency as it considers efficient spectrum utilization by switching users from under-utilized FBSs and free more sub-channels. Figure 5.14 presents the throughput achieved by the femto-cell as a function of various number of FBSs. The transmitted traffic considered here is real-time traffic. The proposed scheme gains the highest throughput
compared to other schemes as it employs adaptive power allocation, considers users QoS requirements and traffic priority which enhances the throughput. Moreover, we notice that the throughput decreases as the number of FBSs increases. It is mainly due to the increase in the probability of collision as the number of FBSs grows.

In this chapter, we embedded CogWNet in LTE networks and optimized the user application QoS parameters. We maximized network throughput, mitigated interference, and finally deployed an intelligent frequency assignment. In the next chapter, we consider machine learning for next generation 5G heterogeneous and H-CRANs. We investigate the resource allocation in 5G Hetnets and H-CRANs with the goal to maximize energy efficiency.
Chapter 6

Machine Learning Approaches for Resource Allocation Toward Efficient Next Generation 5G Networks

The multi-tier network structure with different sizes, transmission powers, and unprecedented numbers of smart and heterogeneous wireless devices is the trends that targets maximize the network capacity in the upcoming 5G networks. It consists of two tiers: primary tier and secondary tier. The primary tier includes high power macrocells that serve MUEs, while the secondary tier comprises picocells, femtocells and D2D communications. H-CRANs is another penetrating trend of 5G that aim at exploiting the advantages of both Hetnets and CRANs and overcome their drawbacks. Comparing with CRANs and Hetnets, H-CRANs have been shown to exhibit significant performance gains though advanced collaborative signal processing and radio resource allocation are still challenging. Inter-tier interference between the macro BSs and RRHs has a severe impact on energy efficiency.

In this Chapter, we investigate the resource allocation problem in the 5G Hetnets with the goal to maximize data rate using online machine learning. As energy efficiency is an important concern in this context, we propose a novel learning technique for power allocation for the secondary tier that maximizes energy efficiency. In addition, we consider the energy efficiency problem from BS stations operation pattern perspective. Thus, a novel traffic offloading algorithm is proposed to control the operation mode of small cells and boost energy efficiency. Finally, the energy effi-
ciency problem is investigated in H-CRANs and a novel resource allocation scheme is proposed. All the machine learning techniques proposed in this chapter are enhanced in a way that reduces the computation complexity and expedites the convergence.

6.1 Cooperative Online Learning Scheme for Resource Allocation in 5G Heterogeneous Networks

This section presents a cooperative online learning scheme which aims at solving the resource allocation and interference problems in the 5G systems [23]. The cooperation considered involves information exchange between secondary tier users of the same type, which is either femtocell or D2D connection. Online learning exploits environment awareness to allocate frequency RBs and control interference. It creates allocation policies without any prior model of the environment (in our settings, a prior model cannot be achieved due to the unplanned placement of the secondary tier users and the dynamics of the wireless environment). Moreover, online learning allows the secondary tier to take actions while they are learning (i.e., without a centralized controller), which reduces the complexity of the system implementation. These features make online learning suitable to be applied at the distributed secondary tier network setting in the form of the so called multi-agent online learning. Our proposed scheme has the following contributions:

- Efficient RBs allocation using cooperative online learning in a priority based distributed fashion, where user location determines its assigned resources. This allocation proceeds in conjunction with system parameters adaptation including transmission power and modulation.

- Cross-tier interference mitigation between primary tier (macro) and secondary tier (femto and D2D links) and co-tier interference between secondary tier devices in the downlink.
• Maintain QoS of both MUEs and secondary tier users. This includes zero outage and SINR above threshold for MUEs, maximum data rate and minimum outage for the secondary tier users, and high level of fairness.

6.1.1 System Model

Our system model consists of a heterogeneous multi-tier network with secondary tier underlaid within the primary (macro) tier coverage with constraint that the interference caused to the MUEs remains below certain threshold as in Figure 6.1. All the underlay network devices including FBSs, and D2D transmitters (DUE) share the same radio resources with the macrocell. The network in Figure 6.1 is a multi-tier heterogeneous since each of the network tiers (i.e., macro tier and underlay tier which comprises femtocells and D2D UEs) has different transmission power range, coverage region and specific set of users with different application requirements. It is assumed that the user association with the base stations (either mbs or FBS) is completed prior to resource allocation. In addition, the potential DUEs are discovered during the D2D session setup by transmitting known synchronization or reference signal (i.e., beacons) [198]. Only one MUE is assumed to be served on each RB to avoid co-tier interference within the macro tier. However, multiple underlay devices compete to reuse the same RB to improve the spectrum utilization. This reuse causes severe cross-tier interference to the MUEs, and also co-tier interference within the underlay tier. The main objective is to allocate resources to the underlay transmitters (FBS or DUE) with the goal of maximizing their data rate while maintaining the SINR of the affected MUE.

Each transmitter in the underlay tier selects one RB from a set of \( N \) RBs for transmission with certain transmission power and modulation order. The underlay transmitters are capable of selecting the transmission power and modulation level from a finite set of power levels, \( p = \{1, 2, 3, ..., P\} \) and finite set of modulation indexes.
Figure 6.1: System Model of 5G Hetnets

\( m = \{1, 2, 3, \ldots, M\} \), respectively. As we have two types of underlay transmissions (i.e., femtocell and D2D), each type selects transmission power and modulation from its corresponding finite set. Each transmitter selects a suitable \( \{n, p, m\} \) level combination. We call this combination transmission package (\( Tp \)). The transmitting FBS or D2D transmitter is denoted by \( y \). Its transmission power over the RB is determined by the vector \( p_y = \{p_y^1, p_y^2, p_y^3, \ldots, p_y^N\} \) where \( p_y^n > 0 \) denotes the transmission power level of the transmitter \( y \) over RB \( n \). The same thing is applicable to the modulation levels. Note that if the RB \( n \) is not allocated to the transmitter \( y \), the corresponding power variable will be \( p_y^{(n)} = 0 \). Since we assume that each underlay transmitter selects only one RB, only one element in the power vector \( p_y \) is non-zero. The frequency band \( b \) is assumed to be divided into two disjoint sub-bands: edge sub-band \( e \), where users experienced the weakest signal and the center band \( c \) with stronger signal as highlighted in Figure 6.1. These two bands differ in the assigned transmission power for the underlay transmissions located within their coverage: users located in \( c \) communicate with less power than the ones located in \( e \) where transmissions occur at the
higher power to compensate for the poor signal experienced. Therefore, we adopt a priority based resource allocation in which UEs with the weakest signal quality located at $e$ benefit from the maximum transmission power of their transmitters. The signal quality is quantified based on the measured SINR and thus, all the underlay users are sorted according to their received SINR.

6.1.2 Problem Formulation

The resource allocation problem in the 5G context with multiple tiers has the objective of maximizing the data rate of the underlay users with minimal interference to the macro tier. This objective relies on certain factors including SINR of underlay UEs and SINR of MUE. In order to calculate the SINR at an underlay receiver $u$ whether it is an (femto UE) or D2D receiver, we need to define the sources of interference that impact its signal assuming that the interfering underlay transmitters $y'$ share the same RB used by $y$. The interference perceived by $u$ is found as follows,

$$I_u = p_k G(k, u) + \sum_{y' \in Y, y' \neq y} \lambda_{y', b} p_{y', b} G(y', u)$$

(6.1)

where $p_k$ is the transmission power of the MBS $k$, $G(k, u)$ is the interference gain of the MBS in the direction of the receiver $u$, $p_{y', b}$ is the transmission power of the interfering underlay transmitters, $\lambda_{y', b}$ represents the load of sub-band $b$ and is defined as the ratio of the number of RBs allocated in sub-band $b$, and the total number of RBs available in that sub-band, and $G(y', u)$ is the gain of underlay transmitter $y$ in the direction of $u$. The gain $G$ of certain transmitter $k$ or $y'$ in the direction of $u$ over RB $n$ is defined as follows,

$$G(y', u) = \beta_{y' u} d_{y' u}^{-\alpha^*}$$

(6.2)

where $\beta_{y' u}$ denotes the channel fading component between $y'$ and $u$ over RB $n$, $d_{y' u}^{-\alpha^*}$ is the distance between $y'$ and $u$, and $\alpha^*$ is the path-loss exponent. The SINR for $u$
over RB $n$ in sub-band $b$ is given by

$$SINR_u = \frac{p_{y,b}G(y,u)}{I_u + \sigma^2}$$

(6.3)

where $p_{y,b}$ is the transmission power of the transmitter $y$ and the $G(y,u)$ is the gain between the underlay transmitter $y$ and the receiver $u$. The variable $\sigma^2 = N_0B_{RB}$ where $B_{RB}$ is the bandwidth corresponding to an RB and $N_0$ is the thermal noise.

Similarly, the SINR for the MUE $z$ over RB $n$ can be written as follows:

$$SINR_z = \frac{p_kG(k,z)}{\sum_{y \in Y} p_{y,b}G(y,z) + \sigma^2}$$

(6.4)

Given the SINR, the data rate of the UE $u$ over RB $n$ in SB $b$ can be calculated according to the Shannon’s formula, i.e., $R_u = B_{RB}(1 + SINR_u)$. The assignment of transmission package i.e. $Tp = \{n, p, m\}$ for any underlay transmitter $y$ to maximize the data rate is denoted by a binary decision variable $x_{y}^{(n,p,m)}$ where

$$x_{y}^{(n,p,m)} = \begin{cases} 1, & \text{if } s \text{ is transmitting over } n \text{ with } p \text{ and } m \\ 0, & \text{Otherwise} \end{cases}$$

(6.5)

Thus, the achievable data rate of a secondary UE $u$ with the corresponding transmitter $y$ is written as,

$$R_{uy} = \sum_{n=1}^{N} \sum_{p=1}^{P} \sum_{m=1}^{M} x_{y}^{(n,p,m)} B_{RB}(1 + SINR_u)$$

(6.6)

In order to maintain the SINR of the MUEs, the interference caused by the underlay transmissions $I_z$ over RB $n$ must be maintained below the threshold of the maximum tolerable interference by the MUE $I_{TH}$ as follows,

$$I_z = \sum_{y \in Y} \sum_{p=1}^{P} \sum_{m=1}^{M} x_{y}^{(n,p,m)}G(y,z)p_{y,b} \leq I_{TH}$$

(6.7)
The resource allocation problem can be expressed by using the following optimization formulation,

\[
\max_{x_{y,(n,p,m)}, p_y, b, m_y, b_y \in Y} \sum_{n=1}^{N} \sum_{p=1}^{P} \sum_{m=1}^{M} x_{y,(n,p,m)} B_{RB}(1 + SINR_{uy}) \quad (6.8)
\]

subjected to,

\[C1 \quad I_z \leq I_{TH} \forall n \in N\]

\[C2 \quad \sum_{n=1}^{N} \sum_{p=1}^{P} \sum_{m=1}^{M} x_{y,(n,p,m)} \leq 1 \forall y \in FBS \sqcup DUE\]

\[C3 \quad x_{y,(n,p,m)} \in \{0, 1\} \forall n \in N, \forall y, \forall p \in P, \forall m \in M\]

where,

\[SINR_u = \frac{p_{y,b} G(y, u)}{p_k G(k, u) + \sum_{y' \in Y, y' \neq y} \sum_{p'=1}^{P} \sum_{m'=1}^{M} x_{y,(n,p,m)} \lambda_{y', b} p_{y', b} G(y', u) + \sigma^2} \quad (6.9)\]

The allocation problem presented in (6.8) aims at maximizing the data rate of the femto UE or DUE while fulfilling the constrains C1, C2, and C3. In the constraint C1, the interference caused to the MUEs by the underlay transmitters on certain RB is limited by a predefined threshold. The constraint in C2 indicates that the number of RBs selected by each underlay transmitter should be at most one and each transmitter can only select one power level and one modulation level at each RB. The binary assignment variable for \(T_p\) selection is represented by the constraint in C3. The resource allocation problem is a combination non-convex non-linear optimization problems. Considering the computational overhead, it not feasible to solve the resource allocation problem by a centralized approach.
6.1.3 Online Learning System Parameters

We consider a cooperative multi-agent online learning system, where each agent is either an FBS or DUE transmitter. Only agents of the same type (i.e., FBS or D2D transmitter) coordinate with each other. The agent collects the environment information or performance indicators from its own and its neighboring agents to define the system state $s_t$ at time $t$, and perform a local action $a_t$. Agents enforce the cooperative learning using a global reward, which comprises the sum of rewards achieved by all the neighboring agents. The state action table (Q-table) is common and shared by the underlay transmitters. Thus, the agents learn together a common strategy by feeding a single Q-table. In addition to fast convergence time and fairness, the system benefits from a diversified experience learned by the cooperating agents.

The fundamental parameters of the online learning system are defined as follows:

- **State**: at time instant $t$, the environment state is defined as,

$$s_t = \{y, n, p, SINR_{cu}, SINR_{eu}, SINR_z\}$$

where $y$ is the underlay tier transmitter, $n$ is the available RB, $p$ is the transmission power of the underlay tier transmitter, $SINR_{cu}$ is the SINR measured at the underlay receiver $u$ in the center band, $SINR_{eu}$ is the SINR measured at the underlay receiver in the edge band, and $SINR_z$ is the SINR measured at the macro receiver. The aggregated state information of the neighboring agents $(s(y'_{\text{all}}))$, is defined as a weighted sum over the performance indicators (of the same type) $(s(y'))$ and it is given by:

$$s(y'_{\text{all}}) = \sum_{y' \in Y} w_{y'} s(y') \quad \text{(6.10)}$$
where $w_{y'y}$ is the weighing coefficient that reflects the degree of neighborhood of agent $y'$ to $y$. The weight represents the normalized traffic flux between agents $y$ and $y'$ with respect to the total traffic flux between $y$ and all its neighbors.

- **Action**: the action at time instant $t$ is defined as $a_t = \{\text{allocation of } n, p_c, p_e, m\}$ where $p_c$ and $p_e$ are the selected transmission power of the underlay transmitters $y$ located at the center band and edge band respectively. $m$ is the transmission modulation level.

- **Reward**: the reward achieved is represented as $r_t = \{R_{uy}, R_{zk}, SE_{uy}\}$ where $R_{zk} = Blog_2(1 + SINR_z)$ is the data rate achieved by the MUE and $SE_{uy}$ is the spectral efficiency achieved at the underlay receiver $u$, which is found by mapping it to $SINR_u$ using quality tables (obtained using a link-level simulator) incorporated within the network simulator. The reward function is evaluated with rationale of maximizing the underlay tier users’ data rate while maintaining the macrocell users’ data rate above certain threshold $R_{zk}'$ as follows,

$$r_t^y = \begin{cases} 
  e^{-(R_{zk}-R_{zk}')^2} - e^{-R_{uy}}, & \text{subjected to } C.1, C.2, C.3 \\
  -2, & \text{otherwise}
\end{cases}$$

As the considered learning scheme is cooperative, the instantaneous global reward is a sum of all the individual rewards of the cooperating agents of the same type and is given by

$$r_y = \sum_{y \in Y} r_t^y$$

### 6.1.4 Resource Allocation Mechanism

The online learning mechanism exploits the collected state information and uses it to allocate resources and control interference. As the exploited learning approach is cooperative, it necessitates sending state-action information from the corresponding
agent Q-table to all the neighbors and receive their state-action information. This information supports the exploitation phase in which the action is selected according to the highest Q-value, which is recursively updated as follows,

$$Q_y(s_y, a_y) = (1 - \alpha)Q_y(s_y, a_y) + \alpha(r_y(s_y, a_y) + \gamma \max_{l \in A_y} Q(s_{y*}, l))$$

where $s_y$ is the current state of the agent $y$ and $s_{y*}$ is the previous state of the agent $y$. However, the learning mechanism relies on the performance metric (i.e., SINR, data rate) measured to take actions in the exploration phase besides their usage to evaluate the reward function. There are three main characteristics of the proposed learning mechanism: First, the environment state information are assumed to be available upon request during learning, which is the task of the employed environment awareness methodology. Second, the mechanism starts with high exploration rate $\epsilon$ and this rate is decreased gradually to ensure that there are enough actions with high Q-value to exploit. Finally, the learning rate $\alpha$ in (6.13) is assumed to be dynamic and follow the Win-or-Learn-Fast principle which states that the learning agent should learn faster when it is losing and more slowly when winning [54]. The learning rates that we used are $\alpha = 0.05$ for rewarded solution and $\alpha = 0.2$ for the punished one. The learning mechanism based on the proposed model is illustrated in Figure 6.2.

The process starts by collecting the network state information. Then, the state information exchange process engages where each agent $y$ shares the row of its Q-table that corresponds to its current state and the optimal action where $Q_y(s_y, a_y) \geq Q_y(s_y, a'_y)$ for all $a'_y \in A$ with all other cooperating agents $y'$ (i.e., underlay tier in the same range). At the same time, it receives the current state and all optimal actions $Q_{y'}(s_{y'}, \cdot)$ from other agents $y'$. This helps the agent $y$ to determine its joint action
Start

Collect current state information

Share state with the neighboring agents

Send

Receive

Check $\epsilon$

Exploration or Exploitation?

Exploration

Select action randomly,
1. Check SINR at underlay receiver to select $p$ and $m$ for the allocated RB.
2. Check SINR at the macro receiver to guarantee MUE QoS.
3. Check $C_1$, $C_2$, $C_3$ for fulfillment

Evaluate reward function for the selected action at $=\{n, p, m\}$ according to equation (14)

Update the Q-value of the action in the Q-table as state-action pair according to equation (15)

Observe the new state $s_t = s_{t+1}$

Exploitation

Select action with the highest Q-value according to Equation (16)

Figure 6.2: Online resource allocation mechanism in 5G Hetnets with the highest cumulative Q-value in the exploitation phase as follows,

$$a_y = \arg\max \left( \sum_{y \in Y} Q(s_y, a_y) \right)$$ (6.14)

The Q-value is found and updated according to the rule in (6.13). The global Q-value found in (6.14) is decomposed into a linear combination of local agent-dependent Q-values: $Q(s, a) = \sum_{y=1}^{Y} Q_y(s, a)$. Thus, if each agent $y$ maximized its own Q-value, the global Q-value will be maximized. Based on this observation, choosing the action based on (6.14) will maximize the global Q-value. After the information exchange, the mechanism decides to employ exploration or exploitation process according to the value of $\epsilon$. In the exploration process, the action to assign RB and the selection of power and modulation is taken randomly at the first trial. Then, the measurement of SINR and the location of the underlay receivers are exploited to improve resource allocation. For instance, if the selected transmission power to communicate with user at $c$ is $p_c$, the allocated power to communicate with a receiver at $e$ will be
\( p_c = \Gamma p_c \) where \( \Gamma \) is decided based on the maximum allowed transmission power. Modulation is selected to be high for high SINR users and vice versa. In addition, the conditions C1, C2, and C3 are checked for satisfaction. On the other hand, the transmission package with highest cumulative Q-value is selected according to (6.14) in the exploitation process. For example, if there are two agents \( i \) and \( j \), each agent has one state \( s \) and three actions \( a_1, a_2 \) and \( a_3 \), the reward for each agent is its capacity and the Q-values for both agents are as follows: \( Q_i(s,a_1) = 1, Q_i(s,a_2) = 2, Q_i(s,a_3) = 3, Q_j(s,a_1) = 2.5, Q_j(s,a_2) = 7 \) and \( Q_j(s,a_3) = 4.5 \). Both agents will choose action \( a_2 \) (the maximum of the summation of the Q-values is \( 2 + 7 \)), thus maximizing the aggregate capacity. After that, the reward function is evaluated for the selected \( T_p \) and the new state is observed. Finally, the Q-table is updated with the new Q-value as in (6.13) according to the state action paired selected. The overhead of information exchange in this cooperative scheme is calculated based on the size of \( T_p \) and the number of cooperative transmitters. So if the number of transmitters is \( Y^* \), then the overhead is \( Y^*(Y^* - 1) \) messages. Each message has the size of \( T_p \).

### 6.1.5 Performance Evaluation

We conducted extensive simulations to demonstrate the capability of using online learning scheme for resource allocation. The considered simulation environment consists of a multi-tier network where multiple femtocells and D2D communications share the spectrum resources with one macrocell in an underlay fashion. Note that the number of D2D transmitters contribute to 20% of the number of the underlay transmitters while the rest are FBSs in all the simulations. The MUEs are uniformly distributed in their respective cell coverage area. Each FBS is assumed to serve one UE which is randomly located in the coverage area. The simulation parameters and the online mechanism related parameters are listed in Table 6.3.

We compare the performance of our online learning resource allocation scheme
<table>
<thead>
<tr>
<th>Simulation Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Number of RBs</td>
<td>10</td>
</tr>
<tr>
<td>number of MUEs</td>
<td>10</td>
</tr>
<tr>
<td>FBS Tx power</td>
<td>12 to 22 dBm</td>
</tr>
<tr>
<td>MBS Tx power</td>
<td>48 dBm</td>
</tr>
<tr>
<td>DUE TX power</td>
<td>-10 to -2 dBm</td>
</tr>
<tr>
<td>Modulation QPSK and coding rate</td>
<td>16-QAM: 1/2, 2/3, 3/4 ; 64-QAM: 1/2, 2/3, 3/4, 4/5</td>
</tr>
<tr>
<td>Path loss</td>
<td>$PL = 127 + 37 \log(d)$, $d$ = distance between underlay transmitter and UE</td>
</tr>
<tr>
<td>Shadowing</td>
<td>log-normal distribution (mean = 0dB, standard deviation = 8dB)</td>
</tr>
<tr>
<td>Thermal noise density</td>
<td>-174 dBm/Hz</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Learning Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Rate ($\alpha$)</td>
<td>dynamic</td>
</tr>
<tr>
<td>Exploration Rate ($\epsilon$)</td>
<td>dynamic</td>
</tr>
<tr>
<td>Discount Factor ($\gamma$)</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Table 6.1: 5G Hetnets environment and online scheme simulation parameters to validate its performance with reference schemes including: Down-SA [199], Joint-RALA [200], and Matching-RM schemes [201]. Down-SA proposes an architecture that consists of a Decision Support System (DSS) and a data collection system to dynamically manage and control the spectrum allocation process in the heterogeneous 5G networks. Joint-RALA proposes a joint algorithm for resource allocation and link adaptation to support carrier aggregation functionality for downlink in LTE-A network. Matching-RM employs matching theory to find a stable match for the resource allocation problem to maximize the throughput of small cells under the cross-tier interference constraint. We evaluate the performance of our scheme compared to others in terms of the underlay average system throughput, spectral efficiency in bits/Hz and fairness of resource allocation among the underlay transmitters compete for the available RBs as in Figure 6.3, Figure 6.4, and Figure 6.5 respectively. Figure 6.3 and Figure 6.4 show the aggregate system throughput with respect to the number of underlay transmitters involved and the spectral efficiency achieved by our allocation scheme in comparison to other schemes respectively. It is clear that our al-
location scheme achieved the highest throughput and spectral efficiency, compared to other schemes. The reason is that our scheme formalizes the allocation problem with throughput maximization objective and it exploits online learning as a tool to solve it. Online learning is convenient to make decisions to perform allocation in such heterogeneous environment that needs plenty of interactions. The fairness performance in terms of Jain’s fairness index $f(x_1, x_2, \ldots) = \frac{\left(\sum_{i=1}^{J} x_i\right)^2}{J \sum_{i=1}^{J} x_i^2}$ as in [202] versus the number of underlay UEs for each scheme is plotted in Figure 6.5. The value of the fairness
Figure 6.5: Jain’s fairness index with different number of active transmissions in 5G Hetnets

index lies in the range between [0, 1], and the value of 1 represents that all users have the same average data rate. We notice that the online allocation scheme achieved the maximum fairness. Only Joint-RALA shows comparable fairness to our scheme. The reason is that Joint-RALA adopts proportional fairness based scheduling.

The average SINR achieved by the MUEs is plotted in Figure 6.6. This measurement aims to evaluate the scheme capability to account for the macro tier QoS. We can notice that our scheme achieved the highest SINR as this is a fundamental constraint in the allocation problem. Matching-RM is the only other scheme account for this constraints. Therefore, it achieved a comparable SINR. Figure 6.7 and Figure 6.8 present the outage ratio versus the number of underlay transmitters for the underlay users and the MUEs respectively. The outage ratio of a particular tier can be expressed as the ratio of the number of UEs supported by a tier with their minimum target SINR and the total number of UEs in that tier. Our online scheme recorded almost zero outage ratio for the underlay tier and minimum outage for the macro tier in comparison with other schemes.

Nevertheless, that Down-SA, Joint-RALA, and Matching-RM are proposed for resource allocation in the downlink of 5G, their performance is limited by several
Figure 6.6: Mean SINR at MUE receiver in 5G Hetnets

Figure 6.7: Average Outage ratio at the underlay tier in 5G Hetnets

drawbacks. For instance, Down-RA does not consider QoS in RB allocation and interference problem is out of its scope although it is the main problem in resource allocation. Joint-RALA proposed an algorithm to allocate RBs while maximizing the achieved data rate. However, power control and macro tier QoS maintenance issue are not taken in consideration. Matching-RM came with the goal to handle both allocation with QoS and interference problems but it follows centralized approach for resource management which is not feasible with all the computational overhead in such network. In addition, none of the proposed schemes consider the impact of the
6.2 Energy Efficient Power Allocation in Multi-tier 5G Networks Using Enhanced Online Learning

In this section, we consider the power allocation problem from the energy efficiency perspective in the multi-tier 5G heterogeneous network. Therefore, we proposed a distributive intuitive online learning power allocation scheme for the STs, which allows each ST to conjecture other STs power allocation strategies with only local information from direct interactions with the environment and making use of the past experience [24]. This accounts for the lack of information exchange among STs, which is essential to reach optimal power allocation. In addition, the proposed scheme maintains QoS represented by signal to interference plus noise ratio (SINR) for all users in the network. The proposed scheme provides the following key highlights: First, the intuition feature allows each ST to update its learning information using its private past experience and this eliminates the overhead of cooperation. Second, the chapter contributes to the general literature of online learning [203] as the traditional online learning relies on full information from all agents in the environment,
which is difficult to achieve in such dynamic environment. Third, the proposed online scheme exploits a brief representation feature that significantly reduces the scheme computation. The brief feature approximates the Q-value of the online learning as a function of much smaller set of variables, which reduces the state space and expedites the algorithm convergence.

6.2.1 System Model

Our system considers the downlink transmission in a spectrum sharing heterogeneous 5G with two tiers: primary tier and secondary tier, where the primary tier consists of the macrocell with its associated MUEs and the secondary tier consists of two types of cells including picocells and femtocells noted as secondary tier cells and D2D communications as in Figure 6.9. All the secondary tier BSs and D2D users are assumed to be uniformly distributed under the coverage of the macrocell. The set of secondary BSs is donated by $N = \{1, 2, ..., N\}$, the set of secondary BSs associated users (SUEs) are denoted by $X = \{1, 2, ..., X\}$, the set of MUEs is represented by $K = \{1, 2, ..., K\}$, and the set of the active D2D pairs as $D = \{1, 2, ..., D\}$. The
dth D2D pair \((d \in D)\) consists of the D2D transmitter \(d_T \in D_T\) and D2D receiver \(d_R \in D_R\), where \(D_T = \{1, 2, \ldots, D_T\}\) and \(D_R = \{1, 2, \ldots, D_R\}\). The set of SUEs associated with the \(n\)th secondary BS is referred as \(X_n\) with assumption that each SUE can associate with at most one secondary BS. Their resource allocation problem includes SUE association and power allocation. For SUE association, each SUE can be associated with a single BS. To specify the SUE association, we define \(\nu_n^x\) as the association indicator for SUE \(x\) with BS \(n\), which is a binary variable. If \(\nu_n^x = 1\), it indicates that \(x\)th SUE \((x \in X)\) is associated with \(n\)th cell and it is zero otherwise.

For transmission power allocation, the \(n\)th secondary BS can select a random power level \(P_L\) from a set of discrete power levels as follows,

\[
P_L = \begin{cases} 
\in [1, P_L^*], & \text{if the } n\text{th BS serves one SUE} \\
0 & \text{otherwise}
\end{cases}
\]

(6.15)

where \(P_L^*\) is the maximum integral level of the secondary BS transmission power. To illustrate, the power allocated to \(n\)th secondary BS \(P_n\) belongs to the set \([0, \frac{1}{P_L}P_{n}^{\text{max}}, \frac{2}{P_L}P_{n}^{\text{max}}, \ldots, \frac{P_d}{P_L}P_{n}^{\text{max}}, \ldots, P_{n}^{\text{max}}]\), where \(P_{n}^{\text{max}}\) is the maximum allowed transmission power for the \(n\)th secondary BS. On the other hand, the \(d\)th D2D pair can select a power level \(P_d \in \{0, 1, 2, \ldots, P_L^*\}\) that satisfies minimum transmission power requirements \(P_d^{\text{min}}\) and ensures that the D2D receiver is located within the proximity of the D2D transmitter \(R_d \leq R_d^{\text{max}}\). The channel inversion power control is exploited to compensate the large scale fading and make the average received power at the D2D receiver larger than the minimum sensitivity \(\eta_{\text{min}}\) [204]. Thus, the D2D proximity is calculated as,

\[
R_d = \left(\frac{P_d P_d^{\text{max}}}{P_L^* \eta_{\text{min}}}\right)^{\alpha}
\]

(6.16)
As a result, the minimum power level allocated to the $d$th D2D pair $P_d$ is found as

$$P_{d}^{\text{min}} = \frac{P^{*} \eta_{\text{min}} P^{\alpha}}{P_{d}^{\text{max}} P_{d}}$$ (6.17)

where $P_{d}^{\text{max}}$ is the maximum allowed transmission power for the D2D transmitter.

The SINR of the $x$th SUE associated with the $n$th secondary BS is calculated as follows,

$$\gamma_{n,x} = \frac{P_{n} G_{n,x}}{I_{d,x} + I_{q,x} + I_{M,x} + \sigma}$$ (6.18)

where $G_{n,x}$ is the power gain between the $n$th secondary BS and the $x$th SUE, and $\sigma$ is the noise power. The aggregate interference at the $x$th SUE from all D2D transmitters is defined as,

$$I_{d,x} = \sum_{d \in D_T} P_{d} G_{d,x}$$ (6.19)

Moreover, the interference at the $x$th SUE from all other secondary BSs $q$ is defined as,

$$I_{q,x} = \sum_{q \in N / n} P_{q} G_{q,x}$$ (6.20)

where $G_{d,x}$ and $G_{q,x}$ are the power gains between the $x$th SUE and both D2D transmitters and the other secondary BSs $q$ respectively. $P_{q}$ is the transmission power of the secondary BS $q$. The interference from the macro BS to the $x$th SUE is defined as follows,

$$I_{M,x} = P_{m} G_{m,x}$$ (6.21)

where $P_{m}$ is the transmission power of the macro BS and $G_{m,x}$ is the power gain between the macro BS and the $x$th SUE. Note that the macro BS interference comes from the macro BS that the SUE operate under its coverage.
The SINR of the $d$th D2D pair is calculated as follows,

$$\gamma_d = \frac{P_d G_{dT,dR}}{I_{y,d} + I_{N,d} + I_{M,d} + \sigma}$$  \hspace{1cm} (6.22)$$

where $G_{dT,dR}$ is the power gain between the $d$th D2D transmitter and the $d$th D2D receiver and $\sigma$ is the noise power. The aggregate interference at the $d_R$th D2D from all other D2D transmitters is defined as,

$$I_{y,d} = \sum_{y \in D_T/d_T} P_y G_{y,dR}$$  \hspace{1cm} (6.23)$$

Moreover, the interference at the $d_R$th D2D from all the secondary BSs that belongs to $N$ is defined as,

$$I_{N,d} = \sum_{n \in N} P_n G_{n,dR}$$  \hspace{1cm} (6.24)$$

where $G_{y,dR}$ and $G_{n,dR}$ are the power gains between the $d_R$th D2D and both D2D transmitters $y$ and other secondary BSs $n$ respectively. $P_y$ is the transmission power of the D2D transmitter $y$. The interference from the macro BS to the $d$th receiver is found as follows,

$$I_{M,d} = P_m G_{m,dR}$$  \hspace{1cm} (6.25)$$

where $G_{m,dR}$ is the power gain between the macro BS and the $d$th D2D receiver.

**6.2.2 Problem Formulation**

To realize the non-cooperative energy efficient power allocation, we define the energy efficiency in the power allocation process as the ratio of the data rate to the power consumed by both secondary BSs and D2D transmitters as follows,

$$EE_i = \frac{B \log_2(1 + \gamma)}{P_i + P_{cc}}$$  \hspace{1cm} (6.26)$$
where $B$ is the bandwidth and $P_{cc}$ is the power consumed by the ST circuit. Note that the index $i$ refers to the ST including both secondary BSs and D2D transmitters. $\gamma$ is the SINR achieved at the secondary receiver whether it is $\gamma_{n,x}$ or $\gamma_d$. Formally, The non-cooperative power allocation optimization problem in the 5G heterogeneous structure is defined as follows,

$$\max_{P_i} EE_i$$  \hfill (6.27)$$

subjected to

C.1: $\gamma_{n,x} \geq \gamma^*_{n,x} \ \forall x \in X$ and $\gamma_d \geq \gamma^*_d \ \forall d \in D_R$

C.2: $\gamma_k \geq \gamma^*_k \ \forall k \in K$

C.3: $\sum_{n \in N} v_n^x = 1 \ \forall x \in X$

C.4: $\sum_{x \in X} v_n^x \leq \kappa \ \forall n \in N$

The constraint C.1 is to guarantee that the SINR of the $x$th SUE and $d$th D2D do not fall below the thresholds $\gamma^*_{n,x}$ and $\gamma^*_d$ respectively. C.2 is the constraint to maintain the SINR of the MUEs above a designated threshold $\gamma^*_k$ provided that the SINR of the macro tier MUE is defined as follows,

$$\gamma_k = \frac{P_m G_{m,k}}{\sum_{n \in N} P_{n} G_{n,k} + \sum_{d \in D_T} P_{d} G_{d,T,k} + \sigma}$$  \hfill (6.28)$$

where $G_{m,k}$, $G_{d,T,k}$ and $G_{n,k}$ are the power gains between the $k$th MUE and macro BS, D2D transmitters $d_T$ and secondary BSs $n$ respectively. Note that C.2 is supported by the assumption that macro BS can exchange its associated MUEs SINR information with both secondary BSs and D2D transmitters as they operate under its coverage. The constraint C.3 indicates that each SUE can be associated with only one secondary BS and C.4 emphasizes that each secondary BS can serve at most $\kappa$ SUEs.
6.2.3 Power Allocation Learning Model

In this section, we establish the ST power allocation model using online learning defined as \( \zeta = (N, D_T, P, EE_i) \). The action space available for all STs is defined as \( P = \prod_{i \in N \cup D_T} P_i \). We consider a slotted time structure for spectrum access for macro and secondary tier during the long time learning process. The continuous action profile \( P_i = [P_{i\text{min}}, P_{i\text{max}}] \) is discretized to be compatible with the online learning framework according to 6.15 in the system model. We designate \( a_i \in A_i = \{0, \frac{1}{P_L} P_{i\text{max}}, \frac{2}{P_L} P_{i\text{max}}, \ldots, \frac{L}{P_L} P_{i\text{max}}, ..., P_{i\text{max}} \} \) as the STs’ action and \( A = \prod_{i \in N \cup D_T} A_i \) as the action space for all STs. Therefore, \( \zeta \) is converted to the discrete form \( \zeta' = (N, D_T, \{A_i\}, \{EE_i\}) \). In the following subsection, we define the main components of the online learning mechanism.

Online Learning Structure

Each ST has the role of a learning agent, which aims to reach optimal power allocation strategy for different network state. The online learning parameters are defined as follows:

- **State**: since there is no cooperation among the competing STs, they only rely on the local observation to define their environment state at certain time slot \( t \). The state observed by ST \( i \) is defined as,

\[
s^t_i = (i, P_i)
\]  

(6.29)

- **Action**: the action is defined as the ST transmission power \( (a_i) = (P_i) \).
• Reward: the reward function is defined for the state/action pair as $R_i(s_i, a_i)$ and is evaluated as follows,

$$R_i(s_i, a_i) = \begin{cases} R_i(a_i) = EE_i, & \text{if C.1 to C.4 are satisfied} \\ 0, & \text{otherwise} \end{cases}$$

Specifically, it is a return of selecting power level action ($a_i$) in state $s_i$ to guarantee the transmission QoS requirement as well as to achieve the desired energy efficiency. This indicates that the reward is achieved if the conditions in C.1 to C.4 are satisfied.

• Transition Function: the transition from state $s_i^t$ to $s_i^{t+1}$ is determined by the ST stochastic behavior. Each ST selects the strategy $\pi_i(s_i)$ independently to maximize its total expected reward. The strategy $\pi_i(s_i)$ is defined to be a probability vector $\pi_i(s_i) = [\pi_i(s_i, 0), ..., \pi_i(s_i, P_{i}^{\text{max}})]$ where $\pi_i(s_i, a_i)$ represents the probability at which the ST $i$ selects action $a_i$ at the state $s_i$.

For the case of having complete information about all other STs strategies $\pi_{-i} = (\pi_1, ..., \pi_{i-1}, \pi_{i+1}, ..., \pi_N)$, the total expected discounted reward of ST $i$ over an infinite time slots is defined as follows,

$$V_i(s_i, \pi_i, \pi_{-i}) = E \left[ \sum_{t=0}^{\infty} \beta^t R_i(s_i^t, \pi_i(s_i^t), \pi_{-i}(s_i^t)), s_i^0 = s_i \right]$$

$$= E[R_i(s_i, \pi_i(s_i), \pi_{-i}(s_i))] + \beta \sum_{s_i' \in S_i} T_{s_i, s_i'}(\pi_i(s_i), \pi_{-i}(s_i))V_i(s_i', \pi_i, \pi_{-i})$$

(6.31)

where $T_{s_i, s_i'}(.)$ is the state transition probability, and

$$E[R_i(s_i, \pi_i(s_i), \pi_{-i}(s_i))] =$$
where $a_{-i}$ represents the action selected by other STs for state $s_i$. In the stochastic learning, each ST has the task to learn the optimal power allocation strategy $\pi^*_i$ for each environment state $s_i$. The following condition must be satisfied in order to reach the optimal strategy $\pi^*_i$ for each ST $i \in \mathbb{N} \cup \mathbb{D}_T$,

$$V_i(s_i, \pi^*_i, \pi^*_{-i}) \geq V_i(s_i, \pi_i, \pi^*_{-i}), \forall \pi_i \in \Pi_i$$  \hspace{1cm} (6.33)

The optimal strategy satisfies the Bellman’s optimality equation [205], that is, for ST $i$

$$V_i(s_i, \pi^*_i, \pi^*_{-i}) = \max_{a_i \in A_i} \{E[R_i(s_i, a_i, \pi^*_{-i}(s_i))]
\ + \beta \sum_{s'_i \in S_i} T_{s_i,s'_i}(a_i, \pi^*_{-i}(s_i))V_i(s'_i, \pi^*_i, \pi^*_{-i})\}$$  \hspace{1cm} (6.34)

where

$$E[R_i(s_i, a_i, \pi^*_{-i}(s_i))] = \sum_{a_{-i} \in A} [R_i(s_i, a_i, a_{-i}) \prod_{j \in N/\{i\}} \pi^*_j(s_j, a_j)]$$

Thus, we can evaluate the optimal Q-value of ST $i$ as the current expected reward plus a future discounted reward when all other STs follow the optimal strategy as follows,

$$Q^*_i(s_i, a_i) = E[R_i(s_i, a_i, \pi^*_{-i}(s_i))]$$
\[+ \beta \sum_{s'_i \in S_i} T_{s_i,s'_i}(a_i, \pi^*_{-i}(s_i))V_i(s'_i, \pi^*_i, \pi^*_{-i})\]  \hspace{1cm} (6.35)

By combining (6.34) and (6.35), we get,

$$Q^*_i(s_i, a_i) = E[R_i(s_i, a_i, \pi^*_{-i}(s_i))]$$
The employed online learning scheme aims to reach the optimal Q-value defined in (6.36) in a recursive way using the information \((a_i, s_i, s_i', \pi_t^i)\) with the two states \(s_i = s_t^i\) and \(s_i' = s_{i+1}^i\) observed at the time slot \(t\) and \(t + 1\) respectively, \(a_i\) and \(\pi_t^i\) are the ST action taken at the end of time slot \(t\) and the power allocation strategy during time slot \(t\) respectively. The update rule for the online learning employed to reach the optimal Q-value is given by,

\[
Q_{t+1}^i(s_i, a_i) = (1 - \alpha^t)Q_t^i(s_i, a_i) + \alpha^t\left\{ \sum_{a_{-i} \in A_{-i}} [R_i(s_i, a_i, a_{-i}) \times \prod_{j \in N_i} \pi_j^t(s_j, a_j)] + \beta \max_{b_i \in A_i} Q_t^i(s_i', b_i) \right\}
\]

(6.37)

where \(\alpha \in [0, 1)\) is the learning rate. Once the transmission power level action \((a_i)\) is selected and the ST \(i\) achieved the expected reward, the corresponding Q-value is updated by combining the old value and the new expected reward.

System Design Challenging Issues

As the goal of this chapter is to create a stochastic energy efficient power allocation scheme that is non-cooperative and can guarantees QoS for different tier users, we can notice in the online learning structure presented requires information about other STs strategies and the reward of each ST is a function of joint actions of other STs. This creates a challenging problem due to the following reasons:

- Every ST may not be aware of the number of other ST existing in the system.
- Each ST can only obtain its local information such as environment state, its transmission strategy and received historical reward.
- The 5G heterogeneous system has a large space. Therefore, the curse of di-
dimensionality increases the required computations and makes it unfeasible to use the typical online learning methodology to maintain the Q-value for each state/action pair, which slows the system convergence.

According to the update rule derived in (6.37), we can deduce that the stochastic power allocation problem cannot be solved directly because STs cannot observe other STs strategies in the non-cooperative power allocation fashion.

### 6.2.4 Power Allocation with Approximated-Intuition Based Online Learning

We account for the problem of power allocation in 5G Hetnets without being aware of other STs power allocation strategies. In addition, the slow convergence problem due to the large space of state/action Q-values in such environment is considered. Thus, we propose a novel approximated-intuition based online learning scheme, which allows each ST to surmise other STs power allocation strategies without explicit information exchange. In addition, it uses a brief representation for the Q-values in which they are approximated as a function of much smaller set of variables. This expedites the convergence and reduces the algorithm related computations.

#### Intuition Based Power Allocation

The intuition idea is derived from the concept that different STs follow similar power allocation strategies at the same network state. To be able to estimate other STs power allocation strategies \( \pi^t_{i-1}(s_i) = (\pi^t_1(s_1), ..., \pi^t_{i-1}(s_{i-1}), \pi^t_{i+1}(s_{i+1}), ..., \pi^t_{N+D}(s_{N+D})) \) using non-cooperative learning scheme, we define an intuition factor as,

\[
\mu^t_i(s_i, a_{-i}) = \prod_{j \in \mathbb{N} \cup D / \{i\}} \pi^t_j(s_j, a_j) \quad (6.38)
\]
for ST $i$ at time slot $t$. This function conjectures the change in the Q-value $Q_{t}^{i+1}(s_i, a_i)$ in the next time slot $t + 1$ as a result of certain strategies employed by the other STs. $\mu^i_t(s_i, a_{-i})$ is assumed to be the only information that the learning agent knows about other STs and it is found based on local information. The probability that ST $i$ experience environment state $s^{t+1}$, which is the same as the probability that ST $i$ achieves a reward $R_t(s_i, a_i, a_{-i})$ is defined as follows,

$$\Gamma_i = \pi^i_t(s_i, a_i)\mu^i_t(s_i, a_{-i}) \quad (6.39)$$

The probability calculated in (6.39) is also the probability that ST $i$ achieves the reward function defined in (6.30). Let us assume that $\delta$ is the number of consecutive time slots that ST $i$ achieved the same reward. Consequently, $\delta$ has an independent and identical distribution with probability $\Gamma_i = \frac{1}{1+\delta'}$, where $\delta'$ is the mean of $\delta$ and can be found by ST $i$ through observing its reward history. Thus, the intuition factor can be estimated as $\mu^i_t(s_i, a_{-i}) = \frac{1}{(1+\delta')\pi^i_t(s_i, a_i)}$ as ST $i$ is aware of its own power allocation strategy $\pi^i_t(s_i, a_i)$. As the action available to ST $i$ is to choose the transmission power according to its strategy $\pi^i_t(s_i)$, the simplest method to express the intuition factor as a function of ST $i$ power allocation strategy is the following expression,

$$\mu^i_t(s_i, a_{-i}) = \mu'_i(s_i, a_{-i}) + w_i[\pi^i_t(s_i, a_i) - \pi'_i(s_i, a_i)] \quad (6.40)$$

where $\mu'_i(s_i, a_{-i})$ and $\pi'_i(s_i, a_i)$ are the reference points for specific intuition and probability of certain action selection, and $w$ is a positive scalar for linearization. The reference points are considered to be given of common knowledge. They are determined with assumption that other STs can observe ST $i$ deviation from its reference points $\pi^i_t(s_i, a_i)$ and $\mu^i_t(s_i, a_{-i})$ by a quantity proportional to $[\pi^i_t(s_i, a_i) - \pi'_i(s_i, a_i)]$. If the reference points are $\mu^i_t(s_i, a_{-i}) = \prod_{j \in N \cup D \setminus \{i\}} \pi_j^*(s_j, a_j)$ and $\pi^i_t(s_i, a_i) = \pi_i^*(s_i, a_i)$, then the optimal intuition factor is $\mu^*_i(s_i, a_{-i}) = \prod_{j \in N \cup D \setminus \{i\}} \pi_j^*(s_j, a_j)$ and this will
lead to an optimal transmission. The STs revise their reference points based on their historical local information about transmissions that achieved maximum Q-value. We define the following rule for the intuition factor of STs to update their reference points in time slot $t$,

$$\mu_i^t(s_i, a_{-i}) = \mu_i^{t-1}(s_i, a_{-i}) + w_i[\pi_i^t(s_i, a_i) - \pi_i^{t-1}(s_i, a_i)]$$  \hspace{1cm} (6.41)

where $\mu_i^t(s_i, a_{-i})$ and $\pi_i^t(s_i, a_i)$ are set to $\mu_i^{t-1}(s_i, a_{-i})$ and $\pi_i^{t-1}(s_i, a_i)$ respectively. This means that each ST believes that any modifications to its current strategy, will induce other STs to perform changes in the next time slot. We consider the deviation of ST $i$ from its reference points as the model to capture the strategies variation of other STs as follows,

$$\mu_i^t(s_i, a_{-i}) - \mu_i^{t-1}(s_i, a_{-i}) = w_i[\pi_i^t(s_i, a_i) - \pi_i^{t-1}(s_i, a_i)]$$  \hspace{1cm} (6.42)

Now, we can adjust the update rule in (6.37) by placing the intuition of ST $i$ in place of the allocation strategies of other STs. The new rule becomes,

$$Q_i^{t+1}(s_i, a_i) = (1 - \alpha^t)Q_i^t(s_i, a_i) + \alpha^t\left\{ \sum_{a_{-i} \in A_{-i}} R_i(s_i, a_i, a_{-i})\right\}$$

$$[\mu_i^t(s_i, a_{-i}) - \mu_i^{t-1}(s_i, a_{-i})] + \beta \max_{b_i \in A_i} Q_i^t(s_i', b_i)$$  \hspace{1cm} (6.43)

The update rule in (6.43) emphasizes the point that ST $i$ exploits its intuition factor variation to estimate how other STs strategies change in the stochastic learning process. Balancing exploration and exploitation is an essential issue in the stochastic power allocation process. Exploration aims to try new allocation strategies so it does not only apply the strategies it already knows to be good but also explores new ones. Exploitation is the process of using well-established strategies. The most common technique to achieve this balance is to use the $\epsilon$-greedy selection [206]. However, this
approach selects equally among the available actions i.e. (the worst action is likely to be chosen as the best one). In order to overcome the drawback of the $\epsilon$-greedy approach, the action selection probabilities are varied as a graded function of the Q-value. The best power level is given the highest selection probability while all other levels are ranked according to their Q-values. The learning algorithm exploits Boltzmann probability distribution to determine the probability of the power allocation action that fulfill the energy efficiency maximization constraints in C.1 to C.4. Thus, ST $i$ selects the action $a_i$ in state $s_i$ at time slot $t$ with the following probability,

$$
\pi_t^i(s_i, a_i) = \frac{e^{Q_t^i(s_i,a_i)/\tau}}{\sum_{b \in A_i} e^{Q_t^i(s_i,b_i)/\tau}}
$$

where $\tau$ is a positive integer that controls the selection probability. With high value of $\tau$, the action probabilities become nearly equal. However, low value of $\tau$ causes big difference in selection probabilities for actions with different Q-values.

**Approximated-Intuition Based Power Allocation**

The computational complexity of the system increases along with the size of the states and action spaces. The simple look-up table where separate Q-value is maintained for each state/action pair is not feasible in large space with massive number of states like our system. Therefore, we propose a brief representation for the Q-values in which they are approximated as a function of much smaller set of variables. The compact representation of $Q$ using function approximator $Q' : S' \times A$ is achieved by employing a vector $\xi = \{\xi_z\}_{z=1}^Z$ to minimize the metric of difference between the optimal Q-value $Q^*(s_i, a_i)$ and the approximated one $Q'^i_t(s_i, a_i, \xi)$. The approximated Q-value is expressed as follows,

$$
Q'^i_t(s_i, a_i, \xi) = \sum_{z=1}^Z \xi_z \psi_z(s_i, a_i) = \xi \psi^T(s_i, a_i)
$$
where $T$ denotes the transpose operator, each scalar $\psi_z(s_i, a_i)$ is defined as the basis function (BF) over $S' \times A$, and $\xi_z$ are the associated weights. The right hand side of (6.45) presents the vectors of the corresponding variables. We use the gradient function $\psi(s_i, a_i)$ to combine the online learning model with the brief representation. As a result, the update rule for the Q-value stated in (6.43) takes the following form,

$$\xi_{t+1}^i \psi^T(s_i, a_i) = \left\{ (1 - \alpha_t) \xi_t^i \psi^T(s_i, a_i) + \alpha_t \left[ \sum_{a_i \in A_i} R_i(s_i, a_i, a_{-i}) \left[ \mu_t^i(s_i, a_i) - \mu_{t-1}^i(s_i, a_i) \right] + \beta \max_{b_i \in A_i} \xi_t^i \psi^T(s_i', b_i) \right]\right\} \psi(s_i, a_i)$$

(6.46)

The gradient function $\psi(s_i, a_i)$ is a partial derivative with respect to the elements of $\xi_t$. Moreover, the probability of selecting certain action presented in (6.44) is updated with the Q-value approximation as follows,

$$\pi_t^i(s_i, a_i) = \frac{e^{\xi_t^i \psi^T(s_i, a_i)/\tau}}{\sum_{b_i \in A_i} e^{\xi_t^i \psi^T(s_i, b_i)/\tau}}$$

(6.47)

The intuition based online learning process with approximated Q-value is illustrated in Algorithm 5. The algorithm starts by initializing the power allocation strategy, intuition factor and the approximated Q-value for each state belong to the reduced state space. Once the state is initialized, certain transmission power is selected for the corresponding ST according to the probability in (6.47). If the conditions C.1 to C.4 are satisfied, then, the reward is achieved and the Q-value, intuition factor and power allocation strategy are updated and the new state is observed.

### 6.2.5 Power Allocation Algorithm Convergence

We prove the convergence of the proposed approximated intuition-based online learning algorithm for power allocation. Our proof relies on exploiting ordinary differential
Algorithm 5 Approximated intuition based online learning algorithm for power allocation

Require: $\pi^t_i(s_i, a_i), t, w_i > 0, \gamma^*_n, \gamma^*_d$ and $\gamma^*_k$

Ensure: Transmission power allocation for STs

1: initialization
2: Let $t = 0$
3: for each $(s_i, a_i \in A_i)$ do
4: initialize power allocation strategy $\pi^t_i(s_i, a_i)$;
5: initialize approximated Q-value $\xi^t_i\psi^T(s_i, a_i)$;
6: initialize intuition factor $\mu^t_i(s_i, a_{-i})$;
7: end for
8: evaluate the state $s_i = s_i^t$
9: while (true) do
10: Select action $a_i$ according to $\pi^t_i(s_i, a_i)$;
11: Measure the received $\gamma_{n,x}, \gamma_d$ and $\gamma_k$ with feedback from the receiver and observe the state $s_i'$ by identifying $P_i$ and comparing SINR;
12: if $(\gamma_{n,x} \geq \gamma^*_n, \gamma_d \geq \gamma^*_d$ and $\gamma_k \geq \gamma^*_k)$ then
13: $R_i(s_i, a_i, a_{-i})$ is achieved;
14: else
15: $R_i(s_i, a_i, a_{-i}) = 0$ as the receiver could not receive the data correctly
16: end if
17: Update $\xi^{t+1}_i\psi^T(s_i, a_i)$ based on $\mu^t_i(s_i, a_{-i})$ according to (6.46)
18: Update $\pi^{t+1}_i(s_i, a_i)$ according to (6.47)
19: Update $\mu^{t+1}_i(s_i, a_{-i})$ according to (6.41)
20: $s_i = s_i^{t+1}$
21: $t = t + 1$
22: end while
equations (ODE) to acquire the necessary conditions for convergence. The following assumptions are required to proceed with the proof:

**Assumption 1.** The basis functions \( \psi_z(s_i, a_i) \) are linearly independent for all \((s_i, a_i)\) and all the properties of \( Q'_t(s_i, a_i) \) in previous discussion are applicable to the dot product for the vectors \( \xi_t^T \psi(s_i, a_i) \).

**Assumption 2.** For every \( z = (1, 2, \ldots, Z) \), \( \psi_z(s_i, a_i) \) is bounded, which means \( E\{\psi_z^2(s_i, a_i)\} < \infty \) and the reward function satisfies \( E\{R^2_i(s_i, a_i, a_{-i})\} < \infty \).

**Assumption 3.** The learning rate satisfies \( \sum_{t=1}^{\infty} \alpha_t = \infty \) and \( \sum_{t=1}^{\infty} (\alpha_t)^2 < \infty \).

**Definition 1.** Let \( \Psi = E[\psi^T(s_i, a_i)\psi(s_i, a_i)] \). For the parameter vector \( \xi \) and a particular network state \( s_i \in S' \), we define a vector \( \psi(s_i, \xi) = [\psi_z(s_i, a_i)] \) for \( z = 1 \rightarrow Z \) where \( a_i \in \{ a_i = \arg \max_{b_i \in A_i} \xi_i \psi^T(s_i, b_i) \} \) is the set of optimal power allocation actions for \( s_i \). We define the following \( \xi \)-dependent matrix:

\[
\Psi = E[\psi^T(s_i, \xi)\psi(s_i, \xi)]
\] (6.48)

**Proposition 1.** With the assumptions 1-3 and Definition 1, the intuition based online learning with approximation converges with probability (w.p) 1, if

\[
\Psi' < \Psi, \ \forall \xi
\] (6.49)

**Proof.** The proof of convergence is linked to finding stable fixed points of the ODE defined based on the expectation of the derivative of the update rule in (6.46) with respect to \( t \) as follows,

\[
\xi_{t+1}^i = E[(\sum_{a_{-i} \in A_{-i}} [\mu_{t}^i(s_i, a_i) - \mu_{t-1}^i(s_i, a_i)]R_i(s_i, a_i, a_{-i}) + \beta \xi_t^i \psi^T(s'_i, \xi_t^i) - \xi_t^i \psi^T(s_i, a_i))]\psi(s_i, a_i)]
\] (6.50)
where $\xi_i^t = \frac{\partial \xi}{\partial t}$ as $\alpha \to 0$. From the definition of the intuition factor in (6.42), we state,

$$\sum_{a_i \in \mathcal{A}} \left[ \mu_i^t(s_i, a_i) - \mu_i^{t-1}(s_i, a_i) \right] R_i(s_i, a_i, a_{-i})$$

$$= \sum_{a_i \in \mathcal{A}} w_i [\pi_i(s_i, a_i) - \pi'_i(s_i, a_i)] R_i(s_i, a_i, a_{-i})$$

(6.51)

By substituting the value of $\pi_i(s_i, a_i)$ from (6.47), and when $\tau$ is large, we obtain,

$$e^{\xi_i^T(s_i, a_i)/\tau} = 1 + \frac{\xi_i^T(s_i, a_i)}{\tau} + \rho(\frac{\xi_i^T(s_i, a_i)}{\tau})$$

where $\rho(\frac{\xi_i^T(s_i, a_i)}{\tau})$ is a polynomial of order $O((\xi_i^T(s_i, a_i)/\tau)^2)$, we can easily find,

$$\sum_{b \in \mathcal{A}_i} e^{\xi_i^T(s_i, b_i)/\tau} = m_i + 1 + \frac{\xi_i^T(s_i, b_i)}{\tau} + \rho(\frac{\xi_i^T(s_i, b_i)}{\tau})$$

where $m_i$ is the number of power levels considered in certain range. Consequently, we get the following,

$$\pi_i(s_i, a_i) = \frac{1}{m_i + 1} + \frac{1}{m_i} \frac{\xi_i^T(s_i, a_i)}{\tau} + \varrho(\frac{\xi_i^T(s_i, b_i)}{\tau})$$

(6.52)

where $\varrho(\frac{\xi_i^T(s_i, b_i)}{\tau})$ is a polynomial of order smaller than $O(\xi_i^T(s_i, a_i)/\tau)$, Note that the coefficient of the polynomial is independent of the vector Q-value. The reference strategy is evaluated according to the historical optimal actions as follows,

$$\pi'_i(s_i, a_i) = \frac{1}{m_i + 1} + \frac{1}{m_i} \frac{\xi_i^T(s_i, \xi)}{\tau} + \varrho(\frac{\xi_i^T(s_i, b_i)}{\tau})$$

(6.53)

By substituting (6.52) and (6.53) in (6.51), we get,

$$\sum_{a_i \in \mathcal{A}} \left[ \mu_i^t(s_i, a_i) - \mu_i^{t-1}(s_i, a_i) \right] R_i(s_i, a_i, a_{-i})$$
\[
\sum_{a_i \in A_i} w_i R_i(s_i, a_i, a_{-i}) \frac{1}{\tau} \left[ \frac{\xi_i \psi^T(s_i, a_i) - \xi_i \psi^T(s_i, b_i)}{\tau} \right] 
\]

\[
= \sum_{a_i \in A_i} \left[ [\mu_i^t(s_i, a_i) - \mu_i^{t-1}(s_i, a_i)] R_i(s_i, a_i, a_{-i}) \right] 
\]

\[
\leq \frac{1 - \beta}{m_i + 1} \left[ \xi_i \psi^T(s_i, a_i) - \xi_i \psi^T(s_i, \xi) \right] 
\]

Let us assume that \( \nu = \frac{1 - \beta}{m_i + 1} \) to simplify the notation. Now, (64) can be expressed as follows,

\[
\xi^t = E \left[ (V[\xi^t \psi^T(s_i, a_i) - \xi^t \psi^T(s_i, \xi^t)] + \beta \xi^t \psi^T(s_i, \xi^t) \right. 
\]

\[
- \xi^t \psi^T(s_i, a_i) \psi(s_i, a_i) \right] \] 

(6.54)

We define two trajectories of the ODE \( \xi^t \) and \( \xi^t \) that have different initial conditions and satisfies \( \xi_0^t = \xi_1^t - \xi_2^t \). Then, we have

\[
\frac{\partial \|\xi_0^t\|^2}{\partial t} = 2(\xi_1^t - \xi_2^t)(\xi_0^t)^T 
\]

\[
= E \left[ (-2V \xi_1^t \psi^T(s_i, \xi_1^t) + 2\beta \xi_1^t \psi^T(s_i, \xi_1^t) \psi(s_i, a_i)(\xi_0^t)^T - 
\]

\[
( -2V \xi_2^t \psi^T(s_i, \xi_2^t) + 2\beta \xi_2^t \psi^T(s_i, \xi_2^t) \psi(s_i, a_i)(\xi_0^t)^T \right] 
\]

\[
+ (2V - 2) \xi_0^t \Psi(\xi_0^t)^T \] 

(6.55)

From Definition 1, we can deduce the following two inequalities,

\[
\xi_1^t \psi^T(s_i', \xi_2^t) \leq \xi_1^t \psi^T(s_i', \xi_2^t) \] 

(6.56)

\[
\xi_2^t \psi^T(s_i', \xi_2^t) \leq \xi_2^t \psi^T(s_i', \xi_2^t) \] 

(6.57)
As the expectation $E$ in (6.55) is taken over different states and different actions, we can define two sets $\Lambda_+ = \{(s_i, a_i) \in S_i \times A_i | \xi_0^T \psi(s_i, a_i) > 0\}$ and $\Lambda_- \subseteq S_i \times A_i - \Lambda_+$. If we combine (6.56) and (6.57) in (6.55), we get,

$$\frac{\partial \|\xi_0\|^2}{\partial t} \leq E[(-2V\xi_0^T \psi(s, \xi_2^i) + 2\beta \xi_0^T \psi(s, a_i)(\xi_0)^T | \Lambda_+]$$

$$+ E[(-2V\xi_0^T \psi(s, \xi_1^i) + 2\beta \xi_0^T \psi(s, a_i)(\xi_0)^T | \Lambda_-]$$

$$+ (2V - 2)\xi_0^T \Psi(\xi_0)^T$$

(6.58)

After the application of Holder’s inequality [207] to the expectation in (6.58), we get

$$\frac{\partial \|\xi_0\|^2}{\partial t} \leq \left(-2V \sqrt{E[(\xi_0^T \psi(s, \xi_2^i))^2 | \Lambda_+]} \right)$$

$$+ 2\beta \sqrt{E[(\xi_0^T \psi(s, \xi_2^i))^2 | \Lambda_+]} \times \sqrt{E[(\psi(s, a_i)(\xi_0)^T)^2 | \Lambda_+]$$

$$+ \left(-2V \sqrt{E[(\xi_0^T \psi(s, \xi_1^i))^2 | \Lambda_-]} + 2\beta \sqrt{E[(\xi_0^T \psi(s, \xi_1^i))^2 | \Lambda_-]} \right) \times \sqrt{E[(\psi(s, a_i)(\xi_0)^T)^2 | \Lambda_-]$$

$$+(2V - 2)\xi_0^T \Psi(\xi_0)^T$$

$$\leq \left(-2V \sqrt{E[(\xi_0^T \psi(s, \xi_2^i))^2]} + 2\beta \sqrt{E[(\xi_0^T \psi(s, \xi_2^i))^2]} \right)$$

$$\times \sqrt{E[(\psi(s, a_i)(\xi_0)^T)^2 | \Lambda_+]$$

$$+ \left(-2V \sqrt{E[(\xi_0^T \psi(s, \xi_1^i))^2]} + 2\beta \sqrt{E[(\xi_0^T \psi(s, \xi_1^i))^2]} \right) \times \sqrt{E[(\psi(s, a_i)(\xi_0)^T)^2 | \Lambda_-]$$

$$+(2V - 2)\xi_0^T \Psi(\xi_0)^T$$
If we apply the definition of $\Psi'$ in Definition 1, we get,

$$
\leq (-2V + 2\beta)\sqrt{\max[\xi_0^1\Psi_1'(\xi_0^1)^T, \xi_0^2\Psi_2'(\xi_0^2)^T]}
\times \left( \sqrt{\mathbb{E}[\psi(s_i,a_i)(\xi_0^1)^T] \Lambda_+} + \sqrt{\mathbb{E}[\psi(s_i,a_i)(\xi_0^2)^T] \Lambda_-} \right) + (2V - 2)\xi_0^1\Psi(\xi_0^1)^T
$$

which means that $\xi_0$ converges to the origin and this confirms that there exists a stable point of the ODE in (6.64). Thus, the proposed intuition based online learning with Q approximation converges w.p 1.

According to the condition in (6.49), we can state that,

$$
\frac{\partial\|\xi_0\|^2}{\partial t} < (-2V + 2\beta)\xi_0^1\Psi(\xi_0^1)^T + (2V - 2)\xi_0^1\Psi(\xi_0^1)^T
= (2\beta - 2)\xi_0^1\Psi(\xi_0^1)^T < 0
$$

Consequently, the stable point $\xi^*$ of the ODE in (6.64) indicates that,

$$
0 = E\left[ \sum_{a_{-i} \in A_{-i}} [\mu_i^t(s_i,a_i) - \mu_i^{t-1}(s_i,a_i)] R_i(s_i,a_i,a_{-i}) \right.
+ \beta\xi^*\psi^T(s_i,\xi^*) - \xi^*\psi^T(s_i,a_i))\psi(s_i,a_i) \bigg] 
$$

and $\xi^*$ can be found as follows,

$$
\xi^* = E\left[ \sum_{a_{-i} \in A_{-i}} [\mu_i^t(s_i,a_i) - \mu_i^{t-1}(s_i,a_i)] R_i(s_i,a_i,a_{-i}) \right. 
$$
\[ 200 + \beta \xi^* \psi^T(s_i', \xi^*) \psi(s_i, a_i)] \Psi^{-1} \]  

(6.62)

As a result, the optimal intuition based approximated online Q-function is defined as follows

\[ Q'(s_i, a_i, \xi^*) = \xi^* \psi(s_i, a_i) \]  

(6.63)

### 6.2.6 Numerical Results

The simulation results obtained to demonstrate the capability of our proposed scheme for energy efficient power allocation are presented in this section. The simulation environment comprises multi-tier heterogeneous 5G network constructed according to the network model, where multiple picocells, femtocells and D2D share the spectrum resources with one macrocell in an underlay fashion. The multi-tier network is composed of one macrocell, 4 picocell, 8 femtocells and variable number of D2D connections. The pico, femto BSs and D2D users are uniformly distributed within an area centered by the macro BS with radiuses of 400m for the macrocell, 75 m for the picocell, and 20 m for the femtocell. All the users are assumed to have identical and independent Rayleigh fading channels. The channel gain is given as \( G = F(d)^{-c} \) where \( F \) is shadowing factor, which is a random number generated with log normal distribution with mean 0 and variance 6 dB, and \( d \) is the physical distance between the user and its BS. The learning rate \( \alpha \) is chosen to be dynamic according to the Win or learn fast principle, which states that the learning agent should learn faster when it is losing and more slowly upon winning [54]. The learning rates that we used are \( \alpha = 0.2 \) for rewarded solution and \( \alpha = 0.6 \) for the punished one. The rest of the simulation parameters are presented in Table 6.2. The service request arrival in each time \( t \) follows Poisson distribution with an arrival rate of \( 6 \lambda \) where \( \lambda = 0.3 \). The thresholds \( \gamma_{n,x}^*, \gamma_d^* \) and \( \gamma_k^* \) are specified as follows, \( \gamma_{n,x}^* \) is 20 dB and 15 dB for pico and femto receivers respectively. The \( \gamma_d^* \) is -8 dB, while \( \gamma_k^* \) is 30 dB. We compare the performance of our proposed scheme with non-cooperative online learning (NCOL)
<table>
<thead>
<tr>
<th>Simulation Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td>20 MHz</td>
</tr>
<tr>
<td>Discount factor ($\beta$)</td>
<td>0.9</td>
</tr>
<tr>
<td>Number of MUEs</td>
<td>10</td>
</tr>
<tr>
<td>Pico BS Tx power</td>
<td>20 to 28 dBm</td>
</tr>
<tr>
<td>Femto BS Tx power</td>
<td>12 to 17 dBm</td>
</tr>
<tr>
<td>Macro BS Tx power</td>
<td>45 dBm</td>
</tr>
<tr>
<td>DUE Tx power</td>
<td>-8 to -2 dBm</td>
</tr>
<tr>
<td>$p_{cc}$ for macro, pico, femto, D2D transmitters respectively</td>
<td>130 W, 15 W, 5 W, 0.1 W</td>
</tr>
<tr>
<td>Thermal noise density ($\sigma$)</td>
<td>-174 dBm/Hz</td>
</tr>
</tbody>
</table>

Table 6.2: 5G multi-tier environment simulation parameters

where STs act selfishly to improve their own rewards. In addition, we compare it to the stackelberg game (SG) theory for power allocation proposed in [87], heuristic based scheme in [95] and sub-optimal algorithm (SOA) proposed in [93]. Note that schemes in [95] and [93] require awareness of the network model and rely on information exchange between the secondary BSs. The proposed scheme performance is demonstrated in four sets of simulations.

In the first simulation set, the average achieved energy efficiency for the whole system is evaluated. The convergence behavior is estimated for each scheme by showing how the system behaves as a function of the number of epoch as in Figure 6.10. The impact of the number of STs on the system energy efficiency is evaluated in Figure 6.11 where there are 4 pico BSs, 8 femto BSs and 6 D2D connections. Figure 6.12 presents the achieved system energy efficiency across the channel gain to noise ratio (CNR) from 0 to 30 dB. Unlike the other schemes, we notice that the proposed scheme converges faster than other schemes due to the approximation of Q-value which reduces the state space and information exchange overhead. Consequently, this validates Proposition 1. Moreover, our scheme records high energy efficiency compared to others. This is due to the fact that intuition about other agents’ strategies through online learning approach improves the quality of the action selected in each state. One observation to note is that the energy efficiency first increases then
Figure 6.10: Average system energy efficiency of different schemes in 5G multi-tier network

Figure 6.11: Impact of the number of STs on system energy efficiency in 5G multi-tier network

decreases as the number of STs increases. This is because when there are only few STs in the heterogeneous 5G network, the inter-cell interference caused by the macro BS and other STs is comparably low and there exists enough signal spaces to compensate for its impact. However, when there are more STs, the inter-cell interference becomes significant and more power is required to satisfy the minimum rate target, resulting in a lower energy efficiency.

The second simulation set demonstrate the scheme performance in terms of SE.
Figure 6.12: Average system energy efficiency under different CNR ratio in 5G multi-tier network

This simulation part measures the speed of convergence as in Figure 6.13, which presents the average spectral efficiency of the whole 5G system. In addition, we study the impact of the number of STs from certain type on the achieved spectral efficiency per each ST of that type. For example, the average spectral efficiency per each femtocell is plotted in Figure 6.14 against the number of femtocells in the system. Another evaluation metric considers the impact of the choice of the maximum transmission power of STs on the system spectral efficiency. We select picocells
Figure 6.14: Average spectral efficiency per each femtocell with variable number of femtocells in 5G multi-tier network as an instance to measure the impact of the maximum transmission power choice. Figure 6.15 shows the achieved system spectral efficiency with variable maximum transmission power of pico BS. It is clear that our scheme achieved the highest spectral efficiency with the fastest convergence among other schemes. Another notice is that the spectral efficiency decreases with the increase in the number of STs (femto BSs) in Figure 6.14. We also observe that the interference threshold imposed by the macrocell transmission prevents any performance improvement even when increasing the pico BS transmission power as in Figure 6.15.

The third simulation set focus on the evaluation of users QoS satisfaction. Figure 6.16 and Figure 6.17 present the average SINR measured at the macro receivers and pico receivers respectively for all the schemes. We notice that our scheme achieved higher SINR than others and maintains SINR above threshold for both receivers. This confirms the satisfaction of constraints C.1 and C.2 as the proposed scheme managed to maintain QoS for both macro and secondary tier users.

In the fourth simulation set, we aim at demonstrating the capability of our learning scheme to produce acceptable results in terms of QoS with variable number of UEs in the network. Thus, we plot as in Figure 6.18 the achieved SINR for MUEs against
Figure 6.15: Average system spectral efficiency as a function of maximum transmission power of pico BS in 5G multi-tier network

Figure 6.16: Average SINR for macro users of different schemes in 5G multi-tier network

the total number of UEs. In addition, the SINR for each pico user, which is selected as a representation for the SUEs is plotted in Figure 6.19. Figure 6.20 presents the achieved SINR for D2D receiver with variable number of UEs. The contribution of each tier UEs in the total number of network UEs is determined with the following percentages: 25 % MUEs, 25 % pico users, 25 % femto users, and 25 % D2D users. We notice that our scheme achieved the highest SINR for both tier users compared to other schemes. In addition, it managed to maintain the minimum SINR required by
Figure 6.17: Average SINR for pico users of different schemes in 5G multi-tier network

Figure 6.18: Average SINR for each macro UE with variable number of UEs in 5G multi-tier network
each tier even with high number of UEs. These results are anticipated as the machine learning implemented by our power allocation scheme enforces the QoS constraints during the learning process.

6.3 Energy Efficient Traffic Offloading in Multi-tier Heterogeneous 5G Networks Using Intuitive Online Learning

In this section, we propose a traffic offloading scheme that offloads traffic from congested macrocells to small cells including picocells and femtocells with considera-
tion of energy efficiency under the condition that the system load of all the cells is maintained below certain threshold [208]. The system load is an essential factor for quantifying energy efficiency as it grows with traffic demand and the amount of interference from the active cells. In addition, the user QoS is also dependent on the system load as when the system load is low, BSs are able to offer their users a good capacity to meet QoS requirements. To achieve efficient traffic offloading, we develop an intuitive online learning methodology, where each learning agent (macrocell) aims
to perform traffic offloading by surmising other agents traffic offloading actions for certain network state without cooperation and information exchange. The network state evolves according to a Discrete-Time Markov decision process (DTMDP), whose statistics depend on the traffic offloading strategy [112]. Hereby, a traffic offloading strategy is defined as a sequence of actions, and each action takes the form of small-cell operations (e.g., switching on a small cell to offload the traffic demands within its respective coverage or switching off a small cell to save energy). Our task is to maximize the overall energy efficiency in the network while satisfying the constraint of flow-level QoS requirement. This work has the following contributions:

- The thesis proposes an online learning scheme to select an efficient traffic offloading strategy and determine the most efficient mode of operation for small cells. The proposed scheme considers cell load as a driving factor to determine the proper small cell operation mode and to control interference caused by small cells activation.

- The proposed online learning scheme deals with curse of the exponential growth in the action space as a result of activating more small cells by introducing the intuition feature to online learning with which the scheme has the ability to make each macro BS conjectures other macro BS offloading strategies to find a joint offloading action without explicit information exchange. This has a positive impact on the learning process as it reduces the scheme computation overhead and improves the selected action quality.

6.3.1 System Model

The considered system model tackles the downlink communication in a multi-tier heterogeneous network. The multi-tier structure consists of a primary tier represented by macrocells, and secondary tier that includes picocells and femtocells that share the same spectrum with the primary tier as in Figure 6.21. This multi-tier structure
operates over discrete time slots each with constant time duration. The service region for each cell is represented by a set of $L$ locations, each being characterized by uniform signal propagation conditions [101]. At each location $l \in L$ in each time slot ($t = 1, 2, \ldots$), the service requests follow Poisson arrival process with arrival rate $\lambda(l, t)$. The size of the requested traffic demand is assumed to be an exponentially distributed variable, which makes the network Markovian [209]. The coverage in the network is maintained by a set of macro BSs $K = \{1, 2, \ldots, K\}$ with $N_k$ small cells (pico and femto) that operate within each macro BS $k \in K$. These small BSs are connected to the macro BS using a logical interface. We choose $L_k$ and $L_n$ to designate the sets of locations covered by the macro BS $k$ and small BS $n$ respectively, where $L_n \subset L_k$.

The word cell and BS are used interchangeably in the rest of the thesis. The small cells are assumed to have an open access for macrocells to offload traffic to them. The ON/OFF operations of the $N_k$ small BSs are dynamically determined by the macro BS (offloading controller) that they operate under its coverage. If a small cell is activated, the macro users within its coverage must associate with the BS that satisfy their QoS whether it is the macro BS or the activated small BS. The advantage

Figure 6.21: System model of multi-tier heterogeneous 5G networks

The ON/OFF operations of the $N_k$ small BSs are dynamically determined by the macro BS (offloading controller) that they operate under its coverage. If a small cell is activated, the macro users within its coverage must associate with the BS that satisfy their QoS whether it is the macro BS or the activated small BS. The advantage
of the small cells is their ability to reduce power consumption when they serve the offloaded users from the macro BSs in the heterogeneous network. In this model, we aim to investigate traffic offloading problem from macro BS to small BSs in an energy efficient fashion. The offloading process is triggered based on the system load of the macro BS. For example, when a macro BS is lightly loaded, the offloading is not necessary and small BSs are maintained in OFF status to save energy. However, the heavily loaded macro BS is allowed to offload traffic to activated small BSs and remains ON to handle the traffic of the macro users that are not offloaded [210]. The macro BS controls the operation mode of the small BSs within its coverage based on the system load information obtained from different BSs. Let $s(t)$ describes the evolution of the network state across time slots and it is defined as the number of users of different BSs at different locations. Every small BS under the coverage of the macro BS $k \in K$ is labeled as $k_n$ where $n = 0, 1, 2, ..., N_k$, $n = 0$ for the macro BS and the location covered by the BS $k_n$ are labeled from 1 to $L_{k_n}$. Thus, the configuration of the macro BS and all small cells within its coverage is defined as follows,

$$s_k(t) = s_{k0}^{L_{k0}}(t), ..., s_{k0}^{L_{k1}}(t), s_{k1}^{L_{k1}}(t), ..., s_{kN_k}^{L_{kN_k}}(t),$$

$$...,$$

(6.64)

Each element in 6.64 represents the number of users at the location covered by a BS within the macro BS $k$ during time slot $t$. The operation mode for small BSs under the coverage of the macro BS $k$ is represented by $A_k(t) = (a_{k1}(t), ..., a_{kN_k}(t))$ with $a_{k_n}(t) = 1$ if the small BS is ON and 0 otherwise for all $n \in \{1, ..., N_k\}$. To proceed with the transmission scheduled to deliver the traffic demand of certain user, let $SL_k(s, a)$ and $SL_n(s, a)$ be the level of utilization of the macro BS $k$ and the small BS $n$ respectively. The average SINR achieved by MUE located at $l \in L_k$ associated
with macro BS $k$ is evaluated as,

$$
\gamma_{kl}(s,a) = \frac{G_{kl}P_k}{\sum_{i \in K/\{k\}} G_{il}P_i S L_i(s,a) + \sum_{n \in N} G_{nl}P_n S L_n(s,a) + \sigma} \quad (6.65)
$$

where $G_{kl}$, $G_{il}$ and $G_{nl}$ are the channel gain between the macro user and the macro BS $k$, other macro BS $i$ and small BS $n$, respectively, $P_k$, $P_i$, $P_n$ are the transmission power of macro BS $k$, other macro BS $i \in K$ and small BS $n \in N$ respectively, and $\sigma$ is the background noise. The SINR for the offloaded macro user associated with small BS $n$ located at $l \in L_n$ is calculated as follows,

$$
\gamma_{nl}(s,a) = \frac{G_{nl}P_n}{\sum_{k \in K} G_{kl}P_k S L_k(s,a) + \sum_{q \in N/\{n\}} G_{ql}P_q S L_q(s,a) + \sigma} \quad (6.66)
$$

where $G_{nl}$, $G_{kl}$ and $G_{ql}$ are the channel gain between the offloaded users and the serving small BS $n$, macro BS $k$ and other small BS $q$, respectively at location $l$, $P_n$, $P_k$, $P_q$ are the transmission power of the serving small BS $n$, macro BS $k \in K$ and other small BS $q \in N$ respectively. The achievable data rates for MUEs associated with the macro BS and MUEs offloaded to small BS are calculated as in $(6.67)$ and $(6.68)$ respectively,

$$
R_{kl}(s,a) = B \log_2(1 + \gamma_{kl}(s,a)) \quad (6.67)
$$

$$
R_{nl}(s,a) = B \log_2(1 + \gamma_{nl}(s,a)) \quad (6.68)
$$

where $B$ is the system bandwidth. The system loads $SL_k(s,a)$ and $SL_n(s,a)$ for both macro and small cell during time slot $t$ are evaluated according to the traffic demand $\zeta(l,t)$ and the obtained data rate in $(6.67)$ and $(6.68)$ as

$$
SL_k(s,a) = \sum_{l \in L_k^t} \frac{\zeta(l,t)}{R_{kl}(s,a)} \times \frac{1}{T} \quad (6.69)
$$
\[ SL_n(s,a) = \sum_{l \in L'_n} \frac{\zeta(l,t)}{R_{nl}(s,a)} \times \frac{1}{T} \]  \hspace{1cm} (6.70)

where \( L'_k(t) = L_k \) excluding \( L'_n(t) \forall n \in N \) and \( L'_n(t) \) are the set of locations associated with the macro BS \( k \) and the small BS \( n \) during the time slot \( t \) respectively, and \( T \) is the duration of the time slot in seconds. Note that \( L'_n(t) = L_n \subset L_k \) for an active small BS and \( L'_n(t) = \phi \) if the small cell is OFF. Consequently, the system load for each cell type can be interpreted as the fraction of time scheduled for serving the requested traffic demand or the probability of causing interference to other cells.

### 6.3.2 Problem Formulation

In this section, we formulate the energy aware traffic offloading in 5G heterogeneous network, which is described by users arrival, generation of service requests, BSs that serve these requests, and system load for each cell. Heavily loaded macro BSs provide poor QoS, which requires activation of the within the coverage small BSs to handle the offloaded traffic. The activation of these small BSs creates interference due to spectrum sharing with the macro cells. The energy efficiency achieved by certain BS is dependent on the system load of that BS. Therefore, for certain network configuration \( s(t) \) and operation mode of small cell \( a(t) \) at time slot \( t \), the energy efficiency of macro BS \( k \in K \) is calculated as follows,

\[ EE_k(s,a) = \frac{R_{kl}(s,a)}{P_{ct}^k + \Gamma_k SL_k(s,a)P_k} \]  \hspace{1cm} (6.71)

where \( P_{ct}^k \) is the power consumed by the BS circuit for signal processing and \( \Gamma_k \) is a linear transmission power dependence factor. Similarly, the energy efficiency for small cell \( n \) is calculated as follows,

\[ EE_n(s,a) = \frac{R_{nl}(s,a)}{P_{ct}^n + \Gamma_n SL_n(s,a)P_n} \]  \hspace{1cm} (6.72)
where $P_n$ is the power consumed by the BS circuit for signal processing and $\Gamma_n$ is a linear transmission power dependence factor. The condition $a_{kn}(t) = 1$ to indicate that energy consumption occurs only if the small cell is active. The total energy efficiency for the entire network in time slot $t$ is found as,

$$EE_{TO}(s, a) = \sum_{k \in K} \left( EE_k(s, a) + \sum_{n \in N} EE_n(s, a) \right)$$

The system load $SL_z(s(t), a(t))$, where $z \in K \cup N$ and $t = 1, 2, 3, \ldots$, is an essential factor for quantifying energy efficiency as it grows with the traffic demand and the interference from the active cells. In addition, the user QoS represented by throughput or delay is also dependent on the system load as when the system load is low, BSs are able to offer their users a good capacity to meet QoS requirements. However, heavily loaded system leads to poor QoS and service outage. In that case, macro BS should trigger the offloading process. The overall goal of the proposed offloading scheme is to find an optimal strategy to perform traffic offloading and control the operation of the small cells that maximizes the whole network energy efficiency with condition that maintains the system load of cells below certain threshold, which implicitly include QoS guarantee and control interference. The optimization problem for finding offloading and small cells activation strategy is formulated as follows,

$$\max_{\pi \in \Pi} EE_{TO}(\pi(s, a)) \ s.t \ C.1 \ SL_z(s(t), a(t)) \leq SL^{th}_z$$

where $\Pi$ is the set of all available traffic offloading strategies, and $SL^{th}_z$ is a predefined threshold for the system load for each cell, the incentive of which is to incorporate the flow-level performance when transmission delay is a concern [10]. For example, when the threshold’s value is small, the BS operates with low usage of resources and
experience high throughput and less delay.

6.3.3 Stochastic Traffic Offloading Using Intuitive Online Learning

The goal of this section is to tackle the offloading optimization problem formulated using our developed intuitive online learning. We define a DTMDP that is associated to each network state, action, transition function and reward. Then, the section describes our proposed intuitive machine learning mechanism for traffic offloading in 5G networks. The reason for choosing online learning is that it is difficult to determine an exact state transition model for a practical DTMDP with large state space. Thus, it is challenging to compute the optimal traffic offloading strategy through applying a model-based dynamic programming algorithms.

Online Learning Model

In this model, each macro BS plays the role of an intelligent agent, which observes the network state \( s(t) \) and its associated action \( a(t) \) in the current time slot \( t \). At the end of each time slot, the reward, which evaluates the selected action is generated and a state transition to the next state is performed. The state, action, transition function and reward for our model are defined as follows,

- **State**: Since there is no cooperation among the macrocells, the network state \( s(t) \) is defined based on the local information according to the configuration of BSs and their coverage locations defined in 6.64.

- **Action**: The action selected by the macrocell \( a(t) \) according to the state information is defined as the operation mode of the small BS, which takes the value \( a_{k_n}(t) = 1 \) if the small BS is activated and 0 otherwise.

- **Reward**: The reward function is the received reward due to the selected action and is defined as \( RW(s,a) = EE_{TO}(s,a) \) where \( EE_{TO}(s,a) \) is the energy
efficiency defined in 6.73. However, the reward function is only achievable if the
constraint C.1 is met. Otherwise, it is evaluated to 0.

- Transition Function: The transition function $T_f$ is evaluated while moving from
the state $s(t)$ to $s'(t+1)$ as a result of the selected offloading action $a_{kn}(t)$ by
the macro BS at time slot $t$. The evolution of the DTMDP is Markovian with
the following state transition probability,

$$T_f(s, a, s') = Pr(s(t+1) = s'| s(t) = s, a(t) = a) \quad (6.75)$$

This transition probability is evaluated according to [112] and it depends on the
arrival of the users in the network.

In non-cooperative online learning [205], each macrocell $k \in K$ selects the strategy
$\pi_k(s_k)$ independently to maximize its total discounted reward. However, independent
learning causes an exponential growth in the number of actions with the increase in
the number of small BSs deployed in the network, which creates a potential challenge
in facilitating the non-cooperative learning. Therefore, it is necessary that each macro
BS becomes aware of other macro BSs traffic offloading strategies and learn coopera-
tively to reach the most appropriate offloading strategy and avoid interference among
macro BSs and small cells. This is achieved by defining a common goal of finding
a joint traffic offloading strategy $\pi \in \Pi$ that maximizes the total energy efficiency.
Thus, the total discounted reward in our model is defined as follows,

$$\max_{\pi_k \in \Pi_k} \left\{ E \left[ \sum_{t=0}^{\infty} \beta^t RW_k(s^t_k, \pi_k(s^t_k), \pi_{-k}(s^t_k)) \right] \right\} \quad (6.76)$$

where $\pi_{-k}(s^t_k) = (\pi_1(s^t_1), \ldots, \pi_{k-1}(s^t_{k-1}), \pi_{k+1}(s^t_{k+1}), \ldots, \pi_K(s^t_K))$ where $\beta$ is the discount factor. The strategy $\pi_k(s_k, a_k)$ is defined as the
probability vector with which the macro BS $k$ selects action $a_k = a_{kn}(t)$ at the state
\( s_k \) and \(-k\) is the index for other macro BSs. The total expected discounted reward of cell \( k \) over infinite iterations and with being aware of other cells strategies \( \pi_{-k} \) is given by,

\[
V_k(s_k, \pi_k, \pi_{-k}) = E[RW_k(s_k, \pi_k(s_k), \pi_{-k}(s_k))]+ \\
\beta \sum_{s'_k \in S_k} T_f(s_k, a, s'_k) (\pi_k(s_k), \pi_{-k}(s_k)) V_k(s'_k, \pi_k, \pi_{-k})
\]

(6.77)

where,

\[
E[RW_k(s_k, \pi_k(s_k), \pi_{-k}(s_k))] = \\
\sum_{(a_k, a_{-k}) \in A} \left[ RW_k(s_k, a_k, a_{-k}) \prod_{j \in K/\{k\}} \pi_j^*(s_j, a_j) \right]
\]

(6.78)

The optimal strategy that the learning agent aims to achieve \( \pi_k^* \) for each environment state and satisfies Bellman’s optimality equation \([131]\) achieves the following total discounted reward,

\[
V_k(s_k, \pi_k^*, \pi_{-k}^*) = \max_{a_k \in A_k} \left\{ E[RW_k(s_k, a_k, \pi_{-k}^*(s_k))] + \right. \\
\beta \sum_{s'_k \in S_k} T_{s_k,s'_k} (a_k, \pi_{-k}(s_k)) V_k(s'_k, \pi_k^*, \pi_{-k}^*)
\]

(6.79)

where

\[
E[RW_k(s_k, a_k, \pi_{-k}^*(s_k))] = \\
\sum_{(a_k, a_{-k}) \in A} \left[ RW_k(s_k, a_k, a_{-k}) \prod_{j \in K/\{k\}} \pi_j^*(s_j, a_j) \right]
\]

(6.80)

We define the optimal Q-value \( Q_k^*(s_k, a_k) \) of the macro BS \( k \) as the sum of the current reward and the future expected rewards when all the macrocells achieve optimal offloading strategy as follows,

\[
Q_k^*(s_k, a_k) = E[RW_k(s_k, a_k, \pi_{-k}^*(s_k))] + \beta
\]
\[ \sum_{s'_k \in S_k} T_{s_k, s'_k} (a_k, \pi^*_k(s_k)) V_k(s'_k, \pi^*_k(s'_k), \pi^*_{-k}(s'_k)) = \]

\[ E[RW_k(s_k, a_k, \pi^*_k(s_k))]+ \]

\[ \beta \sum_{s'_k \in S_k} T_{s_k, s'_k} (a_k, \pi^*_k(s_k)) \max_{b_k \in A_k} Q^*_k(s'_k, b_k) \]

(6.81)

The optimal Q-value is reached in a recursive way using \((a_k, s_k, s'_k, \pi^*_t(s_k))\), where \(s_k = s'_k\) and \(s'_k = s^{t+1}_k\) are the environment states observed by the macro BS \(k\) at iteration \(t\) and \(t + 1\) respectively. Thus, the general online learning updating rule is defined as,

\[ Q^{t+1}_k(s_k, a_k) = (1 - \alpha^t)Q^t_k(s_k, a_k) + \]

\[ \alpha^t \left\{ \sum_{a_{-k} \in A_{-k}} \left[ RW_k(s_k, a_k, a_{-k}) \prod_{j \in K/\{k\}} \pi^t_j(s_j, a_j) \right] \right\} \]

\[ + \beta \max_{b_k \in A_k} Q^t_k(s'_k, b_k) \]  

(6.82)

where \(\alpha^t \in [0, 1)\) is the learning rate. The Q-value increases when the macrocell \(k\) achieves higher reward by combining the old value and the future expected reward.

**Intuitive Online Learning Mechanism for Traffic Offloading**

From the online learning model described in the previous section, we notice that the reward achieved by each macrocell is a function of the offloading strategies of other macrocells. Accordingly, our proposed model is a multi-agent online learning to benefit from the awareness about other macrocells’ offloading strategies and improve the quality of the selected offloading action. However, the offloading problem is challenging in this context as information about the number of other macrocells and their offloading strategies are unavailable. Moreover, exchanging information about these strategies between the macro BSs requires large computation and causes considerable overhead. Local information like environment state, traffic offloading strategy and the received historical rewards is the only information available. Thus, we present our
intuitive online learning mechanism that is able to surmise the offloading strategies of the other macrocells and exploits this information to reach an optimal offloading strategy. Our mechanism relies on the fact that each macro BS is aware of the system loads, activities and coverage locations of all the small BSs operates under its coverage.

We define \( \mu_t^k(s_k, a_{-k}) = \prod_{j \in K \setminus \{k\}} \pi_j^t(s_j, a_j) \) for macrocell \( k \) in time slot \( t \), to be the intuition factor used to determine \( Q_{k}^{t+1}(s_k, a_k) \) value in the next iteration when all the macrocells select actions \( a_{-k} \) according to their corresponding offloading strategies \( \pi_{-k}^t(s_{-k}) \). The macrocell determines the intuition factor \( \mu \) from the local observations. The intuition idea is derived from the concept that different macrocells follow similar traffic offloading strategies at the same network states. The probability of experiencing certain environment state \( s_{k}^{t+1} \) and achieve certain reward \( RW_k \) is defined as

\[
pt = \pi_t^i(s_k, a_k) \mu_t^i(s_k, a_{-k}).
\]

Let \( v \) denotes the number of iterations between any two consecutive iterations in which cell \( k \) achieves the same reward \( RW_k(s_k, a_k, a_{-k}) \), then \( v \) has independent and identical distribution (i.i.d) with probability \( pt \). Consequently, \( pt = \frac{1}{1+v'} \), where \( v' \) is the mean value of \( v \) and is calculated through observations of of local reward history. Therefore, the intuition factor can be estimated from the local historical information about offloading strategies using linear model as follows,

\[
\mu_k^t(s_k, a_{-k}) = \mu_k'(s_k, a_{-k}) - w_k[\pi_k^t(s_k, a_k) - \pi_k'(s_k, a_k)]
\]

(6.83)

where \( \mu_k'(s_k, a_{-k}) \) and \( \pi_k'(s_k, a_k) \) are the reference points for specific intuition factor and strategy selection probability respectively. These reference points are given from common knowledge provided that each macrocell \( k \) assumes that other macrocells observe its deviation from its reference points by a quantity proportional to the deviation of \( \pi_t^i(s_k, a_k) - \pi_k'(s_k, a_k) \). The macrocells are assumed to update these reference points based on their local observation. At time slot \( t \), the macro BS \( k \) sets \( \mu_k'(s_k, a_{-k}) \)
and \( \pi'_k(s_k, a_k) \) to be \( \mu^t_{i-1}(s_k, a_{-k}) \) and \( \pi^{t-1}_k(s_k, a_k) \) respectively, as a result,

\[
\mu^h_k(s_k, a_{-k}) = \mu^t_{i-1}(s_k, a_{-k}) - w_k[\pi^t_k(s_k, a_k) - \pi^{t-1}_k(s_k, a_k)] \tag{6.84}
\]

The definition in (6.84) comes from the fact that if the macro BS \( k \) changes its strategy, this will induce other macro BSs to perform changes in their offloading strategies in the next time slot. Now, every macro BS can conjecture other BSs strategies using the intuition factor. Therefore, the updating rule in (6.82) is modified in away that macro BS \( k \) updates its Q-value with its own information during the stochastic learning as follows,

\[
Q^{t+1}_k(s_k, a_k) = (1 - \alpha^t)Q^t_k(s_k, a_k) + \alpha^t \left\{ \sum_{a_{-k} \in A_{-k}} \mu^h_k(s_k, a_{-k})RW_k(s_k, a_k, a_{-k}) + \beta \max_{b_k \in A_k} Q^t_k(s'_k, b_k) \right\} \tag{6.85}
\]

The exploration vs exploitation balancing is important in stochastic learning. Exploration aims to try new offloading strategies while exploitation seizes the already explored strategies with positive feedback. To balance exploration vs exploitation and avoid the problem of choosing equally among actions in the exploration process as in the \( \epsilon \)-greedy selection \[206\], actions are selected according to graded functions of their Q-value. For example, the offloading strategy with the highest Q-value has the highest selection probability. We exploit Boltzman distribution \[131\] to perform this selection where the macro BS \( k \) selects an action \( a_k \) in state \( s_k \) at time slot \( t \) with the following probability,

\[
\pi^t_k(s_k, a_k) = \frac{e^{Q^t_k(s_k, a_k)}/\tau}{\sum_{b \in A_k} e^{Q^t_k(s_k, b)}/\tau} \tag{6.86}
\]

where \( \tau \) is a positive parameter called the temperature. High temperature causes the action probabilities to be all nearly equal, while low temperature induces a large difference in selection probabilities for actions differ in their Q-values. The intuition
The algorithm starts by checking the state information at the macrocell including its system load. If the macrocell is overloaded, this triggers the need for offloading traffic to small cells under its coverage. The macrocell activates sleeping small cells for traffic offloading based on their coverage locations. Then, the intuitive learning traffic offloading process is initialized. The Q-values for each strategy in different states are evaluated. When the state is initialized, offloading action is selected. After the action of the selected strategy is applied, the system load of the active small cells (condition C.1) is checked to ensure that the cells are not overloaded and the reward function is evaluated only if the system load is below the threshold. At the end, the Q-value, intuition factor, and strategy are updated accordingly. The process is repeated until the system converges to the best traffic offloading strategy.

**Intuitive online learning mechanism Convergence**

The proposed intuitive online learning algorithm converges w.p.1 if the conditions in **lemma 1** proposed in [131], which establishes the convergence of general online learning process updated by a pseudo-contraction operator $\mathcal{H}$ are satisfied provided that $\mathbb{R}$ be the space for all Q-values.

**Lemma 1.** Assume that the learning rate $\alpha^t$ in (6.85) satisfies the sufficient conditions of Theorem in [131], and the mapping $\mathcal{H}^t : \mathbb{R} \to \mathbb{R}$ meets the following condition: there exists a number $0 < \beta < 1$ and a sequence $\varphi^t \geq 0$ converging to zero w.p. 1, such that $||\mathcal{H}^t Q^t - \mathcal{H}^t Q^*|| \leq \beta ||Q^t - Q^*|| + \varphi^t$ for all $Q^t \in \mathbb{R}$ and $Q^* = E[\mathcal{H}^t Q^*]$, then the iteration defined by,

$$Q^{t+1} = (1 - \alpha^t)Q^t + \alpha^t(\mathcal{H}^t Q^t)$$

converges to $Q^*$ w.p. 1.
Algorithm 6 Intuitive online learning algorithm for traffic offloading and small cell operation

Require: $\pi_k^t(s, a)$, $t$, $SL_k^{th}$ $SL_z(s(t), a(t))$

Ensure: offloading strategy and small cells operation

1: if ($SL_k(s(t), a(t)) > SL_k^{th}$) then
2: Learning initialization
3: Let $t = 0$
4: for each $(s_k \in S_k, a_k \in A_k)$ do
5: initialize offloading strategy $\pi_k^t(s_k, a_k)$;
6: initialize Q-value of the selected strategy $Q_k^t(s_k, a_k)$;
7: initialize intuition factor $\mu_k^t(s_k, a_{-k})$;
8: end for
9: evaluate the state $s_k = s_k^t$
10: while (true) do
11: Select action $a_i$ according to $\pi_k^t(s_k, a_k)$;
12: Check $SL_z(s(t), a(t))$
13: if ($SL_z(s(t), a(t)) < SL_z^{th}$) then
14: $RW_i(s_k, a_k, a_{-k})$ is achieved ;
15: else
16: $RW_k(s_k, a_k, a_{-k}) = 0$
17: end if
18: Update $Q_k^{t+1}(s_k, a_k)$ according to (6.85)
19: Update $\pi_k^{t+1}(s_k, a_k)$ according to (6.86)
20: Update $\mu_k^{t+1}(s_k, a_{-k})$ according to (6.84)
21: $s_k = s_k^{t+1}$
22: $t = t + 1$
23: end while
24: end if
We define the mapping operator $H$ as a mapping on a complete metric space $\mathbb{R} \to \mathbb{R}$, where,

$$H^t Q^t_k(s_k, a_k) = \sum_{a_{-k} \in A_{-k}} \mu^h_k(s_k, a_{-k}) R W_k(s_k, a_k, a_{-k})$$

$$+ \beta \max_{b_k \in A_k} Q^t_k(s'_k, b_k)$$

(6.87)

Now, we proceed to prove the two conditions in Lemma 1 with the mapping operator defined in (6.87).

**Proposition 2.** For $K$-agent stochastic learning,

$$Q^* = E[H^t Q^*]$$

(6.88)

where $Q^* = (Q^*_1, \ldots, Q^*_K)$

**Proof.** As

$$Q^*_k(s_k, a_k) = E[R W_k(s_k, a_k, \pi^*_k(s_k))] +$$

$$\beta \sum_{s'_k \in S_k} T_{s_k, s'_k}(a_k, \pi^*_k(s_k)) \max_{b_k \in A_k} Q^*_k(s'_k, b_k)$$

$$= \sum_{s'_k \in S_k} T_{s_k, s'_k}(a_k, \pi^*_k(s_k)) \left\{ \sum_{a_{-k} \in A_{-k}} R W_k(s_k, a_k, a_{-k}) \right\}$$

$$\prod_{j \in K \setminus \{k\}} \pi^*_j(s_j, a_j) + \beta \max_{b_k \in A_k} Q^*_k(s'_k, b_k)$$

With optimal intuition factor defined in Section IV.B as $\mu^*_k(s_k, a_{-k}) = \prod_{j \in K \setminus \{k\}} \pi^*_j(s_j, a_j)$. Thus,

$$Q^*_k(s_k, a_k) = E[H^t Q^*_k(s_k, a_k)]$$

for all $s_k \in S_k$ and $a_k \in A_k$.

To prove the second condition in Lemma 1, we need to define the distance be-
between any two Q-values $Q$ and $Q' \in \mathbb{R}$ as follows,

$$||Q - Q'|| = \max_{k \in K} \max_{s_k \in S_k} \max_{a_k \in A_k} |Q_k(s_k, a_k) - Q'(s_k, a_k)|$$

(6.89)

**Proposition 3.** $\mathcal{H}^t$ is a contraction mapping operator in stochastic online learning

**Proof.**

$$||\mathcal{H}^t Q - \mathcal{H}^t Q'|| =$$

$$\max_{k \in K} \max_{s_k \in S_k} \max_{a_k \in A_k} |\mathcal{H}^t Q_k(s_k, a_k) - \mathcal{H}^t Q'(s_k, a_k)|$$

Then, we substitute the value of $\mathcal{H}^t Q_k(s_k, a_k)$ from (24),

$$||\mathcal{H}^t Q - \mathcal{H}^t Q'|| =$$

$$\max_{k \in K} \max_{s_k \in S_k} \max_{a_k \in A_k} \left| \sum_{a_{-k} \in A_{-k}} \left[ \mu_k^v(s_k, a_{-k}) - \mu_k^i(s_k, a_{-k}) \right] \right| \times$$

$$RW_k(s_k, a_k, a_{-k}) + \beta \max_{b_k \in A_k} \left[ \max_{b_{k} \in A_k} Q_k(s_k, b_k) - \max_{b_{k} \in A_k} Q'_k(s_k, b_k) \right]$$

$$\leq \max_{k \in K} \max_{s_k \in S_k} \max_{a_k \in A_k} \left| \sum_{a_{-k} \in A_{-k}} \left[ \mu_k^v(s_k, a_{-k}) - \mu_k^i(s_k, a_{-k}) \right] \right| \times$$

$$RW_k(s_k, a_k, a_{-k}) + \beta \left| \max_{b_k \in A_k} Q_k(s_k, b_k) - Q'_k(s_k, b_k) \right|$$

$$\leq \max_{k \in K} \max_{s_k \in S_k} \max_{a_k \in A_k} \left| \sum_{a_{-k} \in A_{-k}} \left[ \mu_k^v(s_k, a_{-k}) - \mu_k^i(s_k, a_{-k}) \right] \right| \times$$

$$RW_k(s_k, a_k, a_{-k}) + \beta ||Q - Q'||$$

We can use the definition of intuition factor and reference points in (6.83) to get,

$$\sum_{a_{-k} \in A_{-k}} \left[ \mu_k^v(s_k, a_{-k}) - \mu_k^i(s_k, a_{-k}) \right] RW_k(s_k, a_k, a_{-k})$$

$$= - \sum_{a_{-k} \in A_{-k}} w_k \left[ \pi_k(s_k, a_k) - \pi'_k(s_k, a_k) \right] RW_k(s_k, a_k, a_{-k})$$

(6.90)
The definition of $\pi_k(s_k, a_k)$ is borrowed from (6.86) with assumption of large $\tau$ as follows,

$$e^{Q_k(s_k, a_k)/\tau} = 1 + \frac{Q_k(s_k, a_k)}{\tau} + \psi\left(\frac{Q_k(s_k, a_k)}{\tau}\right)$$

where $\psi$ is a polynomial of order $O\left(\left(\frac{Q_k(s_k, a_k)}{\tau}\right)^2\right)$. Thus,

$$\sum_{b \in A_k} e^{Q_k(s_k, a_k)/\tau} = a_k + 1 + \sum_{b \in A_k} \left[\frac{Q_k(s_k, b)}{\tau} + \psi\left(\frac{Q_k(s_k, b)}{\tau}\right)\right]$$

It can be proved that,

$$\pi_i(s_k, a_k) = \frac{1}{a_k + 1} + \frac{1}{a_k + 1} + Q_k(s_k, a_k) + \eta\left(\frac{Q_k(s_k, b)}{\tau}\right)_b$$ (6.91)

where $\eta\left(\frac{Q_k(s_k, b)}{\tau}\right)_b$ is a polynomial of order smaller than $O\left(\frac{Q_k(s_k, a_k)}{\tau}\right)$. Similarly,

$$\pi'_k(s_k, a_k) = \frac{1}{a_k + 1} + \frac{1}{a_k + 1} + Q'_k(s_k, a_k) + \eta\left(\frac{Q'_k(s_k, b)}{\tau}\right)_b$$ (6.92)

Substituting (6.92) and (6.91) in (6.90), we get,

$$\sum_{a_{-k} \in A_{-k}} \left[\mu_k(s_k, a_{-k}) - \mu'_k(s_k, a_{-k})\right] RW_k(s_k, a_k, a_{-k}) =$$

$$\sum_{a_{-k} \in A_{-k}} w_k RW_k(s_k, a_k, a_{-k})$$

$$\frac{1}{a_k + 1} \left[Q_k(s_k, a_k) - Q'_k(s_k, a_k)\right] + \eta\left(\frac{Q_k(s_k, b)}{\tau}\right)_b - \eta\left(\frac{Q'_k(s_k, b)}{\tau}\right)_b$$

With large $\tau$ assumption,

$$\left|\sum_{a_{-k} \in A_{-k}} \left[\mu_k(s_k, a_{-k}) - \mu'_k(s_k, a_{-k})\right] RW_k(s_k, a_k, a_{-k})\right|$$

$$\leq \frac{1 - \beta}{a_k + 1} |Q_k(s_k, a_k) - Q'_k(s_k, a_k)|$$
This brings us to the conclusion that,

$$
\|H'Q - H'Q'\| \leq \max_{k \in K} \max_{s_k \in S_k} \max_{a_k \in A_k} 1 - \beta \frac{a_k + 1}{a + 1}
$$

$$
|Q_k(s_k, a_k) - Q'_k(s_k, a_k)| + \beta\|Q - Q'\|
\leq \frac{1 - \beta}{a + 1}\|Q - Q'\| + \beta\|Q - Q'\| \leq \frac{\beta a + 1}{a + 1}\|Q - Q'\|
$$

where \( a = \min_{k \in K} a_k \), therefore,

$$
\frac{\beta a + 1}{a + 1} < 1
$$

This implies that \( H' \) is a contraction mapping operator.

We clearly notice that we applied Lemma 1 to our intuition based learning model in (6.85) and prove its convergence by satisfying the conditions of Lemma 1 through Proposition 1 and Proposition 2 provided that \( \tau \) is large enough.

### 6.3.4 Numerical Results

We simulate our proposed traffic offloading scheme to demonstrate its capability to boost the energy efficiency and maintain QoS for all users in the system. We establish two-tier 5G network, which is composed of 4 macrocells, 7 picocells, and 10 femtocells in a \( 3 \times 3 \) km\(^2\) square area as in Figure 6.22. The macro BSs are located at equal distance apart while the other cells are distributed within macrocells coverage. The radius of the macrocell, picocells and femtocells are \( \sqrt{2}/2 \) km, 0.2 km, and 0.1 km respectively. The entire area is divided into 3600 small locations, where each location accounts for a small area with a resolution of 50 \( \times \) 50 m\(^2\). The channel gains are fixed as \( G_{nl} = d_{nl}^{-\epsilon} \) for all \( n \in K \cup N \) and \( l \in L \), where \( d_{nl} \) is the physical distance between BS \( n \) and the center of location \( l \) and \( \epsilon \) is the path loss factor. Table 6.3 presents the values for other simulation parameters. Each simulation time is supposed to be 40 second to avoid switching small BSs including pico and femto BSs ON and OFF.
The coverage locations are assumed to be heavily loaded with an arrival rate of $\lambda = 6$.

<table>
<thead>
<tr>
<th>Simulation Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td>20 MHz</td>
</tr>
<tr>
<td>Discount factor ($\beta$)</td>
<td>0.9</td>
</tr>
<tr>
<td>Learning rate $\alpha$</td>
<td>0.5</td>
</tr>
<tr>
<td>Noise power</td>
<td>$4 \times 10^{-21}$ W/Hz</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>0.01</td>
</tr>
<tr>
<td>Iteration time $d$</td>
<td>40 sec</td>
</tr>
<tr>
<td>Path Loss model</td>
<td>$PL = 28.3 + 22 \log_{10}(ds)$</td>
</tr>
<tr>
<td>$p_{ct}$ for macro, pico, femto BS respectively</td>
<td>130 W, 15 W, 5 W, 0.1 W</td>
</tr>
<tr>
<td>Pico BS Tx power</td>
<td>23 dBm</td>
</tr>
<tr>
<td>Femto BS Tx power</td>
<td>14 dBm</td>
</tr>
<tr>
<td>Macro BS Tx power</td>
<td>45 dBm</td>
</tr>
</tbody>
</table>

Table 6.3: 5G Hetnets environment simulation parameters

$\lambda_0$ where $\lambda_0 = 0.3$ users/time slot is the arrival rate of macro users. We assume a predefined threshold for the system load $SL^h_z = 0.6$ to balance energy efficiency and QoS tradeoff, while the data rate threshold varies according to the type of BS that the user is associated with.

We compare the performance of our proposed traffic offloading scheme with the schemes proposed in [90], [211] and [212] in addition to the standard, where there
is no offloading mechanism implemented. The work proposed in [90] exploits linear programing to develop a heuristic based offloading algorithm. The scheme proposed in [211] called (JPUA) proposed a mechanism with joint macro BS transmit power control and load-aware user association to reduce the interference impact in a two-tier Hetnets. The work in [212] (GDSCO) dynamically changes the operating states (on and off) of the small BSs, while keeping the macro BS ON with the goal to reduce the network power consumption where users are uniformly distributed in the network. The considered system evaluation is divided into three categories: convergence evaluation, energy efficiency with various network load evaluation, and QoS of the macro users associated with the macro BS and offloaded users associated with active small cells.

Convergence Evaluation

The convergence of the proposed traffic offloading scheme is evaluated in terms of the reduced traffic load and the achieved energy efficiency. The traffic load is found as the average system loads of all cells during the learning process. This evaluation demonstrates the capability of the proposed scheme to maintain the system load of the macrocells at acceptable level as well as the small cells. The achieved total traffic load during the learning process is potted in Figure 6.23. The figure shows that our conditional traffic offloading scheme records the minimum average traffic load compared to the heuristic scheme and the standard. This indicates that the scheme managed to balance the traffic load over the cells that constitute to the network and this is due to the conditional offloading using enhanced online learning that conjectures other macro BSs traffic offloading actions. In addition, we plotted the achieved average energy efficiency of the network against the number of epoch in Figure 6.24. The results presented in this figure reveal the advantage of the proposed scheme as it achieved the maximum energy efficiency. In addition, we notice that the
standard, which deactivates the small cells all the time achieves initial low total traffic load as it operates using fewer cells. The average traffic load and energy efficiency achieved by other schemes are not competent to our scheme. The fast convergence of our scheme is due to the impact of the intuitive feature of the proposed online learning, which is able to conjecture other macrocells offloading strategies and this reduces the overhead of information exchange among base stations.
Energy Efficiency with Various Network Load

In this part, we evaluate the achieved energy efficiency as a function of the average normalized traffic load of the macrocells and the number of the small cells. The energy efficiency comparison between the proposed scheme and other schemes is presented in Figure 6.25 for different normalized network load conditions. The figure shows that our scheme recorded the highest energy efficiency compared to other schemes with performance improvement of 14% over the best competing GDSCO in [212] when the load is high. Moreover, it shows that the total energy efficiency decreases as the network load increases. This is due to the fact that with low load, more base stations can be inactive and this maintains the interference at low levels and reduces energy consumption. Another energy efficiency evaluation is presented in Figure 6.26 for

![Figure 6.25: Average system energy efficiency under different network loads in 5G Hetnets](image)

Figure 6.25: Average system energy efficiency under different network loads in 5G Hetnets

energy efficiency vs variable number of small cells. The figure demonstrates that the proposed scheme outperforms the other schemes regardless the number of small cells involved as it performs learning to exploit the most efficient traffic offloading strategy unlike other schemes. In addition, we can observe that the system energy efficiency is improved with the increasing number of small cells. The reason is that small cells
enhance the energy efficiency as they consume less power than the macrocells when traffic is offloaded to these cells. However, the energy efficiency saturates at certain number of small cells with little improvement as when the network gets dense, the interference between those cells enforces higher transmission power and consequently more power consumption.

**Users QoS Satisfaction**

In this part, we investigate the experience of the macro users as a result of their traffic offloading to small cells. In addition, this evaluation accounts for QoS of the users, which are not offloaded and remain associated with the macro BS. The CDF plots in Figure 6.27 and Figure 6.28 present the data rate for macro users associated with the macro BS and the macro users offloaded and associate with small cells respectively.

For the case when the network is heavily loaded (i.e. the normalized network load is above 0.8). The minimum data rate requirements considered in this evaluation is 2 Mbps for the macro user. We can notice that the proposed scheme managed to keep the outage (when the user expected data rate is lower than the required rate) at less than 5 % for macro users in both association cases, where other schemes including
heuristic based in [90], JPUA in [211], the GDSCO in [212] recorded outage ratio of 25\%, 16\%, and 12\% respectively for the macro users which remain associated with the macro BS. On the other hand, the outage ratio for the offloaded users for the heuristic, JPUA and GDSCO are 29\%, 18\%, and 11\% respectively as in Figure 8. This evaluation indicates that the conditional offloading proposed scheme with improved online learning smoothes the offloading process and boost the performance not only at the level of energy efficiency but also at the level of QoS for the affected
users. Moreover, the evaluation confirms that the system load has a large impact not only on the energy efficiency, but also on user QoS.

6.4 Sophisticated Online Learning Scheme for Green Resource Allocation in 5G Heterogeneous Cloud Radio Access Networks

We propose a green resource allocation scheme for the downlink of H-CRANs between RRHs and their associated UEs that aims at maximizing energy efficiency while satisfying the UE QoS requirements and mitigate the inter-tier interference. Allocation includes RB assignment and power allocation [25]. The scheme is developed using centralized and decentralized enhanced online learning approaches. The centralized allocation is performed at a designated controller integrated with the Baseband Unit (BBU), thanks to the macro BS, which is interfaced to the BBU pool for coordinating the inter-tier interference and exchange resource allocation control signals. This alleviates the capacity and time delay constraints on the fronthaul and supports the burst traffic efficiently. The decentralized approach relies on macro BSs cooperation and local information analysis to let the macro BSs make decisions to reach the appropriate allocation. The proposed learning methodology functions in conjunction with an enhanced spectrum partitioning that classifies the available RBs according to the users traffic priority and location. For example, UEs with high QoS traffic and located at the cell edge consume RBs that belongs to RBs set that is dedicated for high priority users. Moreover, compact state representation and Q-value approximation are employed to handle the curse of dimensionality during the learning process due to the large state space. To the best of our knowledge, there is no solutions in the literature for resource allocation in H-CRANs with consideration of energy efficiency using machine learning techniques. One of the advantages of online learning is that it is model-free, which facilitates its usage in dynamic heterogeneous networks.
The key contributions of this part are summarized as follows,

- We develop an enhanced spectrum partitioning model that divide the available spectrum into two RBs sets, where each set is dedicated for certain group of users according to their location and QoS requirements.

- A centralized joint RBs and power allocation scheme for RRH and their associated UEs is proposed, which relies on a single controller integrated to the BBU. This controller acquires the state information and selects the most appropriate actions that enhance energy efficiency and guarantee QoS requirements for UEs from different tiers.

- To eliminate the single point of failure possibility and exploit the environment awareness capability of the macro BSs, we propose a decentralized resource allocation scheme, where the macro BSs utilize their local information and the BBU pool data to select the most appropriate resource allocation actions.

- The proposed online learning model incorporates compact state representation and Q-value approximation to handle the curse of dimensionality, and augment the algorithm convergence.

- We implement both resource allocation schemes using software defined radio testbed to demonstrate the schemes capability to perform efficient resource allocation in terms of energy efficiency, BS capacity and the encountered BER.

### 6.4.1 System Model and Problem Formulation

The architecture of the considered H-CRANs system is presented in Figure 6.29 in which a two-tier cellular network is shown, the macro tier and the small cells tier. The macro BSs denoted by $u \in U$ are underlaid with small cells (RRHs) denoted by $s \in S$, and the BBU pool performs all baseband processing functionalities. The
resource allocation process for RRHs is executed at the controller integrated with the BBU pool in the centralized approach and it is achieved through cooperation between macro BSs in the distributed approach. The inter-RRH interference between RRHs is jointly coordinated at the BBU pool. The network users in this model are classified into two categories: MUEs denoted by index $m \in M$, which are the users associated with the macro BS, while RUEs with index $n$ are the users associated with RRH. The backhaul interface is utilized to link the macro BSs with the BBU for control exchange. All the RRHs are connected to the BBU pool by the fronthaul links to facilitates data processing, transmissions, and cloud computing. The system RBs are denoted by $k \in K$ with total bandwidth $B$. To enhance the spectral efficiency and mitigate the inter-tier interference, the available spectrum is partitioned into two sets of RBs: the first set denoted by $\Gamma_1$ incorporates RBs dedicated to RUEs with high QoS requirements or located at their corresponding cell edge. The second set of RBs denoted by $\Gamma_2$ comprises RBs to be shared between RUEs with low QoS or located at the cell center.
and MUEs [213]. This specific partitioning improves the spectrum efficiency compared to the traditional schemes and avoids random exploration of machine learning in RBs allocation, which expedites the learning speed of convergence. The QoS requirement is defined as the minimum data rate of the RUEs. The sets $\mathbf{N} = \{1, 2, \ldots, N\}$ and $\mathbf{q} = \{N + 1, \ldots, N + q\}$ represent the users associated with RRH and occupying the sets $\Gamma_1$ and $\Gamma_2$ respectively. Let us assume that $a^k_n$ is the RB allocation indicator, $a^k_n = 1$, if the RB $k$ is allocated to RUE $n$ associated with RRH $s$ and $0$ otherwise.

For an RUE $n$ served by RRH $s$ on RB $k$, the received signal can be written as,

$$ \eta^k_{s,n} = P^k_{s,n}g^k_{s,n}\rho^k_n + I^k_n + N_0 $$

(6.93)

where $P^k_{s,n}$ is the transmission power of RRH $s$ allocated to RUE $n$ on RB $k$, $\rho^k_n$ is the transmitted information symbol for RUE $n$ on RB $k$, $N_0$ is the noise power, and $g^k_{s,n} = H^k_{s,n}l_{s,n}$ is the channel gain between RRH $s$ and RUE $n$ on RB $k$. Note that $H^k_{s,n}$ and $l_{s,n}$ refer to the fast fading coefficient and the path loss between RRH $s$ and RUE $n$ respectively. The encountered interference denoted by $I^k_n$ is calculated according to the category of the RUE $n$ determined by the set of RBs it accesses. Therefore, $I^k_n$ is calculated as follows,

$$ I^k_n = \begin{cases} 
\sum_{i=1, i\neq n}^{N} (\sum_{r\in \mathbf{S}, r\neq s} P^k_{r,i}g^k_{r,i}) \ k \in \Gamma_1 \\
\sum_{m=1}^{M} P^k_{u,m}g^k_{u,m} + \sum_{i=N, i\neq n}^{N+q} (\sum_{r\in \mathbf{S}, r\neq s} P^k_{r,i}g^k_{r,i}) \ k \in \Gamma_2 
\end{cases} $$

(6.94)

where $P^k_{u,m}$ is the transmission power of macro BS $u$ allocated to MUE $m$ on RB $k$, and $g^k_{u,m} = H^k_{u,m}l_{u,m}$ is the channel gain between macro BS $u$ and MUE $m$ on RB $k$. Note that the index $i$ represents the other RUEs served by other RRH $r$. The signal
to interference and noise ratio (SINR) for \( n \)th RUE occupying \( k \)th RB is given by,

\[
\gamma_{s,n}^k = \frac{P_{s,n}^k g_{s,n}^k}{I_n^k + N_0}
\] (6.95)

Similarly, the received signal by an MUE \( m \) served by macro BS \( u \) is written as follows,

\[
\eta_{u,m}^k = P_{u,m}^k g_{u,m}^k \rho_{m}^k + I_m^k + N_0
\] (6.96)

where \( \rho_{m}^k \) is the received information symbol and \( I_m^k \) is the encountered interference by the MUE \( m \) allocated to RB \( k \) and it is evaluated as follows,

\[
I_m^k = \sum_{i=N}^{N+q} \left( \sum_{s \in S} P_{s,i}^k g_{s,i}^k \right)
\] (6.97)

The SINR achieved by MUE \( m \) utilizing RB \( k \) and associated with macro BS \( u \) is found as follows,

\[
\gamma_{u,m}^k = \frac{P_{u,m}^k g_{u,m}^k}{I_m^k + N_0}
\] (6.98)

The capacity for all RUEs associated with RRH \( s \) is expressed as follows,

\[
C_s = \sum_{n=1}^{N+q} \sum_{k=1}^{K} a_{s,n}^k B \log_2(1 + \gamma_{s,n}^k)
\] (6.99)

Similarly, the capacity of all MUEs associated with the macro BS \( u \) is found as follows,

\[
C_u = \sum_{m=1}^{M} \sum_{k=1}^{K} a_{u,m}^k B \log_2(1 + \gamma_{u,m}^k)
\] (6.100)

where \( a_{u,m}^k \) is the RB resource allocation indicator similar to \( a_{s,n}^k \). The power consumption of the H-CRANs system is composed of the power consumed by RRHs and their related links and the power consumed by macro BSs [214]. The power
consumption of RRH $s$ is evaluated as follows,

$$PC_s = P_{ct} + P_f + P_e + \sum_{n=1}^{N+q} \sum_{k=1}^{K} P_{s,n}^k a_{s,n}^k$$  \hspace{1cm} (6.101)$$

where $P_{ct}$ is the circuit power, $P_f$ is the power consumption of the fronthaul links, and $P_e$ is the power consumed in signal exchange.

The energy efficiency ($EE$) of the considered H-CRAN system is defined as follows,

$$EE = \frac{S \times C_s + C_u}{S \times PC_s + PC_u}$$  \hspace{1cm} (6.102)$$

where $PC_u$ is the power consumed by the macro tier. However, H-CRANs is established based on dense RRH deployment (i.e. the number of RRHs is much larger than macro BSs). The inter-tier interference from RRHs to MUEs remains constant when the density of RRHs is sufficiently high. Thus, the downlink spectral efficiency and energy efficiency performances for the macro BSs can be assumed to be stable [97][214]. Therefore, with sufficiently high $S$, the energy efficiency can be defined as,

$$EE = \frac{C_s}{PC_s}$$  \hspace{1cm} (6.103)$$

The allocation of RBs $a_{s,n}^k$ and transmission power $P_{s,n}^k$ with objective of EE maximization in the downlink of H-CRANs subjected to QoS requirements for all RUEs and MUEs and the fronthaul links constraints is formulated as follows,

$$\max_{P_{k,n}^k, a_{s,n}^k} EE \hspace{1cm} s.t \hspace{1cm} (6.104)$$

C1: $\sum_{k=1}^{K} a_{s,n}^k \leq z_n$, $\forall s \in S$

C2: $a_{s,n}^k \in \{0, 1\}$ $\forall s \in S$, $\forall k \in K$

C3: $\sum_{k=1}^{K} a_{s,n}^k B \log_2(1 + \gamma_{s,n}^k) \geq \theta$, $\forall n \in N$
The constraint $C_1$ limits the number of allocated RBs to RUE $n$ to $z_n$ RBs, which prevents the cloud from greedily allocating all the available RBs to its RUEs in order to leave some RBs for the other network tiers. $C_2$ indicates that $a_{s,n}^k$ is a binary variable. The capacity constraints $C_3$ and $C_4$ ensure that the achieved capacity for RUE with high QoS requirements and low QoS requirements RUE is above the thresholds $\theta$ and $\theta^*$ respectively. $C_5$ is the constraint to guarantee QoS for the MUE $m$. The transmit power on an unallocated RB is enforced to be zero and the maximum allowed power for each RRH is indicated in $C_6$ as $P_{s,max}$. Finally, $C_7$ is a fronthaul constraint, which limits the number of baseband signals transmitted on the fronthaul link between the cloud and RRH $s$. $\chi_{max}$ is defined as the maximum number for the transmitted signals on the fronthaul link between the cloud and RRH $s$ and $\phi$ is a step function that takes value 1 if $\sum_{k=1}^{K} P_{s,n}^k > 0$ and 0 otherwise. $P_{s,n}^k < 0$ indicates that RRH $s$ does not serve RUE $n$ and that the fronthaul link between the cloud and RRH $s$ does not carry the baseband signal for RUE $n$. On the other hand, if there is at least one RB $k$ such that $P_{s,n}^k > 0$ then RRH $s$ serves RUE $n$ and the fronthaul link between the cloud and RRH $s$ is active. The resource allocation problem in (6.104) is tackled using enhanced online learning to reach ultimate allocation of RB and transmission power to guarantee capacity in the designated sub-bands and inter-tier interference mitigation.
6.4.2 Online Learning Model for Resource Allocation

It is challenging to determine the exact state transition model by applying a model-based dynamic programing algorithm for the following reasons: The spectrum partitioning based RB allocation considered in this work is not only location dependent but also considers QoS requirements. Moreover, it is not trivial to list all the states and action pairs to migrate from one state to another, and it is not practical to pre-define the state transition model in problem solving. For these reasons, we choose online Q-learning [205] for resource allocation in the H-CRANs specified model. The basic concept of online learning is described as follows, when the network is in state $x^t$ at time step $t$, a finite number of possible actions $y^t$, which are elements of the action space $Y$ can be selected. As a result, a reward is received, which is the network feedback for the action selected at state $x^t$. In this section, we describe the learning model employed to achieve resource allocation that maximizes the network energy efficiency. We suppose that the resource allocation action is chosen by the network controller in the centralized approach and it is a result of the cooperative decision made by the macro BSs in the distributed one within a slotted time step $t$. The considered learning model for resource allocation is defined as $\zeta = (N, q, \gamma^k_{s,n}, \gamma^k_{u,m}, a^k_{s,n}, P^k_{s,n}, EE)$.

The online learning parameters are defined as follows:

- **State:** the environment state at certain time step $t$ is defined as, $x^t_n=(n, RUE$ location, $\theta, \theta^*, \delta, \gamma^k_{s,n}, \gamma^k_{u,m})$. The state information is acquired from the BBU and from the macro BSs that are assumed to be aware of the small cells operate under their coverage.

- **Action:** the action $y^t_n = (a^k_{s,n}, P^k_{s,n})$ is defined as the allocation of RB and transmission power, RRH allocated to its associated RUEs.
• Reward: the reward function is the energy efficiency and is defined as,

$$R(x, y) = EE(x, y)$$  \hspace{1cm} (6.105)$$

This indicates that the reward is achieved if the conditions in C1 to C7 are satisfied.

• Transition Function: for a given resource allocation strategy $\pi \in \Pi$, the state transition probability is defined as follows,

$$T(x, y, x') = Pr(x(t + 1) = x'| t = x, y^t = y)$$  \hspace{1cm} (6.106)$$

Basically, the strategy $\pi$ is defined as the probability of selection of action $y$ at state $x$.

The optimal Q-value of the online learning model is defined as the current expected reward plus a future discounted reward as follows,

$$Q^*(x, y) = E[EE(x, y)] + \beta \sum_{x' \in X} T_{x,x'}(y) \max_{y' \in Y} Q^*(x', y')$$  \hspace{1cm} (6.107)$$

where $T_{x,x'}(.)$ is the state transition probability and $\beta$ is the discount factor. The optimal Q-value $Q^*_n(x, y)$ is learned by updating the Q-value function on the transition from state $x$ to state $x'$ under the action $y$ in time step $t$ as follows,

$$Q^{t+1}(x, y) = (1 - \alpha^t)Q^t(x, y) + \alpha^t[EE(x, y)$$

$$+ \beta \max_{y' \in Y} Q^t(x', y')]$$  \hspace{1cm} (6.108)$$

where $\alpha^t \in (0, 1]$ is the learning rate. The initial Q-value for all $(x, y)$ is arbitrary. The considered online learning model here is a stochastic approximation method.
that solves the Bellman’s optimality equation associated with the DTMDP. Online learning does not require explicit state transition probability model and it converges with probability one to an optimal solution if $\sum_{t=1}^{\infty} \alpha^t$ is infinite, $\sum_{t=1}^{\infty} (\alpha^t)^2$ is finite, and all state action pairs are visited infinitely often \[131\]. Balancing exploration and exploitation is an essential issue in the stochastic learning process. Exploration aims to try new allocation strategies so it does not only apply the strategies it already knows to be good but also explore new ones. Exploitation is the process of using well-established strategies. The most common technique to achieve this balance is to use the $\epsilon$-greedy selection \[206\]. However, this approach selects equally among the available actions i.e. (the worst action is likely to be chosen as the best one). In order to overcome this drawback, the action selection probabilities are varied as a graded function of the Q-value. The best power level is given the highest selection probability, while all other levels are ranked according to their Q-values. The learning algorithm exploits Boltzmann probability distribution \[215\] to determine the probability of the resource allocation action that fulfills the energy efficiency maximization constraints in C1 to C7. Thus, the action $y$ in state $x$ is selected at $t$ with the following probability,

$$
\pi^t_n(x,y) = \frac{e^{Q^t(x,y)/\tau}}{\sum_{y' \in Y} e^{Q^t(x,y')/\tau}}
$$

(6.109)

where $\tau$ is a positive integer that controls the selection probability. With high value of $\tau$, the action probabilities become nearly equal. However, low value of $\tau$ causes big difference in selection probabilities for actions with different Q-values. One issue to report is that the 5G H-CRANs system has a large space. Therefore, the curse of dimensionality increases the required computations and makes it unfeasible to use the typical online learning methodology to maintain the Q-value for each state/action pair, which slows the system convergence.
6.4.3 Centralized Approximated Online Learning Resource Allocation Scheme

In this approach, the resource allocation process is performed at a dedicated controller that is integrated with the BBU pool and the macro BSs act as brokers between the controller and the RRHs for control exchange with the goal of inter-tier interference mitigation and energy efficiency maximization. Resource allocation and data processing signals from the controller to the macro BSs are sent through X1 and S1 interfaces respectively, which are obtained from definitions of the 3GPP standards. All the entities in the network including RUEs, MUEs, RRHs report the channel state information and RUEs location to the macro BS that they operate under its coverage in a hierarchical manner. The channel state information includes path loss and channel gains from the serving RRH and the macro BS to the RUEs and MUEs. All the macro BSs provide this information to the controller through the control exchange interface. The controller exploit the reported information and QoS requirements to select the proper RB and transmission power using online learning. The allocation decision made by the controller is sent to the RRH through the macro BS. Note that SINR is the state information exploited in action selection and it is determined according to the channel information reported and the allocated transmission power.

The computational complexity of the system increases along with the size of the states and action spaces. The simple look-up table where separate Q-value is maintained for each state/action pair is not feasible in large space with massive number of states like our system. Therefore, we propose a brief representation for the Q-values in which they are approximated as a function of much smaller set of variables to account for the curse of dimensionality. The brief representation of Q-value focuses on a countable state space $X^*$ using the function $Q' : X^* \times Y$, which is referred as a function approximator. The parameter vector $\xi = \{\xi_z\}^Z_{z=1}$ is adopted to approximate the Q-value by minimizing the metric of difference between $Q^*(x, y)$ and $Q'(x, y, \xi)$
for all \((x, y) \in X^* \times Y\). Thus, the approximated \(Q'\) value is formalized as follows,

\[
Q'(x, y, \xi) = \sum_{z=1}^{Z} \xi_z \psi_z(x, y) = \xi \psi^T(x, y)
\] (6.110)

where \(T\) denotes the transpose operator and the vector \(\psi(x, y) = [\psi_z(x, y)]_{z=1}^{Z}\) with a scalar function \(\psi_z(x, y)\) defined as the basis function (BF) over \(X^* \times Y\), and \(\xi_z (z = 1, \ldots, Z)\) are the associated weights. A gradient function \(\psi(x, y)\), which is a vector of partial derivative with respect to the elements of \(\xi^t\), is used to combine the typical online learning model defined in (16) with the linearly parametrized approximated online learning proposed.

The Q-value update rule in (6.108) is reconstructed to include the parameter vector updates as follows,

\[
\xi^{t+1} \psi^T(x, y) = \{(1 - \alpha^t)\xi^t \psi^T(x, y) + \\
\alpha^t \left[ EE(x, y) + \beta \max_{y' \in Y} \xi^t \psi^T(x, y') \right]\}\psi(x, y)
\] (6.111)

The probability of selecting certain action presented in (6.109) is updated with the Q-value approximation as follows,

\[
\pi^t(x, y) = \frac{e^{\xi^t \psi^T(x, y)/\tau}}{\sum_{y' \in Y} e^{\xi^t \psi^T(x, y')/\tau}}
\] (6.112)

The online learning process with approximated Q-value is illustrated in Algorithm 7.

The algorithm takes the QoS requirements for RUEs and MUEs as input to check the quality of the strategies selected and they are compared to the capacity achieved by different BSs. The algorithm selects action strategies according to (6.112). If the conditions C1 to C7 are satisfied, then, the reward is achieved. Finally, the Q-value and resource allocation strategy are updated according to (6.111) and (6.112) respectively, and the new state is observed.
Algorithm 7 Centralized approximated online learning algorithm for resource allocation in 5G H-CRANs

Require: $\pi^t(x,y)$, $t = 1, \delta, \theta^*, \theta^*, \gamma_{s,n}, \gamma_{u,m}$

Ensure: RB and $P_{s,n}$ allocation for RUE

1: initialization of Learning
2: for each $(x, y \in Y')$ do
3: initialize resource allocation strategy $\pi^t(x,y)$;
4: initialize approximated Q-value $\xi^t \psi^T(x,y)$;
5: end for
6: while (true) do
7: evaluate the state $x = x^t$
8: Select action $y$ according to $\pi^t(x,y)$ in (6.112);
9: if (C1 to C7 are satisfied ) then
10: $R(x,y)$ is achieved
11: else
12: $R(x,y) = 0$
13: end if
14: Update $\xi^t \psi^T(x,y)$ according to (6.111)
15: Update $\pi^t(x,y)$ according to (6.112)
16: $x = x^{t+1}$
17: $t = t + 1$
18: end while

To demonstrate Algorithm 7 convergence, we found the necessary conditions for convergence of the proposed approximated learning resource allocation. To start the proof, we introduce the following definition and assumptions [216] [217].

Definition 2. Let $\Psi = E[\psi^T(x,y)\psi(x,y)]$. For the parameter vector $\xi$ and a particular network state $x \in X^*$, we define a vector $\psi(x,\xi) = [\psi_z(x,y)]$ for $z = 1 \rightarrow Z$ where $y \in \mathbb{N} = \{y = \arg \max_{y' \in Y} \xi \psi^T(x,y')\}$ is the set of optimal joint resource allocation actions for $x$. We define the following a $\xi$-dependent matrix:

$$\Psi' = E[\psi^T(x,\xi)\psi(x,\xi)] (6.113)$$

Assumption 4. The basis functions $\psi_z(x,y)$ are linearly independent for all $(x,y)$ and all the properties of $Q^t(x,y)$ in previous discussion are applicable to the dot product for the vectors $\xi^t \psi^T(x,y)$. 
Assumption 5. For every \(z = (1, 2, \ldots, Z)\), \(\psi_z(x, y)\) is bounded, which means \(E\{\psi_z^2(x, y)\} < \infty\) and the reward function satisfies \(E\{R^2(x, y)\} < \infty\).

Assumption 6. The learning rate satisfies \(\sum_{t=1}^{\infty} \alpha^t = \infty\) and \(\sum_{t=1}^{\infty} (\alpha^t)^2 < \infty\).

Proposition 4. With the assumptions 4-6 and Definition 2, the approximated online learning algorithm converges with probability (w.p) 1, if

\[\Psi^t > \Psi, \quad \forall \xi\] (6.14)

Proof. The proof of convergence requires finding stable fixed points of the ordinary differential equations (ODE) associated with the update rule in (6.111), which can be written as.

\[\xi^t = E[EE(x, y) + \beta \xi^t \psi^T(x', \xi^t) - \xi^t \psi^T(x, y)]\psi(x, y)\] (6.115)

where \(\xi^t = \frac{\partial \xi}{\partial t}\) as \(\alpha \to 0\). We define two trajectories of the ODE \(\xi_1^t\) and \(\xi_2^t\) that have different initial conditions and satisfies \(\xi_0^t = \xi_1^t - \xi_2^t\). Then, we have

\[
\frac{\partial \|\xi_0^t\|^2}{\partial t} = 2(\xi_1^t - \xi_2^t)(\xi_0^t)^T = 2\beta E[\xi_1^t \psi^T(x', \xi_1^t)\psi(x, y)(\xi_0^t)^T - \xi_2^t \psi^T(x', \xi_2^t)\psi(x, y)(\xi_0^t)^T - 2\xi_0^t \psi(\xi_0^t)^T]
\] (6.116)

From the definition of \(\psi(x, \xi)\) in Definition 2, we can deduce the following two inequalities,

\[\xi_1^t \psi^T(x', \xi_1^t) \geq \xi_1^t \psi^T(x', \xi_2^t)\] (6.117)

\[\xi_2^t \psi^T(x', \xi_2^t) \geq \xi_2^t \psi^T(x', \xi_1^t)\] (6.118)

As the expectation \(E\) in (6.116) is taken over different states and different actions, we can define two sets \(\Lambda_+ = \{(x, y) \in X \times Y | \xi_0^t \psi^T(x, y) > 0\}\) and \(\Lambda_- \in X \times Y - \Lambda_+\).
If we combine (6.117) and (6.118) in (6.116), we get,

\[
\frac{\partial \|\xi_t^0\|^2}{\partial t} \geq 2\beta (E[\xi_t^0 \psi^T(x', \xi_t^1)]\psi(x, y)(\xi_t^0)^T|\Lambda_+]
\]

\[
+ E[\xi_t^0 \psi^T(x', \xi_t^1)]\psi(x, y)(\xi_t^0)^T|\Lambda_-) - 2\xi_t^0 \Psi(\xi_t^0)^T
\]

(6.119)

After the application of Holder’s inequality to the expectation in (6.119), we get,

\[
\frac{\partial \|\xi_t^0\|^2}{\partial t} \geq 2\beta \left( \sqrt{E[(\xi_t^0 \psi^T(x', \xi_t^1))^2|\Lambda_+]} \times \sqrt{E[(\psi(x, y)(\xi_t^0)^2|\Lambda_+]} + \sqrt{E[(\xi_t^0 \psi^T(x', \xi_t^1))^2|\Lambda_-]}
\]

\[
\times \sqrt{E[(\psi(x, y)(\xi_t^0)^2|\Lambda_-)]} - 2\xi_t^0 \Psi(\xi_t^0)^T
\]

\[
\geq 2\beta \left( \sqrt{E[(\xi_t^0 \psi^T(x', \xi_t^1))^2|\Lambda_+]} \times \sqrt{E[(\psi(x, y)(\xi_t^0)^2|\Lambda_+]} + \sqrt{E[(\xi_t^0 \psi^T(x', \xi_t^1))^2|\Lambda_-]}
\]

\[
\times \sqrt{E[(\psi(x, y)(\xi_t^0)^2|\Lambda_-)]} - 2\xi_t^0 \Psi(\xi_t^0)^T
\]

(6.120)

If we apply the definition of \(\Psi'\) in Definition 2, we get,

\[
\geq 2\beta \sqrt{\max [\xi_t^0 \Psi'(\xi_t^1)^T, \xi_t^0 \Psi'(\xi_t^1)^T]}
\]

\[
\times \sqrt{E[(\psi(x, y)(\xi_t^0)^2|\Lambda_+]} + \sqrt{E[(\psi(x, y)(\xi_t^0)^2|\Lambda_-)]} - 2\xi_t^0 \Psi(\xi_t^0)^T
\]

(6.120)

According to the condition in (6.114), we can state that,

\[
\frac{\partial \|\xi_t^0\|^2}{\partial t} \geq 2\beta \xi_t^0 \Psi(\xi_t^0)^T - 2\xi_t^0 \Psi(\xi_t^0)^T = (2\beta - 2)\xi_t^0 \Psi(\xi_t^0)^T < 0
\]

(6.121)

which means that \(\xi_t^0\) converges to the origin and this confirms that there exists a stable point of the ODE in (24). Thus, the proposed online learning with Q approximation
converges w.p 1.

Consequently, the stable point $\xi^*$ of the ODE in (6.115) indicates that,

$$0 = E[EE(x, y) + \beta \xi^t \psi^T(x', \xi^t) - \xi^t \psi^T(x, y)]$$

(6.122)

and $\xi^*$ can be found as follows,

$$\xi^* = E[EE(x, y) + \beta \xi^* \psi^T(x', \xi^*)] \Psi^{-1}$$

(6.123)

As a result, the approximated online Q-function is stated as follows,

$$Q'(x, y, \xi^*) = \xi^* \psi(x, y)$$

(6.124)

6.4.4 Decentralized Multi-Agent Online Learning Resource Allocation Scheme

Even with the Q-function approximation using compact representation, the number of actions can grow exponentially with the number of cells deployed in the network. Thus, to have a centralized controller for resource allocation is not the best practice. As the macro BSs and can locally manage their operations, there is a possibility to deploy a decentralized resource allocation scheme using online learning that functions through the macro BSs. All the macro BSs can learn in a cooperative manner how to make local decisions for resource allocation for the RUEs associated with RRHs. In this way, the resource allocation task is achieved using multi-agent online learning, where macro BSs represent the agents. The macro BSs estimate the SINR according to the channel state information reported by the network nodes and the power allocated, and execute the online learning algorithm for resource allocation.

We assume that the macro BSs learn in a team Markov game defined as $GE =$
\{U, X', Y, T, R\} with the common goal to find a joint resource allocation strategy \(\pi\) that mitigates the inter-tier interference and maximizes energy efficiency, where \(U\) is the set of the macro BSs. The optimal Q-value \(Q^*(x, y)\) for all \((x, y) \in X' \times Y\), defines the optimal resource allocation strategy and capture the markovian game structure.

For each network state \(x \in X'\), the action in the team Markov game is generated by the \(U\) independent macro BSs in a decentralized fashion. The decided action \(y\) at state \(x\) is considered optimal if \(Q^*(x, y) \geq Q^*(x, y')\) for all \(y' \in Y\). The macro BSs are assumed to learn using the compact representation model stated in (6.111).

Consequently, we can deduce the following proposition.

**Proposition 5.** For the Markovian game \(GE\), the decentralized multi-agent online learning algorithm converges w.p. 1 if the condition in **Proposition 2** holds.

**Proof.** If we consider each macro BS as a single controller that follows a stationary resource allocation strategy, then the Markovian game is a DTMDP. As a result, the proof follows the same procedure of the proof of **Proposition 2**. \(\square\)

To proceed with the multi-agent learning, the following assumptions are made,

**Assumption 7.** The resource allocation strategy of different macro BSs do not alter significantly in similar network states.

**Assumption 8.** The initial network state \(x^t\) evolves following Harris recurrent Markov chain [219].

According to **Assumption 7**, each macro BS can conjecture the allocation strategy of other macro BSs without explicit cooperation through the use of historical knowledge if it encounters the same network state. The similarity between network states can be measured in terms of Hamming distance [220] denoted by \(DH(x, x')\).

We define the historical knowledge up to time step \(t\) using \(\sigma\)-algebra as follows,

\[
F(t) = \sigma(\{x(b), y(b)\}_{b=1}^{t}, \{R(x(b), y(b))\}_{b=1}^{t-1})
\]

(6.125)
where the information of each experienced network state \(x(b)\), each performed joint action \(y(b)\) and network energy efficiency (the reward) \(R(x(b), y(b))\) can be obtained from the BBU. At each time step \(t\), each macro BS checks the Hamming distance between the current state \(x(t)\) and state \(x(b)\) in \(F(t)\). Then, it creates a sample set \(X_F(x^t, F(t))\), which includes \(F\) most recent observations from \(F(t)\) that has a minimum value of \(\sum_{f=1}^{F} DH(x^t, x^b)\). Now, the common reward \(R_c(x, y)\) that all macro BSs receive after they perform a joint resource allocation action, which is defined as \(y \in Y\), is set to 1 if \(y = \arg\max_{y \in Y} Q'(x, y', \xi^*)\) and 0 otherwise. Moreover, we define \(Y_u\) for each macro BS as the set of joint actions that gives the reward 1 in state \(x^t\). The decentralized multi-agent learning for resource allocation process in H-CRAN is illustrated in Algorithm 8.

Note that \(\mu\) and \(\nu\) are two integers that satisfy \(1 \leq \mu \leq F \leq \nu\). The algorithm states that when \(t < \nu\), the resource allocation process according to the probability in (6.111) is engaged to determine the resource allocation strategy. From \(t = \nu + 1\), each macro BS selects \(\mu\) records of \(Y_F(X_F(x^t, F(t)))\) from the \(F\) joint actions for \(X_F(x^t, F(t))\). If the following conditions C8 and C9 are met, the macro BS \(u\) selects the allocation action \(y_u(b^*)\), where \(y_u(b^*) = \max_b\{b | y(b) \in Y_F(X_F(x^t, F(t))) \cap Y_u(x^t)\}\).

- **C8:** there exists allocation action \(y = (y_u, y_{-u}) \in Y_u(x^t)\) such that \(y'_{-u} = y_{-u} \forall y' = (y'_u, y'_{-u}) \in Y_F(X_F(x^t, F(t)))\)

- **C9:** there exists at least one action \(y \in Y_F(X_F(x^t, F(t))) \cap Y_u(x^t)\)

However, if C8 and C9 are not met, macro BS \(u\) will select an action from \(Y'_u(x^t) = \{y_u | y_u = \arg\max_{y_u} R'_c(x^t, y_u)\}\), where

\[
R'_c(x^t, y_u) = \sum_{y_{-u}} R_c(x^t, y) \frac{T'_u(x^t, y_{-u})}{\nu}
\]  

(6.126)

\(\nu\) are cases randomly drawn from \(F\) most recent actions, \(T'_u(x^t, y_{-u})\) is the number of times the conjectured action of other macro BS \(y_{-u}\) in state \(x^t\) is performed. The
Algorithm 8 Decentralized online learning algorithm for resource allocation in 5G H-CRANs

Require: $\pi^t(x, y)$, $t = 1, \delta, \theta^*, \theta, \gamma^k_{x, y}, \gamma^k_{u, m}$ and $1 \leq \mu \leq F \leq \nu$

Ensure: RB, $P^k_{s,n}$ allocation for RUEs

1: initialization of Learning
2: for each $(x, y \in Y')$ do
3: initialize resource allocation strategy $\pi^t(x, y)$;
4: initialize approximated Q-value $\xi^t\psi^T(x, y)$;
5: end for
6: while (true) do
7: evaluate the state $x = x^t$
8: if ($t < \nu + 1$) then
9: Select action $y$ according to $\pi^t(x, y)$ in (6.112);
10: if (C1 to C7 are satisfied) then
11: $R(x, y)$ is achieved
12: else
13: $R(x, y) = 0$
14: end if
15: else
16: Update $Y_u(x^t) = \{y|R_c(x, y) = 1\}$ for $x^t$
17: Randomly select $Y_F(X_F(x^t, F(t)))$ out of $F$ joint actions associated with $X_F(x^t, F(t))$
18: Calculate $R'_c(x, y)_{j}$ according to (6.126) and populate $Y'_u(x^t)$
19: if (C8 and C9 hold) then
20: select the action from $Y_F(X_F(x^t, F(t))) \cap Y_u(x^t)$
21: else
22: select an action from $Y'_u(x^t)$
23: end if
24: end if
25: Update $\xi^{t+1}_i\psi^T(x, y)$ according to (6.111)
26: Update $\pi^{t+1}_i(x, y)$ according to (6.112)
27: $x = x^{t+1}$
28: $t = t + 1$
29: end while
convergence of \( \{\xi^t\} \) to the optimal \( \xi^* \) arises as a result of Proposition 5. With Assumptions 7 and 8, the team stage game \( GE \) is reduced to a team game under network states \( X_F(x^t, F(t)) \). According to theorem 1 in [221], the U macro BSs coordinate resource allocation strategy for all \( x^t \) as long as \( \nu \leq F/(\rho_{GE} + 2) \), where \( \rho_{GE} \) is the length of the shortest path in the best response graph of team stage game \( GE \) [222]. This confirms that the decentralized resource allocation algorithm will converge with probability one.

6.4.5 Testbed Implementation

The H-CRANs system is implemented using USRP-N210 front-ends to demonstrate the performance of the proposed resource allocation scheme. The logical architecture of the considered H-CRANs in testbed implementation for both centralized and decentralized resource allocation consists of four main components: centralized BBU pool for baseband processing, macro BSs, RRHs represented by low power pico BSs, and the resource allocation controller integrated to the BBU pool. The heterogeneous network including macro and pico BSs are connected to the BBU pool through wired links.

As SINR is utilized by the online learning model for state information, this creates a need for a communication protocol that facilitates the acquisition of channel state information. The communication protocol performs three tasks: synchronization, SINR estimation and resource allocation. Therefore, the resource allocation frame structure is divided into three stages:

- Synchronization Stage: in this state each macro BS performs synchronization with the BSs operating under its coverage through broadcasting a beacon frame periodically to compensate for any misalignment in clock frequencies of each node. Moreover, the beacon frame includes some information about the network such as the number of time steps in each stage, and the time step occupancy in
the following stages.

- Acquisition Stage: this state is a TDMA-based state, where each UE report its channel state information to its corresponding pico BS or macro BS. It is composed of three time slots, where the first and second slots are spared for the macro BS and its associated MUEs to acquire their channel state information to find the channel gains of all links for each RB. The third slot is dedicated for pico BSs to report their channel state information while other pico BSs estimate the channel gain with that pico BS on the same RB.

- Resource Allocation Stage: at this state, the online learning-based resource allocation algorithms are executed whether at the controller as in the centralized approach or through macro BSs in the decentralized approach.

**Implementation Setup**

We exploit GNU radio [182] as an SDR development platform to create digital signals. GNU is an open-source software development toolkit that provides signal processing blocks to implement wireless protocols. As GNU radio can only handle digital data, RF front ends are required to shift the baseband signal to the desired center frequency. USRP-N210 from Ettus Research [223] is utilized as the RF front end in our setup with CBX daughter board that can operate in frequency range of 1 GHz to 6 GHz with one antenna. In addition, we add two Dell servers as SDR processors that run GNU radio to perform the baseband processing and include the controller implementation for resource allocation. The centralized and the decentralized resource allocation topologies considered in the testbed implementation are presented in Figure 6.30 and Figure 6.31 respectively. We have two macro BSs, four pico BSs, four RUEs, and two MUEs, where each of them is represented by one USRP-N210. In the centralized resource allocation approach, only one dell server is utilized as the controller and the
unit for baseband processing. However, two servers are used, where each one processes the resource allocation algorithm implemented in each macro BS in the distributed resource allocation approach.

Figure 6.30: Centralized testbed implementation topology of 5G H-CRANs

Figure 6.31: Decentralized testbed implementation topology of 5G H-CRANs
The conceptual models for each node in the network for both centralized and decentralized approaches for resource allocations are presented in Figure 6.32. This model consists of the following modules: communication protocol module, which is common in all nodes and it performs the tasks mentioned before. The online learning algorithm module, which includes the implementation of the resource allocation process in Algorithm 7 and 8. This module is implemented in the controller in the centralized approach and at the macro BS in the distributed methodology. GNU Radio PHY module performs the physical layer functionality supported by the physical layer modules in the GNU Radio software. Finally, SINR estimation module, which relies on the probe block in GNU Radio, where the probe performs an average magnitude square process on the samples acquired during the sensing time. The acquired results from the probe block is further used to estimate the signal and the noise power. Signal power estimation is performed by averaging the acquired samples from the probe block over certain time while the noise power is estimated by considering the variance of the acquired samples.

Figure 6.32: Conceptual model of the resource allocation system in 5G H-CRANs
Table 6.4 presents the parameters used for testbed implementation including the communication protocol, configuration and online learning parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Communication protocol parameters</strong></td>
<td></td>
</tr>
<tr>
<td>Acquisition stage time slots</td>
<td>3</td>
</tr>
<tr>
<td>Sensing duration</td>
<td>10 ms</td>
</tr>
<tr>
<td>Modulation scheme</td>
<td>gmsk</td>
</tr>
<tr>
<td><strong>Configuration Parameters</strong></td>
<td></td>
</tr>
<tr>
<td>Macro BS Tx power</td>
<td>20 dBm</td>
</tr>
<tr>
<td>System bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Carrier frequency</td>
<td>2.4 GHz</td>
</tr>
<tr>
<td>Sampling rate</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Antenna gain</td>
<td>3 dBi</td>
</tr>
<tr>
<td><strong>Online Learning Related Parameters</strong></td>
<td></td>
</tr>
<tr>
<td>Learning rate $\alpha$</td>
<td>0.5</td>
</tr>
<tr>
<td>Discount factor $\beta$</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Table 6.4: Parameters of the 5G H-CRANs system in Testbed Implementation

**Testbed Experimental Results**

In this section, we present the performance evaluation of our proposed centralized and decentralized approaches compared to the equal power allocation standard scheme. The logical structures in Figure 6.30 and Figure 6.31 are exploited as the network setup, where we have H-CRANs that incorporates two macro BSs. Each macro BS has two small BSs operate under its coverage and one MUE. One of the small cells communicates with an RUE 1 with high QoS requirements while the other cell is connected to low QoS RUE 2. RUE 1 traffic type is video file streaming while RUE 2 and MUE receive a small TCP file. The considered evaluation of the network operation targets UEs achieved capacity, network energy efficiency and UEs BER. Figures 6.33, 6.34, and 6.35 present the average capacity achieved by the small cell with high QoS requirements RUE, the small cell with low QoS requirements RUE, and macro BS with one MUE respectively.
We notice that our proposed scheme in both approaches outperforms the standard, where fixed power allocation is assumed in the achieved capacities for all users. Moreover, the proposed scheme maintains the capacity above the threshold. In Figure 6.33, the centralized approach for resource allocation converges to capacity of 20 bits/sec/Hz while the decentralized approach saturates at 18.5 bits/sec/Hz. However, the capacity of the low QoS requirements RUE is 14 bits/sec/Hz and 12.8 bits/sec/Hz for both centralized and distributed resource allocation approaches respectively as in Figure 6.34. Finally, the capacity for the macro UE is measured to be 15 bits/sec/Hz for the centralized approach and 14 bits/sec/Hz for the distributed approach as in
Figure 6.35: Average macro BS capacity in 5G H-CRANs system

The convergence time for the centralized scheme is 300 time steps while it is 330 for the decentralized scheme, where the time duration for each step is 0.065 ms.

The BER evaluations for RUE 1, RUE 2 and MUE are plotted in Figures 6.36, 6.37, and 6.38 respectively. The proposed resource allocation scheme in both approaches succeeded to maintain BER below the BER threshold which is 0.001 for RUE 1 as in Figure 6.36 and below BER threshold for RUE 2 which is 0.005 as in Figure 6.37. In addition, the recorded BER for the MUE using both decentralized and centralized
approaches for resource allocation is maintained at 0.00038 and 0.00022 respectively as in Figure 6.38. Thus, our scheme reduces the BER by a factor of 10 compared to the standard approach.

Figure 6.37: Average BER experienced by low QoS RUE in 5G H-CRANs system

Figure 6.38: Average BER experienced by MUE in 5G H-CRANs system

The average energy efficiency for the small cell that supports RUE 1 and the small cell that supports RUE 2 are plotted in Figures 6.39 and 6.40 respectively as a function of time step with duration of 0.065 ms. Figure 6.39 shows that our proposed scheme achieves efficiency of 1.9 bps/Hz/W and 1.75 bps/Hz/W using the centralized and decentralized approaches compared to 0.98 bps/Hz/W achieved by the standard.
On the other hand, our scheme achieves energy efficiency of 1.45 bps/Hz/W and 1.33 bps/Hz/W in both centralized and decentralized approaches respectively for RUE 2. Figures 6.39 and 6.40 also note the convergence time for both approaches, which is 300 for the centralized and 330 for the distributed.

Figure 6.39: Average energy efficiency for high QoS pico BS in 5G H-CRANs system

Figure 6.40: Average energy efficiency for low QoS pico BS in 5G H-CRANs system

The system evaluation in Figures 6.33 to Figure 6.40 demonstrates the capability of the implemented sophisticated online learning in resource allocation reflected by the superior results recorded. Moreover, we notice that the proposed scheme converges within short time thanks to the approximation of online learning Q-value that reduces
the difference between the approximated Q-value and the optimal one. Another point to report from the evaluation is that the decentralized approach converges at rate close to the centralized approach. This is because of the conjecture concept exploited in the decentralized learning as in Algorithm [8] which does not require explicit information exchange between the macro BSs and relies on estimation of other BSs action based on the historical knowledge for similar network states. This highlights a significant potential for deploying H-CRANs with distributed resource allocation approach in spite of the little loss in the achieved performance compared to the centralized approach. This is because the centralized approach requires a dedicated controller for resource allocation, which increases the complexity of the network topology and makes it subjected to sudden failure if the controller is down.

6.4.6 Numerical Results

In this section, we verify the performance of the proposed scheme in terms of energy efficiency, spectral efficiency, and QoS. The evaluation environment consists of three macro BSs with 21 MUEs, 15 RRHs with 47 RUEs accessing $\Gamma_1$ and 28 RUEs sharing the spectrum with MUEs in $\Gamma_2$. It is assumed that the path-loss model is expressed as $31.5 + 40*\log_{10}(d)$ for RRH to RUE link and $31.5 + 35*\log_{10}(d)$ for macro BS to RUE and RRH to MUE links, where $d$ is the distance between the transmitter and receiver. Fast-fading coefficients are all generated as independent and identically distributed Rayleigh random variables with unit variances. The data rate thresholds per both types of RUEs and MUE $\theta, \theta^*$ and $\delta$ are assumed to be 2 Mbps, 512 Kbps, and 1.2 Mbps, respectively. We assume that the number of BFs is equal to the number of RBs that are within the corresponding sub-band. The rest of the simulation parameters are presented in Table [6.5]. The evaluation of the proposed scheme including energy efficiency, spectral efficiency, and data rate is investigated in the following sections.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>User distribution</td>
<td>uniform</td>
</tr>
<tr>
<td>Number of RBs</td>
<td>50</td>
</tr>
<tr>
<td>Total bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Thermal noise power</td>
<td>-112 dBm</td>
</tr>
<tr>
<td>Macro BS transmission power</td>
<td>43 dBm</td>
</tr>
<tr>
<td>Back-haul power consumption $P_{bh}$</td>
<td>23 dBm</td>
</tr>
<tr>
<td>Antenna gain for macro/RRH</td>
<td>17/6 dB</td>
</tr>
<tr>
<td>Pico maximum transmission power</td>
<td>25 dBm</td>
</tr>
<tr>
<td>Trials per experiment</td>
<td>1000</td>
</tr>
</tbody>
</table>

Table 6.5: Parameters of the 5G H-CRANs system in Simulation

**Energy Efficiency Evaluation**

In this evaluation, we are interested in measuring the speed of convergence and the achieved energy efficiency for both types of RUEs. Thus, we conduct three simulations to evaluate the system convergence, achieved energy efficiency with variable maximum transmission power of RRHs and energy efficiency against the SINR threshold of MUEs. First, we plot energy efficiency as function of the number of time steps in Figure 6.41. The performance of our centralized and decentralized resource allocation is compared to two schemes including the standard with fixed power allocation in which the same power is allocated for all RBs and the scheme proposed in [97], which aims at tackling the energy efficiency problem in H-CRANs denoted by (EE-
Moreover, we include typical online learning resource allocation, which is online Q-learning without the enhancement of the brief representation proposed for the Q-value in which they are approximated as a function of much smaller set of variables. The inclusion of this typical online learning is to demonstrate the advantage of approximation of online learning in speed of convergence. From Figure 6.41, we notice that our scheme converges after 300 iterations faster than typical online learning and EE-HCRAN and achieved the highest energy efficiency.

Second, we plot the achieved energy efficiency against the maximum transmission power of RRH ($P_{\text{max}}$) in Figure 6.42. It is observed that the average energy efficiency is monotonically non-decreasing function of $P_{\text{max}}$. With small value of $P_{\text{max}}$, the energy efficiency increases until it saturates at $P_{\text{max}}$ of 23 dBm. This is due to the fact that the compared schemes aim to balance the system energy efficiency and the power consumption. The further increase in transmission power will result in degradation in energy efficiency. We notice that both versions of our scheme achieved higher energy efficiency than EE-HCRAN.

![Figure 6.42: Energy efficiency with variable $P_{\text{max}}$ in 5G H-CRANs system](image)

The third evaluation focuses on the system with variable SINR threshold of MUEs. Figure 6.43 presents the energy efficiency achieved versus the SINR threshold of MUEs accessing $\Gamma_2$ with $P_{\text{max}} = 25$ dBm. We notice that the proposed scheme outperforms
both EE-HCRAN and the standard schemes. Figure 6.43 reveals that when SINR threshold is not large, the energy efficiency is stable with the increasing threshold because the inter-tier interference is not severe thanks to the proposed RB allocation strategy in conjunction with sophisticated online learning. This indicates that the proposed solution mitigates the inter-tier interference and provide higher bit rates for MUEs than the other schemes.

Figure 6.43: Energy efficiency with variable SINR threshold for MUEs in 5G H-CRANs system

**Spectral Efficiency and QoS Evaluation**

In this section, we evaluate the proposed scheme performance in terms of spectral efficiency and QoS represented by the data rate achieved by both RUEs accessing $\Gamma_1$ and $\Gamma_2$, and MUEs. Thus, we plot the average system spectral efficiency against the time steps and maximum transmission power of RRH in Figures 6.44 and 6.45 respectively. Figure 6.44 emphasizes the speed of convergence achieved by our scheme compared to others. Our scheme is the fastest with the highest spectral efficiency. In Figure 6.45, spectral efficiency with variable maximum transmission power shows that both approaches of our scheme record the best level of spectral efficiency compared to other schemes.
QoS requirements stated in C2, C3 and C4 are investigated in this evaluation. Figures 6.46, 6.47 and 6.48 present the CDF of the data rate for both RUEs accessing sub-band $\Gamma_1$ and $\Gamma_2$, and the CDF of the data rate achieved by MUEs respectively. We notice that our scheme (Centralized RA) and (Decentralized RA) is the only scheme that managed to have more than 97% of the users above the specified thresholds $\theta$, $\theta^*$ and $\delta$ compared to EE-HCRAN that records 78% above threshold, and standard with 65%. This evaluation demonstrates the capability of the online learning scheme to allocate RB and power efficiently while maintaining QoS of users.
at the maximum level.

Figure 6.46: Data rate CDF for RUEs with high QoS requirements in 5G H-CRANs system

Figure 6.47: Data rate CDF for RUEs with low QoS requirements in 5G H-CRANs system

Other Evaluations Aspects and Discussion

We consider the case of dynamic users who join and leave the system frequently. These users may be viewed as distributed randomly. Their arrival is modeled as homogeneous Poisson point process with intensities $\lambda = 3$. Different users have independent duration to stay in the system. We notice the impact of considering users mobility on the achieved energy efficiency, spectral efficiency and data rate in the results of Figures 6.41, 6.44, 6.46, 6.47, and 6.48. Figure 6.41 and Figure 6.44 reveal that there
is a drop of 27% in the energy efficiency and spectral efficiency as a result of user mobility when comparing the performance of our scheme in the case of stationary users (S) and mobile users (M). However, our scheme has only a drop of 14% in the data rate achieved for both types of RUEs and MUEs and all the achieved rates are still above the threshold. The degradation in performance in the mobile users case is due to the difficulty to acquire the state information as it changes frequently. From the presented results, we notice that our proposed scheme with both approaches outperforms the EE-HCRAN scheme in terms of the achieved energy efficiency with gain of 16% for the decentralized resource allocation and 24% for the centralized resource allocation. The use of machine learning in resource allocation is superior compared to the simple convex optimization used in EE-HCRAN for resource allocation as online learning does not require specific model of the network environment. This is a significant factor as 5G H-CRANs is a dynamic environment, which cannot be tied to specific model. Therefore, online learning, which adopts learning from experience approach is a good fit for resource allocation problem. In addition, the compact representation for states and approximation for Q-value contribute to the enhancement of speed of convergence in our scheme. Another issue to note in the proposed scheme is that the decentralized approach achieved comparable results in the evaluation which
reflects the advantage of the conjecture feature in the exploited learning approach. This feature eliminates the need for explicit cooperation between the macro BSs to exchange information that improve the quality of the selected action. In addition, this relieves the BBU pool from resource allocation responsibility and reduces the signal processing overhead. Another factor for the superior performance achieved is the sophisticated RB allocation mechanism that considers both location and QoS of the RUEs. Moreover, appropriate power allocation has a considerable contribution to the achieved results.

<table>
<thead>
<tr>
<th>Evaluation Metric</th>
<th>Proposed Centralized Testbed</th>
<th>Proposed Decentralized Testbed</th>
<th>Proposed Centralized (S) Simulation</th>
<th>Proposed Decentralized (S) Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy efficiency for high QoS RUEs</td>
<td>2 bps/Hz/W</td>
<td>1.7 bps/Hz/W</td>
<td>2.15 bps/Hz/W</td>
<td>1.92 bps/Hz/W</td>
</tr>
<tr>
<td>Energy efficiency for low QoS RUEs</td>
<td>1.45 bps/Hz/W</td>
<td>1.25 bps/Hz/W</td>
<td>1.62 bps/Hz/W</td>
<td>1.34 bps/Hz/W</td>
</tr>
<tr>
<td>Spectral efficiency</td>
<td>285 bps/Hz</td>
<td>260 bps/Hz</td>
<td>302 bps/Hz</td>
<td>275 bps/Hz</td>
</tr>
<tr>
<td>Average data rate for high QoS RUEs</td>
<td>5.6 Mbps</td>
<td>5 Mbps</td>
<td>6 Mbps</td>
<td>5.55 Mbps</td>
</tr>
<tr>
<td>Average data rate for low QoS RUEs</td>
<td>2 Mbps</td>
<td>1.6 Mbps</td>
<td>2.2 Mbps</td>
<td>1.78 Mbps</td>
</tr>
<tr>
<td>Average data rate for MUEs</td>
<td>2.1 Mbps</td>
<td>1.73 Mbps</td>
<td>2.3 Mbps</td>
<td>1.9 Mbps</td>
</tr>
<tr>
<td>Convergence Speed (time steps)</td>
<td>300</td>
<td>325</td>
<td>290</td>
<td>310</td>
</tr>
</tbody>
</table>

Table 6.6: Testbed and numerical results comparison

Finally, we compare the testbed results and numerical results in terms of energy efficiency for high QoS RUEs, energy efficiency for low QoS RUEs, system spectral efficiency, data rate for both types RUEs and MUEs, and convergence speed. The results are compared on the basis of the optimal (maximum) value reached for energy efficiency and spectral efficiency and on average for data rate. Table 6.6 presents the comparison results. We notice that the numerical results record better values than the testbed implementation. The reason for that little win is due to real time interfering means such as other wireless devices in the lab. In addition, hardware limitation on processing and synchronization with the controlling unit may degrade the achieved results.
Chapter 7

Concluding Remarks

CR is an exciting communication paradigm, which have a potential to enable easier management, self-organization and better performance of the future radio networks. Nonetheless this paradigm has been studied over the past years, There are still large number of challenges that obstacle their path. One of the contributions of this thesis is to provide practical realization of cognitive radios. We have proposed algorithmic solutions for facilitating adaptation and optimization in key resource sharing problems such as channel allocation and radio parameters optimization. Additionally, we have introduced our cognitive resource management framework (CogWnet), which encompasses tools and mechanisms for cross-layer optimization, cognitive operations and advanced management of future networks. In addition, resource allocation problem in the future 5G Hetnets and 5G H-CRANs are tackled using enhanced online learning. The resource allocation is formalized with the goal to maximize energy efficiency under the channel conditions and QoS constraints. At the beginning of this thesis, the basic concepts of CR and SDRs are introduced. Mitola’s vision was reviewed and as a result, CR can be classified into two types: the first one deal with DSA and improving spectrum utilization. Moreover, the upcoming 5G technology with all its merits to maximize network capacity, provide efficient radio resource management, and accommodate heterogeneous users applications is presented. The challenges encountered to fulfill the 5G vision and in resource allocation are also highlighted.

Following Mitola’s idea on fully reconfigurable radio that can learn, analyze, de-
cide and act upon the environmental stimuli, we have proposed CogWnet that could enable easier realization of the cognitive cycle. This framework is one of the first cognitive RRM architectures on a system level that suggested all the necessary components to implement the cognition cycle. The design of CogWnet has been performed in a way that several key design principles such as extensibility, flexibility, portability and reasonable complexity are guaranteed. CogWnet is component based, highly modular and easily extensible to include new functionalities. Furthermore, it has been designed to facilitate easier interaction and exchange of information between the layers of the protocol stack by means of generic interfaces. A central layer of the architecture is a so called decision-making layer, which is responsible for coordination of actions, resource scheduling and decision-making. The decision-making layer reaches optimal decisions exploiting information from different sources in the radio environment and historical data that have been generated from off-line learning. One of the novelties in CogWnet is the optimization techniques, which is tightly coupled to the decision-making layer and can be used to perform local and global optimization based on the rich information provided through the interfaces. The AI algorithms utilized for the decision-making functionality in CogWnet are described in Chapter 4. These algorithms including decision trees, genetic algorithm, case based reasoning, artificial neural networks, and reinforcement learning, are exploited in different approaches for radio parameters configuration. These approaches include single, hybrid and supervised decision-making engines. The performance of these methodologies is evaluated and compared to highlight the advantage of considering each of them and highlight related tradeoffs.

Further in this thesis, CogWnet was integrated with LTE cellular network to improve its efficiency. The overall goal for the integrated system was to optimize spectrum allocation, mitigate interference, maximize throughput and reduce complexity. Radio environment awareness and optimization algorithms are used to improve net-
work efficiency and respond to changes in network conditions. Optimization starts with receiving periodic channel information. SINR, traffic load and BER were used to tune modulation, power, frequency, and bandwidth.

Another important contribution of this thesis is the efficient resource allocation in the 5G networks including Hetnets and H-CRANs trends. Power allocation problem is tackled for the downlink transmission in a spectrum sharing multi-tier 5G environment, where small cells and D2D access the spectrum in an underlay fashion. We proposed an enhanced online learning based scheme to allocate transmission power to reduce the overall power consumption while maintaining QoS for both primary tier and secondary tier. The online learning exploits an intuition based approach to account for the impact of other users transmissions on the selected transmission power strategy. In addition, a green resource allocation scheme in H-CRANs network following online learning based centralized and decentralized approaches is proposed. RBs and transmission power are allocated subjected to inter-tier interference and capacity constraints. The centralized approach places a dedicated controller integrated to the BBU pool to perform resource allocation while the decentralized approach seizes the macro BSs awareness about RRHs operates under their coverage and assign them the task of resource allocation in a distributed fashion.
REFERENCES


[23] ——, “A cooperative online learning scheme for resource allocation in 5g systems,” in IEEE International Conference on Communications (ICC), May 2016, pp. 1–7.


[26] ——, “Enhanced machine learning scheme for energy efficient resource allocation in 5g heterogeneous cloud radio access networks,” in IEEE Symposium on


PUBLICATIONS


