CROSSFIRE

Uncoordinated network strategies for enhanced interference, mobility, radio resource, and energy saving management in LTE-Advanced networks

FP7 Contract Number: 317126

Deliverable D3.1: Architectural aspects for SONs and cognition in LTE-A

WP3

<table>
<thead>
<tr>
<th>Contractual Date of Delivery:</th>
<th>February 2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Date of Delivery:</td>
<td>February 2014</td>
</tr>
<tr>
<td>Responsible Beneficiary:</td>
<td>IQU</td>
</tr>
<tr>
<td>Contributing Beneficiaries:</td>
<td>KCL, UOC, NEC, IQU</td>
</tr>
<tr>
<td>Security:</td>
<td>Public</td>
</tr>
<tr>
<td>Nature</td>
<td>Report</td>
</tr>
<tr>
<td>Version:</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Document Information

Version Date: 28/2/2014
Total Number of Pages: 92

Authors

<table>
<thead>
<tr>
<th>Name</th>
<th>Organisation</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panagiotis Trakas</td>
<td>UoC</td>
<td><a href="mailto:ptrakas@uoc.gr">ptrakas@uoc.gr</a></td>
</tr>
<tr>
<td>Eirini Liotou</td>
<td>UoA</td>
<td><a href="mailto:eliotou@di.uoa.gr">eliotou@di.uoa.gr</a></td>
</tr>
<tr>
<td>Rudraksh Shrivastava</td>
<td>NEC</td>
<td><a href="mailto:rudraksh.shrivastava@neclab.eu">rudraksh.shrivastava@neclab.eu</a></td>
</tr>
<tr>
<td>Christoforos Vlachos</td>
<td>KCL</td>
<td><a href="mailto:christoforos.vlachos@kcl.ac.uk">christoforos.vlachos@kcl.ac.uk</a></td>
</tr>
<tr>
<td>Giorgos Kollias</td>
<td>IQU</td>
<td><a href="mailto:gkollias@iquadrat.com">gkollias@iquadrat.com</a></td>
</tr>
</tbody>
</table>

Document History

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date</th>
<th>Modification</th>
<th>Contact Person</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>08/12/13</td>
<td>First ToC</td>
<td>N. Zorba</td>
</tr>
<tr>
<td>0.2</td>
<td>22/12/13</td>
<td>Second ToC</td>
<td>N. Zorba</td>
</tr>
<tr>
<td>0.5</td>
<td>26/1/2014</td>
<td>First input from partners</td>
<td>N. Zorba</td>
</tr>
<tr>
<td>1.0</td>
<td>28/2/2014</td>
<td>Final deliverable</td>
<td>N. Zorba</td>
</tr>
</tbody>
</table>
Executive Summary

This document shows the activities carried out within the consortium about architectural aspects for SONs and cognition within the wireless domain in general and in the in LTE-A standard in particular, where each ESR has concentrated on a specific topic within the SON field. A preliminary review of the state of the art in the selected topic is accomplished by each ESR in order to realize the current status on the science in the research arena, while identifying green spots that need to be filled by their work.

The first topic that is tacked in this document is the Device-to-Device (D2D) communication that is motivated by the emergence of new applications and user entertainment demands such as video sharing, content distribution and location-aware services (e.g., Foursquare on Facebook), all of them resulted in the introduction of different use cases for the Device-to-Device communications operating in a cellular network. D2D communication was firstly considered to support multi-hop relay in cellular networks (i.e. MCNs), where the transmission in multiple steps via several mobile stations is allowed in order to optimize the total network throughput. Other significant usage cases have been also identified like video dissemination, multicasting, peer-to-peer communication (P2P), machine-to-machine communication (M2M), cellular offloading and other cases. Within the SoA review, we will describe some of the challenges and research subjects that were studied and most of them remain at the centre of attention. Compliant with the research flow, as most of the works so far were handling the underlaying inband concept of D2D communication, we review all the D2D approaches to get familiar with the covered and uncovered research issues.

The second topic is the handover (HO) that is carried out between two cells while the user is moving, in order to keep the connection to the user. HO is one of the main topics within SON as it is automatically done without the user’s knowledge. A detailed review of the SoA is first presented showing all the HO options while concentrating on the LTE-A standard. We discussed all options on intracell HO, intercells HO and interRAT HO and compared their performance benefits and drawbacks. Moreover, we reviewed the Heterogeneous Networks (HetNETs) and showed the challenges faced to tackle multiple kinds of applications at the same time and how to optimize the resources allocation over them. The joint consideration of HO and HetNETs seems promising as well a required strategy to deal with the current and future usage profiles, which attracted our attention as a future work.

The third topic is the traffic offloading over access technologies, as during the past few years, mobile network operators have given a lot of attention and funding in traffic offloading. Subsequently, the research community has been proposing a lot of interesting ideas regarding its efficient implementations. We first provide the SoA in traffic offloading from two different views. The first refers to the utilization of SONs and cognitive radio, while the other targets its economic issues. We concentrate in our current research on the financial aspects and the transactions between different participants that small cell traffic offloading incorporates. We aim to bring the attention on a network sharing scheme that will become common in the future. This scheme refers to particular individuals that have in their possession small cell infrastructure, which can be utilized by the current mobile network operators and service providers. Such individuals could be municipal authorities, public services and schools or athletic organizations, which own such infrastructure.

The last topic is related to SON and cognition for decentralized network management where we concentrate on the concept of virtual cells in TD-LTE system, which enable users residing in overlapping coverage regions of different cells to utilize resources from multiple base stations. The virtual cell scheme addresses the problem of pseudo-congestion by efficiently managing, scheduling and sharing the available resources in both UL and DL directions. It provides improved resource flexibility and real time adaptation to serve the varying traffic demands in UL and DL directions with enhanced overall system performance. We tackle the resource management and provide a new point of view for allocation that considers the issue of cross-slot interference that arises in case of dynamic cell specific re-configuration in TD LTE systems.
Table of Contents

List of Figures .................................................................................................................. 5
List of Tables .................................................................................................................... 7
List of Abbreviations and Acronyms ................................................................................. 8
1. Introduction .................................................................................................................. 11
2. Architectural aspects for SONs and cognition for device to device communications .... 12
   2.1 State-of-Art analysis ................................................................................................. 14
       2.1.1 Device to Device communications .................................................................. 14
       2.1.2 Resource Allocation ....................................................................................... 20
   2.2 Current research workflow ...................................................................................... 25
       2.2.1 Scenarios and methodology .......................................................................... 25
       2.2.2 Performance analysis .................................................................................... 27
   2.3 Future research planning ......................................................................................... 27
3. Architectural aspects for SONs and cognition for handovers optimization ............... 29
   3.1 Handover Procedures - State of the Art ................................................................. 29
       3.1.1 Network Entities and Interfaces .................................................................... 29
       3.1.2 UE States ........................................................................................................ 31
       3.1.3 Handover Types ............................................................................................ 32
       3.1.4 Handover Procedure in detail ........................................................................ 33
       3.1.5 Intra-LTE, Inter-LTE and Inter-RAT Handovers ............................................ 35
       3.1.6 Handover Triggering Events .......................................................................... 39
   3.2 HetNEts and Standardization .................................................................................. 44
   3.3 HetNEts and Handover Procedures ........................................................................ 49
   3.4 Work plan and future results ................................................................................ 53
4. Architectural aspects within SONs for Traffic Offloading with small-cells .............. 56
   4.1 Traffic offloading in LTE-A ..................................................................................... 56
       4.1.1 Local IP Access (LIPA) and Selected IP Traffic Offload (SIPTO) .................. 56
       4.1.2 Network Sharing Scenarios and Use Cases ................................................. 65
   4.2 State of the Art on traffic offloading with small-cells .......................................... 67
       4.2.1 SONs and cognition aspects ......................................................................... 67
       4.2.2 Financial aspects .......................................................................................... 70
5. Architectural aspects for SONs and cognition for decentralized network management methods, policies and algorithms ................................................................. 78
   5.1 SON Architectures ................................................................................................. 78
       5.1.1 SON-Load Balancing and Information Exchange for TD-LTE System .......... 80
       5.1.2 SON Interference Management ..................................................................... 82
   5.2 SON Optimization and Management ...................................................................... 84
       5.2.1 Virtual Cell Concept-Overview ..................................................................... 84
       5.2.2 Self-Organized Network Optimization and Management ............................. 85
       5.2.3 Virtual Cell Distributed SON Resource Management ............................... 86
   5.3 Conflict Handling ................................................................................................... 86
6. Conclusion .................................................................................................................... 87
References ....................................................................................................................... 88
List of Figures

Figure 2.1: Smartphones' penetration in market according to GSMA ......................................................... 12
Figure 2.2: Mobile data volume growth estimation in Europe (2012–2017) .................................................. 13
Figure 2.3: D2D communication figure ....................................................................................................... 13
Figure 2.4: SIP Initiation figure in D2D ....................................................................................................... 16
Figure 2.5: D2D use cases paradigms (from the work of Lei et al. [2]) ............................................................ 17
Figure 2.6: Resource allocation scheme with FFR use in D2D-supporting cellular networks ([24]) ................. 21
Figure 2.7: Column generation method model ([29]) .................................................................................... 24
Figure 2.8: D2D Intra-cluster re-transmissions with two multicast senders .................................................. 25
Figure 2.9: Single Cell Scenario deployment with n=20 users ...................................................................... 26
Figure 2.10: Timeslot scheduling depending on the number of links ............................................................ 27
Figure 2.11: 7-cell scenario with total number of 20 D2D and cellular UEs per cell ........................................ 27
Figure 3.1: S1 Interface Architecture [40] .................................................................................................... 30
Figure 3.2: Handover example .................................................................................................................... 32
Figure 3.3: Intra-RAT and Inter-RAT mobility ............................................................................................... 33
Figure 3.4: Message Sequence Diagram of Handover Procedure for Macro to Macro case [59] ...................... 34
Figure 3.5: Intra-LTE (Intra-MME/SGW) Handover Using the X2 interface [45] ........................................... 36
Figure 3.6: Intra-LTE (Intra-MME/SGW) Handover Using the S1 interface [45] ........................................... 36
Figure 3.7: Inter-MME Handover Using the S1 Interface (Without changing S-GW) [45] ............................... 37
Figure 3.8: Inter-MME/SGW Handover Using the S1 Interface [45] .............................................................. 37
Figure 3.9: Inter-RAT Handover (E-UTRAN to UTRAN) preparation and execution phase [45] .................... 38
Figure 3.10: Inter-RAT Handover (UTRAN to E-UTRAN) preparation and execution phase [45] ................. 39
Figure 3.11: Entering and Leaving Condition for Event A1 ........................................................................ 41
Figure 3.12: Entering and Leaving Condition for Event A2 ........................................................................ 41
Figure 3.13: Entering and Leaving condition for Event A3 ......................................................................... 42
Figure 3.14: Entering and Leaving condition for Event A4 ......................................................................... 42
Figure 3.15: Entering and Leaving condition for Event A5 ......................................................................... 43
Figure 3.16: Entering and Leaving condition for Event B1 ......................................................................... 43
Figure 3.17: Entering and leaving condition for Event B2 ......................................................................... 44
Figure 3.18: Moving towards to HetNets layout [46] .................................................................................. 44
Figure 3.19: Overall E-UTRAN Architecture with deployed HeNB GW [44] ............................................... 46
Figure 3.20: 3GPP HetNet scenarios [51] ...................................................................................................... 47
Figure 3.21: User Case scenario defined by 3GPP [50] ............................................................................... 48
Figure 3.22: Co-Channel Deployment of macro and small cells [50] ............................................................ 48
Figure 3.23: Handover failure case 1 [52] .................................................................................................... 50
Figure 3.24: Handover failure case 2 [52] .................................................................................................... 50
Figure 3.25: Number of Ping-Pongs = 3 [52] ............................................................................................... 50
Figure 3.26: CROSSFIRE architecture ..................................................................................................... 55
Figure 4.1: LIPA solution for HeNodeB [73] ................................................................................................ 57
Figure 4.2: UE Requested PDN Connectivity to LIPA ................................................................................ 59
Figure 4.3: S1 Release procedure when LIPA PDN connection for UE exists ............................................... 59
Figure 4.4: UE Triggered Service request procedure if LIPA PDN connection existed ............................... 60
Figure 4.5: L-GW Triggered Service request procedure if LIPA PDN connection existed ............................ 60
Figure 4.6: SIPTO for EPC macro network with S4 SGSN [73] .................................................................... 62
Figure 4.7: E-UTRAN Tracking Area Update with S-GW change ............................................................... 63
Figure 4.8: UE or MME requested PDN disconnection ................................................................................ 63
Figure 4.9: MME triggered S-GW relocation ................................................................. 64
Figure 4.10: Cognitive functional architecture mapped to the LTE offloading scenario [77] ................................................................. 67
Figure 4.11: System overview for offloading of LTE [77] .................................................................................................................. 68
Figure 4.12: The two SSRs for a generic FAP (the black dot at the center): the outer dotted circle and the inner dashed circle represent, respectively, the macro SSR and the femto SSR [78] ................................................................. 69
Figure 4.13: JAB-approach with two MSs, one picoBS and one BS [79] ................................................................................................. 69
Figure 4.14: Cognitive hybrid division duplex scheme [80] ................................................................................................................. 70
Figure 4.15: An instance of the data offloading scenario [82] ............................................................................................................. 71
Figure 4.16: Each BS is managed by a different MNO and serves a set of MUs [83] ............................................................................... 71
Figure 4.17: A cognitive radio network with one primary owner, two primary users, and four secondary users [84] ................................................................. 72
Figure 4.18: MultiMUE: auction with multiple MUEs and femtocells [85] ................................................................................................. 72
Figure 4.19: Sample cellular sector and its WiFi regions [86] .................................................................................................................. 73
Figure 4.20: Hierarchical dynamic game framework for pricing, spectrum sharing, and service selection [87] .... 73
Figure 4.21: The network scenario used in [88] ........................................................................................................................................... 74
Figure 4.22: System model [90] ......................................................................................................................................................... 75
Figure 4.23: Policy based offloading architecture and mechanism [91] ................................................................................................. 75
Figure 4.24: A general scenario topology ............................................................................................................................................. 76
Figure 5.1: Centralized SON .......................................................................................................................... 78
Figure 5.2: Distributed SON ......................................................................................................................................................... 79
Figure 5.3: Hybrid SON ............................................................................................................................................................... 79
Figure 5.4: Uplink / Downlink sub-frame configurations for LTE TDD (TD-LTE) ................................................................. 80
Figure 5.5: Cross-Slot interference representation for Cell Specific Dynamic UL/DL Configuration ..................... 81
Figure 5.6: A simple example of virtual cell with the UE utilizing resources from both eNBs .......... 84
List of Tables

Table 2-1: Simulation parameters for D2D underlaying LTE network .......................................................... 28
Table 3.1: Clarification of HeNB Access [48] ................................................................................................. 45
Table 3.2: Handover Cases Supported by X2 interface [44] ........................................................................ 46
## List of Abbreviations and Acronyms

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2G</td>
<td>Second Generation</td>
</tr>
<tr>
<td>3G</td>
<td>Third Generation</td>
</tr>
<tr>
<td>3GPP</td>
<td>Third Generation Partnership Project</td>
</tr>
<tr>
<td>4G</td>
<td>Fourth Generation</td>
</tr>
<tr>
<td>ACP</td>
<td>ACcess Permission</td>
</tr>
<tr>
<td>ACCS</td>
<td>Autonomous Component Carrier Selection</td>
</tr>
<tr>
<td>AP</td>
<td>Access Point</td>
</tr>
<tr>
<td>AL</td>
<td>Adaptation Layer</td>
</tr>
<tr>
<td>APN</td>
<td>Access Point Name</td>
</tr>
<tr>
<td>B3G</td>
<td>Beyond Third Generation</td>
</tr>
<tr>
<td>BS</td>
<td>Base Station</td>
</tr>
<tr>
<td>CA</td>
<td>Carrier Aggregation</td>
</tr>
<tr>
<td>CCCH</td>
<td>Common Control Channel</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CHDD</td>
<td>Cognitive Hybrid Division Duplex</td>
</tr>
<tr>
<td>CM</td>
<td>Cognitive Manager</td>
</tr>
<tr>
<td>CM-RM</td>
<td>Cognitive Manager for Resource Management</td>
</tr>
<tr>
<td>CM-SM</td>
<td>Cognitive Manager for Spectrum Management</td>
</tr>
<tr>
<td>CN</td>
<td>Core Network</td>
</tr>
<tr>
<td>CNME</td>
<td>Central Network Management Entity</td>
</tr>
<tr>
<td>CoMP</td>
<td>Coordinated Multi-Point Transmission and Reception</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel State Information</td>
</tr>
<tr>
<td>CSG</td>
<td>Closed subscriber Group</td>
</tr>
<tr>
<td>CU</td>
<td>Cellular User</td>
</tr>
<tr>
<td>D2D</td>
<td>Device-to-Device</td>
</tr>
<tr>
<td>DCBCAA</td>
<td>Distributed Capacity Based Channel Allocation Algorithm</td>
</tr>
<tr>
<td>DCA</td>
<td>Dynamic Channel Allocation</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier Transform</td>
</tr>
<tr>
<td>DL</td>
<td>Downlink</td>
</tr>
<tr>
<td>DSL</td>
<td>Digital Subscriber Line</td>
</tr>
<tr>
<td>DOFF</td>
<td>Data OFFloading game</td>
</tr>
<tr>
<td>E-RAB</td>
<td>E-UTRAN Radio Access Bearer</td>
</tr>
<tr>
<td>eNB</td>
<td>E-Node B</td>
</tr>
<tr>
<td>EPC</td>
<td>Evolved Packet Core</td>
</tr>
<tr>
<td>E-UTRAN</td>
<td>Evolved UMTS Terrestrial Radio Access Network</td>
</tr>
<tr>
<td>FAP</td>
<td>Femtocell Access Point</td>
</tr>
<tr>
<td>FDD</td>
<td>Frequency Division Duplex</td>
</tr>
<tr>
<td>FEC</td>
<td>Forward Error Correction</td>
</tr>
<tr>
<td>FFR</td>
<td>Fractional Frequency Reuse</td>
</tr>
<tr>
<td>FDMA</td>
<td>Frequency Division Multiple Access</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FO</td>
<td>Fixed line Operator</td>
</tr>
<tr>
<td>FRF</td>
<td>Frequency Reuse Factor</td>
</tr>
<tr>
<td>FSTP</td>
<td>Feasible Simultaneous Transmission Pattern</td>
</tr>
<tr>
<td>FSSG</td>
<td>Femtocell Service Selection Game</td>
</tr>
<tr>
<td>GP</td>
<td>Greedy Physical</td>
</tr>
<tr>
<td>GSM</td>
<td>Global System for Mobile communications</td>
</tr>
<tr>
<td>GSMA</td>
<td>Group Speciale Mobile Association</td>
</tr>
<tr>
<td>HARQ</td>
<td>Hybrid Automatic Repeat-Request</td>
</tr>
<tr>
<td>HPLMN</td>
<td>Home Public Land Mobile Network</td>
</tr>
<tr>
<td>HeNB</td>
<td>Home E-Node B</td>
</tr>
<tr>
<td>HM</td>
<td>Hysteresis Margin</td>
</tr>
<tr>
<td>HO</td>
<td>Handover</td>
</tr>
<tr>
<td>HSS</td>
<td>Home Subscriber Server</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>IHO</td>
<td>Intracell Handover</td>
</tr>
<tr>
<td>I-CA</td>
<td>Iterative – Combinational Auction</td>
</tr>
<tr>
<td>IC</td>
<td>Interference Cancellation</td>
</tr>
<tr>
<td>ICIC</td>
<td>Inter-Cell Interference Cancellation</td>
</tr>
<tr>
<td>IFFT</td>
<td>Inverse Fast Fourier Transform</td>
</tr>
<tr>
<td>IMT-Advanced</td>
<td>International Mobile Telecommunications - Advanced</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>ISD</td>
<td>Inter-Site Distance</td>
</tr>
<tr>
<td>ITU</td>
<td>International Telecommunications Union</td>
</tr>
<tr>
<td>ITU-R</td>
<td>International Telecommunications Union - Radio communications</td>
</tr>
<tr>
<td>JAB</td>
<td>Joint Access and Backhaul</td>
</tr>
<tr>
<td>JV</td>
<td>Joint Venture</td>
</tr>
<tr>
<td>JOMSRAP</td>
<td>Joint Mode Selection, Resource Allocation and Power Allocation</td>
</tr>
<tr>
<td>LOS</td>
<td>Line-of-Sight</td>
</tr>
<tr>
<td>LTE</td>
<td>Long Term Evolution</td>
</tr>
<tr>
<td>LTE-A</td>
<td>Long Term Evolution - Advanced</td>
</tr>
<tr>
<td>L-GW</td>
<td>Local Gateway</td>
</tr>
<tr>
<td>LIPA</td>
<td>Local IP Access</td>
</tr>
<tr>
<td>M2M</td>
<td>Machine-to-Machine</td>
</tr>
<tr>
<td>MAC</td>
<td>Medium Access Control</td>
</tr>
<tr>
<td>MCL</td>
<td>Minimum Coupling Loss</td>
</tr>
<tr>
<td>MCS</td>
<td>Modulation and Coding Scheme</td>
</tr>
<tr>
<td>MILP</td>
<td>Mixed Integer Linear Programming</td>
</tr>
<tr>
<td>MINLP</td>
<td>Mixed Integer Non-Linear Programming</td>
</tr>
<tr>
<td>MIMO</td>
<td>Multiple Input / Multiple Output</td>
</tr>
<tr>
<td>MIP</td>
<td>Mixed Integer Programming</td>
</tr>
<tr>
<td>MME</td>
<td>Mobile Management Entity</td>
</tr>
<tr>
<td>MTC</td>
<td>Machine Type Communication</td>
</tr>
<tr>
<td>MBS</td>
<td>Mobile Base Station</td>
</tr>
<tr>
<td>MNO</td>
<td>Mobile Network Operator</td>
</tr>
<tr>
<td>MO</td>
<td>Mobile Operator</td>
</tr>
<tr>
<td>MSP</td>
<td>Macrocell Service Provider</td>
</tr>
<tr>
<td>MTS</td>
<td>Minimum Time of Stay</td>
</tr>
<tr>
<td>MU</td>
<td>Multi-User</td>
</tr>
<tr>
<td>MU-MIMO</td>
<td>Multi-User Multiple-Input Multiple-Output</td>
</tr>
<tr>
<td>NAS</td>
<td>Non Access Stratum</td>
</tr>
<tr>
<td>NRP</td>
<td>Number of Receivers-based Packing</td>
</tr>
<tr>
<td>OFDMA</td>
<td>Orthogonal Frequency-Division Multiple Access</td>
</tr>
<tr>
<td>OPEX</td>
<td>Operational Expenses</td>
</tr>
<tr>
<td>P2P</td>
<td>Peer-to-Peer</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PDN-GW</td>
<td>Packet Data Network - Gateway</td>
</tr>
<tr>
<td>Pcell</td>
<td>Primary Cell</td>
</tr>
<tr>
<td>PDCP</td>
<td>Packet-Data Control</td>
</tr>
<tr>
<td>PGW</td>
<td>Packet Data Network Gateway</td>
</tr>
<tr>
<td>PHY</td>
<td>Physical</td>
</tr>
<tr>
<td>PLMN</td>
<td>Public Land Mobile Network</td>
</tr>
<tr>
<td>PP</td>
<td>Ping-Pong</td>
</tr>
<tr>
<td>PO</td>
<td>Primary Spectrum Owner</td>
</tr>
<tr>
<td>PU</td>
<td>Primary User</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality of Service</td>
</tr>
<tr>
<td>RAN</td>
<td>Radio Access Network</td>
</tr>
<tr>
<td>RB</td>
<td>Resource Block</td>
</tr>
<tr>
<td>REG</td>
<td>Regulatory databases</td>
</tr>
<tr>
<td>RL</td>
<td>Reinforcement Learning</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>RLC</td>
<td>Radio Link Control</td>
</tr>
<tr>
<td>RN</td>
<td>Relay Nodes</td>
</tr>
<tr>
<td>RR</td>
<td>Round Robin</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>RLF</td>
<td>Radio Link Failure</td>
</tr>
<tr>
<td>RRC</td>
<td>Radio Resource Control</td>
</tr>
<tr>
<td>RAT</td>
<td>Radio Access Technology</td>
</tr>
<tr>
<td>RRM</td>
<td>Radio Resource Management</td>
</tr>
<tr>
<td>RRAT</td>
<td>Radio Resource Allocation Table</td>
</tr>
<tr>
<td>RSRP</td>
<td>Reference Signal Received Power</td>
</tr>
<tr>
<td>RSSI</td>
<td>Reference Signal Strength Indicator</td>
</tr>
<tr>
<td>RSRQ</td>
<td>Reference Signal Received Quality</td>
</tr>
<tr>
<td>RSSI</td>
<td>Reference Signal Strength Indicator</td>
</tr>
<tr>
<td>SAE</td>
<td>System Architecture Evolution</td>
</tr>
<tr>
<td>Scell</td>
<td>Secondary Cell</td>
</tr>
<tr>
<td>SGSN</td>
<td>Serving GPRS Support Node</td>
</tr>
<tr>
<td>S-GW</td>
<td>Serving Gateway</td>
</tr>
<tr>
<td>SIPTO</td>
<td>Selected IP Traffic Offload</td>
</tr>
<tr>
<td>SINR</td>
<td>Signal to Interference Noise Ratio</td>
</tr>
<tr>
<td>SIP</td>
<td>Session Initiation Protocol</td>
</tr>
<tr>
<td>SS</td>
<td>Spectrum Sensing</td>
</tr>
<tr>
<td>SSP</td>
<td>Small cell Service Provider</td>
</tr>
<tr>
<td>SSR</td>
<td>Spectrum Sensing Region</td>
</tr>
<tr>
<td>SU</td>
<td>Secondary User</td>
</tr>
<tr>
<td>TA</td>
<td>Tracking Area</td>
</tr>
<tr>
<td>TDD</td>
<td>Time Division Duplex</td>
</tr>
<tr>
<td>TTI</td>
<td>Transmission Time Interval</td>
</tr>
<tr>
<td>TTT</td>
<td>Time-to-Trigger</td>
</tr>
<tr>
<td>UE</td>
<td>User Equipment</td>
</tr>
<tr>
<td>UL</td>
<td>Uplink</td>
</tr>
<tr>
<td>UM</td>
<td>Unacknowledged Mode</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal Mobile Telecommunications System</td>
</tr>
<tr>
<td>URI</td>
<td>Uniform Indicator</td>
</tr>
<tr>
<td>VCG</td>
<td>Vickrey-Clarke-Groves</td>
</tr>
<tr>
<td>VPLMN</td>
<td>Visited Public Land Mobile Network</td>
</tr>
<tr>
<td>WCDMA</td>
<td>Wideband Code Division Multiple Access</td>
</tr>
<tr>
<td>WiMAX</td>
<td>Worldwide Inter-operability for Microwave Access</td>
</tr>
<tr>
<td>WINNER</td>
<td>Wireless World Initiative New Radio</td>
</tr>
</tbody>
</table>
1. Introduction

Self-organizing networks (SONs) with autonomous operation, control and maintenance have the potential to provide significant improvements in the way the next generation networks such as LTE/LTE-Advanced are managed and configured. It is anticipated that SONs with software defined networking (SDN) and virtualization as enabling technologies will be able to optimize the network based on real time traffic conditions and user requirements with least human intervention and reduced complexity. This would lead to significant expenditure savings for the operators in network deployment, operation and management. SONs will allow network operators to introduce new and innovative services in an easy and fast manner, enabling more diverse business opportunities.

It is a key technology to enable network operation and management in a multi-tenant environment with more granular control of the network, efficient and targeted policy application while providing traffic isolation and security. With SONs, network operators would be able to easily and rapidly test new approaches and ways for network management in real time without any impact on operational traffic therefore would be able to serve users in an efficient way without impacting their quality of service (QoS). Hence, such functionalities in modern communication networks are essential to cope with the boost in mobile data traffic that is expected to grow even further in the coming years.

As the radio access technologies and network architectures are evolving to overcome the limitations of existing networks and address the rising demands and expectations of the end-user, the approach and notion of how these networks are managed and operated is also evolving. The increasing heterogeneity in either access technology, topological layers for example, macro, pico, femto or spectrum (multiple frequency bands, maybe aggregated) has made it necessary for operators to maintain and operate distinct access, backhaul and core networks. This could lead to increased OPEX and CAPEX for the operators. Thus, it clearly indicates that deploying and operating networks in near future should be cost effective and less time consuming. SONs with SDN and virtualization as enabling technology offer an attractive solution for operators to reduce network expenses. Over the past few years SON has received significant attention and strong interest in this domain has been reflected many recent activities. For example, under the framework of Next Generation Mobile Networks (NGMN) Alliance, network use cases have been formulated for different stages of network roll out, [94] namely, network planning (e.g. eNode B configuration parameters), network deployment (e.g. transport parameter settings), network optimization (e.g. capacity, coverage, performance driven radio parameters) and network maintenance (e.g. software upgrades and conflict management).

While NGMN mainly provides economical and technical guidelines [95], [96], 3GPP is responsible for the standardization of related network optimization procedures, components and interfaces. Relevant 3GPP specifications related to LTE-SON are covered by different releases. LTE Release 8 includes fundamental specifications for LTE systems that are currently deployed by most operators [97], [98]. In addition to LTE system fundamentals, it also cover basic specifications for home eNodeB (HeNB) components, procedures to self-establish network equipment and automatic neighbor relation list management [99]. However, SON algorithms themselves are not standardized and have not been included in Release 8 [100]. Specifications for enhanced HeNB and studies focused on self-organization for HeNB, self-organized coverage and capacity optimization and self-healing have been included in LTE Release 9 specifications. 3GPP Releases 8 and 9 specifies some of the most important SON related objectives which includes interference control, coverage optimization (CCO), mobility load balancing, mobility robustness optimization (MRO) and energy saving mechanisms for homogeneous topology. Information regarding further developments, use cases together with required functionalities, evaluation criteria, expected results impact on specifications and interfaces, etc. is reflected in [101]. In 3GPP releases 10 and 11, SON for LTE-Advanced heterogeneous networks (HetNets) was recommended for technical requirement study and work items related to standardization is expected to start soon. Specifications related to LTE system features; especially with respect to future LTE-Advanced systems are being continuously extended in further 3GPP releases (10 onwards).

In addition to 3GPP and NGMN Alliance activities, several other initiatives have been established to investigate and contribute self-optimization and self-configuration in next generation wireless networks [102]. The project named End-to-End Efficiency, funded by European Union (EU)-Seventh Framework Program (FP7), covers some SON-related use cases, such as handover optimization and Intercell interference coordination [103]. Similarly, the ICT-FP7 Socrates project defined detailed classification and SON use case taxonomy [109]. It also consider SON aspects, such as integrated handover parameter optimization, load balancing, automatic generation of initial insertion parameters and cell outage optimization. Another project named the Celtic Gandalf under the Celtic initiative [104] contributed at a very early stage, for example to automated troubleshooting and automatic tuning of network parameters [105].
2. Architectural aspects for SONs and cognition for device to device communications

The increasing user trend of wireless devices utilization (e.g. smartphones, tablets) and the need of the operators’ side to encounter its demands caused the booming development of the networks towards Long Term Evolution – Advanced (LTE - A) in order to provide reliable mobile services for the fourth generation (4G) technologies. According to a report that was conducted in 2013 from Group Speciale Mobile Association Intelligence (GSMA Intelligence) ([1]), smartphone intrusion in Europe is one of the most statistically dominant among regions around the world through years (Figure 2.1). In 2013, the levels of smartphones’ penetration in Western Europe, where the largest European markets of EU5 (UK, Italy, Spain, Germany and France) state, were almost 65% and is predicted to increase more in the upcoming years.

![Smartphones' penetration in market according to GSMA](image)

Mobile users, influenced by the attractive prices of the offered services from mobile broadband providers and their coverage reliability, increase mobility conditions and drive networks into encountering an enhanced volume of data that arises in time. In year 2017, the estimated data traffic is going to correspond at almost 1.5 million TBs per month for Western Europe and 0.8 million TBs for Central and Eastern European countries, which means almost 5 times over 2013 (Figure 2.2). However, the upgrade and evolution of physical infrastructures towards LTE-Advanced include some beneficial facts for handling this issue and its consequences. The most significant advantages that this networking type offers are the ability of increasing data rates (peak data rate up to 1 Gbps) and leverage more spectrum through the Carrier Aggregation (CA) property, enhance spectrum efficiency via advanced MIMO (Multiple Input / Multiple Output) antenna techniques and, last but not least, increase cell capacity by introducing small cells to offload the macro-cell base station.

The heterogeneous nature of the evolved networking systems is characterised by different access type components integrated into one compound entity. The general operators’ idea is to build on a strong interaction and co-operation between every element that belongs to the backhaul and Radio Access Network (RAN) in order to achieve better performance and cost-efficient solutions. In order to manage the total data volume by introducing innovative methods and additional components to the existing infrastructures, several works have been made to accept the concept of small cells, based on 3GPP standards (micro / pico-cells and other low power cell types), and Wi-Fi to help the operational aspects and offload of macro-cells.
One of the most intriguing Wi-Fi techniques that is gaining continuous interest not only for the benefits that offers but also for the researchers’ effort to propose its soft injection to the LTE systems is that of the Device-to-Device communication (D2D) concept. Device-to-Device communication is expected to become a promising component for IMT-Advanced cellular networks. As these heterogeneous networks are characterized by high bit rates and high Quality of Service (QoS) for myriads of subscribers, D2D offer some important advantages to this direction, most of them described in the below sections.

In this type of communication, two user equipments (UEs) are directly connected and communicate between each other (Figure 2.3) by using either the cellular spectrum (i.e. inband) or the unlicensed spectrum (i.e. outband), unlike the traditional communication via the base station (BS).

The scope of introducing such a technique is to enable new serving opportunities and offload cellular eNBs by exploiting the advantages that the local, short range wireless peer-to-peer communications can offer. In section 2.1, we provide an overview of D2D communications and cite some of the most important works that were proposed in the research community.
2.1 State-of-Art analysis

Several works concerning D2D communications have been carried out the last years. The topic of interest was not only the architectural and communication aspects of it, but also the functional improvement of specific networking issues in cellular networks such as spectral efficiency, power control, interference management and resource allocation schemes in order to achieve good overall system performance in a cellular wireless network that is supported by the LTE-based device-to-device communication. In sub-section 2.1.1 we briefly describe the D2D concept in its current status and we continue with some reference to existing literature regarding interference management, power control and spectrum-based works, whereas in sub-section 2.1.2 we concentrate on resource allocation and prove through a series of proposals the significance of finding an optimal, or more feasibly, a near-optimal way of assigning radio resources for transmission purposes in cellular networks.

2.1.1 Device to Device communications

Current data traffic is mainly met on video sharing applications and social networking activities that require high exchanging rates but could possibly be helped and controlled by the proximity gain that a D2D connection can offer. Other than high bit rates, the short distance between two mobile users operating in a cellular network can provide ameliorated channel conditions, lower delays and more efficiency abilities leading to better communication conditions. Additionally, the existence of both D2D and cellular pairing facilities can lead the system to enhanced spectrum utilization, as well as the advantage of opting between cellular and D2D mode is offered. As the operation of D2D communication does not require any hardware installation or change, the operational expenditures remain stable, while from the mobile user’s side new charging policies can take part. Except for the proximity gain that depicts the main characteristics of a D2D connection, there are three other gains that enable fast and many times more secure access to the available spectrum. One of them is the hop gain, where the transmission is executed in one step rather than two hops (communication via the base station). Furthermore, the aforementioned facility of choosing between cellular or D2D mode can offer additional reliable wireless services in a local level and it is referred as pairing gain. Last but not least is the reuse gain, a very significant property of this type of communication, where cellular and D2D links can share the same radio resources at the same time, if this is allowed by the interference levels caused by the involved user equipments.

2.1.1.1 D2D Operational Aspects, Benefits and Challenges

Device-to-Device (D2D) communication has motivated the operators’ interest not only for ameliorating specific technical aspects of the LTE-based cellular network, but also for establishing novel ways for reducing operational expenses (OPEX). However, the conventional D2D technologies that are used until nowadays, such as Wi-Fi and Bluetooth working at the unlicensed band of 2.4 GHz, are currently inadequate as they require manual pairing in order to manage communicational needs between two devices and also they are not integrated into the cellular network (they work in exempt band). Additionally, their quite small distance and the possibility that many wireless devices exist within this range sometimes complicates users’ communication ability. Also, traditional D2D technologies could possibly not be able to meet the user demands due to some technical limitations that cellular networks impose. Finally, concerning revenues, wireless operators might not have any profit by using these technologies as they work independently ([2]). So far, wireless operators have not used these D2D technologies in standardization purpose such as in Global System for Mobile Communications (GSM), Universal Mobile Telecommunications System (UMTS) or 3rd Generation Partnership Project (3GPP) LTE and consequently LTE Release 8 and further (LTE-Advanced). Lastly though, due to the advent of new services and applications (many of them are described in a below paragraph), this seems to be shifting towards D2D standardization and the research interest is being motivated by the bunch of advantages that could be able to provide to future system performance. D2D communication technology can be categorized into two control modes considering the level of control, when operating in a cellular network: (a) fully-controlled and (b) loosely-controlled D2D communication. In the first mode, the cellular network (CN) applies full control over D2D communication setup and operation, including both user and control plane functions, such as resource / power allocation and connectivity setup, respectively. Therein, cellular and D2D links can share the same spectrum in the licenced band (i.e. underlay property) and are able to spatially reuse the available radio resources. The resource allocation procedure is executed either dynamically or semi-statically by associating dedicated resources with each D2D connection. On the other hand, in the second category, operators execute the access authentication process for the D2D deployed UEs, D2D UEs can operate independently and most of the times without the intervention of the operators. Therein, D2D communications can make use of either the unlicensed band (Wi-Fi or Bluetooth) or to be assigned dedicated carriers on the licensed band.

There are some notable advantages that can be able to provide the LTE-A cellular networks with helpful means to enhance their overall performance. These advantages are stated in the following list:
Radio resources reuse leads to an improved spectral efficiency and, consequently, better network throughput.

- Offloaded cellular network, saving processing effort from the base station.
- Capacity enhancement.
- Improved efficiency in terms of power and energy.
- Extended network coverage.
- Low communication delays and higher data rates between the directly communicating users.
- Low energy consumption.

However, the operation of such communication type is accompanied by some challenges that many of them are still open and need to be solved in order to provide efficient and reliable coexistence between D2D and cellular devices. The most important of these challenges are: Interference management, power control, resource allocation and spectrum management, capacity-based improvement, self-organization and cognition of the direct links.

The control of the abovementioned issues and the satisfaction of the requirements that the LTE-A networking standards pose is the subject of current interest for the research community that investigates the compatibility and expandability of D2D communication in cellular OFDMA-based networks.

In the next paragraph, we briefly describe some protocols and initiation steps to establish a connection between two D2D devices.

### 2.1.1.2 D2D Communication Activation

As it can be easily conceivable, in order to softly inject device-to-device services in the LTE-Advanced networks, some architectural additions are needed. Current literature offers two methodologies of D2D communication enabling based on Session Initiation Protocol (SIP) and Internet Protocol (IP). Having already moved from the circuit-switched to the packet-switched way of connection, System Architecture Evolution (SAE), the core network architecture of 3GPP’s LTE-A standard, is characterized by the cooperation of the mobile management entity (MME) and the Packet Data Network (PDN) gateway to take care of the UE’s context by setting up the SAE bearers and the connectivity and communication tunnels between the UEs and the serving PDN gateway. Session setup which is depicted in the figure below (Figure 2.4) requires the according steps ([3]):

1. The session is initiated by one of the deployed user equipments.
2. PDN-GW detects IP traffic to be served within the same serving eNB or between eNBs that serve different cells.
3. If the traffic fulfils specific criteria, the gateway recognizes it as candidate D2D traffic.
4. eNB requests from the user equipment to check if D2D communication offers better performance in terms of throughput.
5. If step 4 holds, and both D2D users support allow D2D connectivity, eNB sets up a D2D bearer.
6. eNB(s) keeps the SAE bearer between the UE and PDN-GW for cellular transmissions, even if the D2D communications and connection setup is successful.
The serving PDN-gateway (PDN-GW in Figure 2.4) is able to initiate a D2D session by detecting the potential D2D traffic (it knows which UE is served by which eNB due to IP processing). Also, a D2D session can be directly initiated via the users and/or applications by selecting the SIP Uniform Indicator (URI) format with a local extension like .direct or .short etc. with the URI of the destination UE. Afterwards, the eNB requests from the controlled UE to measure not only if the D2D devices are in communication range but also if this type of communication offers better throughput. If the aforementioned conditions are fulfilled, the eNB sets up a D2D bearer to allow direct communication between the two involved UEs. Therefore, there is no need to involve the eNB or the SAE gateway into the routing procedure as UEs can transmit via the D2D connection.

2.1.1.3 D2D Usage cases and Applications

The emergence of new applications and user entertainment demands such as video sharing, content distribution and location-aware services (e.g. Foursquare on Facebook) resulted in the introduction of different use cases for the Device-to-Device communications operating in a cellular network. D2D communication was firstly considered to support multi-hop relay in cellular networks (i.e. MCNs) ([2]), where the transmission in multiple steps via several mobile stations is allowed in order to optimize the total network throughput. Only recently the existing literature focused on other significant usage cases like video dissemination, multicasting, peer-to-peer communication (P2P), machine-to-machine communication (M2M), cellular offloading and other cases.

Doppler et al. ([3]) investigate a traffic scenario where D2D users exchange big content such as videos and big data files that derive from a local media server which contains mainly video files and general information material. They study a scenario of a LTE-Advanced cellular network that operates in the band of 100 MHz, divided (equally) into five sub-bands, and they distribute a number of cellular and D2D users in each cell. Their objective is to prove that, under specific assumptions, D2D cooperation with the typical cellular mode can result in enhancing the total system throughput in the deployed area.

In terms of multicasting, an innovative hybrid automatic repeat-request (HARQ) mechanism is proposed in [4] to provide a reliable multicast transmission scheme for D2D communications integrated into the cellular networks and, in order to signify the supremacy of their work, they compare it with the traditional D2D multicast techniques.

Apart from relaying purposes, D2D is considered to be a very helpful and innovative technology to provide wireless peer-to-peer services, where the transmission endpoints are D2D-supportive mobile devices. D2D assists the system by offloading the cellular network as two proximate end users (in terms of distance) can become independent from the base station’s control and communicate directly between each other. A specific example of the aforementioned ability is any photo, video or meeting/conference downloadable material exchange through the
users’ smartphones or other wireless devices in an indoor environment. Furthermore, the use of D2D is “hidden” in video games’ multi-play mode, which is a co-operative gaming mode where a number of friends can play together using local wireless connection. Also, another example of such use case is the widely known “check-in” on Facebook, a mean of location-aware social networking where the user reveals his presence at places using a mobile website, text messaging or selecting from a list of “venues” that the application locates close to him. Such kind of users can interact among each other.

As device-to-device communication is gaining continuous interest, it was about time for researchers to conceptually combine it with the machine type communication (MTC) perspective, where machines (not only UEs but also other devices) have the ability to directly communicate with each other. Pratas and Popovski describe a representative example in [5], where they propose two network-assisted D2D-based mechanisms for enabling the cooperation between existing cellular devices and machine-type devices that are supported by D2D links and they prove the gains that these schemes can offer.

Furthermore, compliant with the idea of D2D communication to encounter the cellular overload rather than relying on more expensive solutions of adding extra hardware equipment or spectrum, Bao et al. [6] investigate the idea for mobile operators to sample the locations of the deployed users and create a data spot map that includes information about where Wi-Fi hotspots are located and are available for each user in its coverage area. Therefore, while the operators can retrieve the available content for each user when someone enters a specific spot, they can schedule peer-to-peer data exchange to “listen” from devices within the spot.

![D2D use cases paradigms](image)

**Figure 2.5:** D2D use cases paradigms (from the work of Lei et al. [2]).

Looking into the industrial field, several companies, among which a representative colossus is LG Electronics, fight for commercializing and introducing the concept of D2D connection in order to provide public safety communications. In case of cellular network malfunction or infrastructure problem, D2D communications, provided that they smoothly coexist with the cellular links in time and frequency, can replace the conventional cellular linking method and offer an alternative and secure option for the system recovery and proper operation.

### 2.1.1.4 D2D Research Classification

Taking into consideration the available literature, D2D communications can be conceptually divided and classified in two major categories: inband and outband communication. The incentive to choose the first type of communication is most of the times the high control over cellular spectrum (i.e. licensed spectrum), as the interference in the unlicensed spectrum can become rampant and deteriorate the system performance and degrade the Quality of Service (QoS) levels (can be further explained in [7]).

Inband D2D communication can additionally be classified into the underlay and overlay category. In the first category, cellular and D2D links can use the cellular spectrum (i.e. they share the same radio resources) in order to
enhance the spectral efficiency, whereas in the other category the BS of each cell decides and dedicates specific cellular radio resources to the D2D links. The main drawback of the use of inband D2D is the resulting interference that may occur due to the coexistence of D2D and cellular pairs, but, through a series of research proposals that will be referred and compared in the next sub-chapters, the levels of interference can be mitigated or, optimally, cancelled by introducing some innovative resource allocation techniques.

On the other hand, outband D2D communication utilizes the available unlicensed spectrum. The objective of using such type of communication is to encounter any interference issue between a cellular and a D2D link that cause intra-cell degradation. However, as easy as it may sound to confront with such issues by introducing the outband method, there is a trade-off regarding cost and complexity. This complexity arises from the need to adopt other wireless technologies, like Bluetooth or Wi-Fi and give the control of this technology to the cellular network (i.e. controlled outband D2D). This condition means that the devices involved in this methodology should support both technologies and might require a hardware update concerning their operation leading to an extra cost.

In this section, we will summarily describe some of the challenges and research subjects that were studied and most of them remain at the centre of attention. Compliant with the research flow, as most of the works so far were handling the underlaying inband concept of D2D communication, we review below the current state-of-art approaches to get familiar with the covered and uncovered research issues.

2.1.1.4.1 Interference Management

One of the most significant aspects in underlaying inband D2D communications is the proper management of interference levels between the cellular and D2D users that reuse the same radio resources. Interference mitigation/cancellation methods have been proven to be very important while they enhance the system capacity without changing the network infrastructure. In the D2D transceiving procedure, interference can be possibly eliminated from the receiver by modulating the received interfering signal and then cancel it ([8]). Yet, this technology is characterized by its complexity in terms of modulation and the limitations considering the total robustness to multiple transmitters. Additionally, the challenge of contriving the proper modulation scheme for achieving the best recovering results (i.e. Forward Error Correction - FEC) is still open and under investigation. In Error! Reference source not found., Min et al. proposed a novel interference manipulation scheme to improve the reliability of D2D receivers and, generally, the communication during the uplink period without mitigating the transmitting power of cellular users. They introduced three receiving methods for reliably demodulating when D2D UEs share cellular radio resources. The first method demodulates the desired signal by firstly treating the interference as noise. In the second mode, an interference cancellation (IC) scheme is preceding and then demodulates the desired signal. Lastly, in the third mode, the proposed solution is to require the retransmission of the interference from the base station to the D2D receiver in order to generate the desired interference increase and afterwards apply the interference cancellation process to obtain the desired signal. Their criterion for choosing the best receive technique for any interference case has to do with the minimum outage probabilities among them and they conclude by suggesting the first mode for low interference regimes, the second for high interference cases and the proposed third for the middle interference regimes.

Before the aforementioned, another interference cancellation technique was proposed in [10] by allocating a dedicated control channel for all D2D users. This scheme comes to make the uplink resources sharing understandable and try to avoid near-far interference between D2D and cellular users in case of D2D transmission in a hybrid network. After the cellular user listens to the signalling of this common control channel (D2D CCCH) and measures its SINR, the eNB monitors the acquired information and decides the cease of scheduling cellular UEs on the resource blocks that are used by the D2D UEs, just in case the measured SINR is above a predefined threshold. Then, the eNB broadcasts the allocated resources and the position of the cellular UEs in the dedicated channel and helps D2D UEs avoid using the same resource blocks and therefore interference issues. Assuming a scenario with 200 uniformly distributed cellular users and 50 cluster-based D2D users within a cell, their simulations show that they can enhance the average system throughput by 374% in comparison to a no-interference-cancellation scheme. In [11], Peng et al. investigate a similar problem of a hybrid system with two cellular users and one D2D pair that reuses the same resources and address the uplink interference issue by proposing two solutions to avoid the resulting interference among users. In this work, D2D pair traces the interference caused by all UEs as they already know the resource blocks allocated to each one of them and, consequently, they can avoid using the same resource blocks allocated to close-distanced cellular UEs. Their proposal is based on transmitting the expected interference from D2D communication on a specific cellular resource block to all D2D deployed users.

After this, D2D UEs adjust their transmitting power and resource block allocation scheme in a way that interference is tolerable. Again, in comparison to a no-interference-based scheme, the system throughput is
The two lastly referred modes are very useful for interference mitigation needs. By the time that the D2D links do not cause interference to the distributed cellular users and vice versa, they can use their maximum transmitting power to provide the best performance. Yet, the issue of which mode is mostly preferable is still open and under research, as these modes may not utilize optimally the available resources but, conversely, degrade the total network throughput and QoS standards.

Not so many studies have been made so far regarding the mode selection in D2D communications. In [15], Doppler et al. propose an algorithm for the optimal selection of possible sharing modes in a single-cell scenario (with one cellular user and one D2D pair), taking into consideration the interference effect for each possible mode and the D2D and cellular link quality, and then expand the case for multi-cell environments. Here, they do not consider only the received signal strength over the D2D communicating link or the distance criterion for all terminals. They introduce a cellular rate restriction to prioritize the cellular user and select the optimal mode that provides the highest among all sum-rates and at the same time satisfies the SINR constraint of the cellular network. Even though their results show that their method provides reliable D2D communication, the authors did not utilize any power control mechanism to ensure acceptable interference levels. On the contrary, [16] analyses the mode selection issue along with a power control technique and optimum resource allocation for the involved users that are subjected to spectral efficiency and energy restrictions. Yu et al. defined the sum-rate of the D2D and cellular communications by applying the Shannon capacity formula and with no specific assumptions on the background cellular networks. They firstly study a greedy sum-rate maximization problem where the two communication types compete services. Then, they set some priorities to the cellular users for obtaining the available radio resources with a guaranteed minimum transmission rate. They conclude by pointing out that the optimum power control and resource allocation for the aforementioned sharing modes can be found by solving the optimization problem in closed form or searching from an infinite set.
2.1.1.4.3 Power Control

Power efficiency has lastly become a subject that attracts the interest of the research community in inventing new algorithms and mechanisms to mitigate interference among all users in a cellular LTE-A network that is supported by the concept of D2D communication. By introducing a proper power allocation scheme, operators can be able to improve their network’s spectrum utilization and provide efficient coordination among all users, where more D2D UEs can share the same radio resource. An important factor that should not be overlooked is also the energy efficiency, since mobile user equipments are characterized by limited battery energy and should be able to maintain it in good levels to perform their operations.

One of the first and most representative works done so far is that of [17], where the authors propose a simple power control scheme to limit the interference levels by reducing the D2D devices’ transmission power in a cellular network. In their simulations Yu et al. considered a single-cell scenario where D2D communication underlays the cellular network. They firstly formulated the SINR distribution of both cellular and D2D link by assuming uniform distributions of the cellular and D2D users and then applied their power reduction method. The same authors proceeded few months later ([18]) with an analysis of the same network environment and studied power optimization in order to maximize the sum-rate (i.e. maximization problem of the ratio between the power optimized case and the fixed transmit power case) in non-orthogonal and orthogonal sharing cases between cellular and D2D UEs. In [19], Gu et al. propose a dynamic scheme to achieve a proper power management for a D2D transmission link to reduce interference and improve the overall system performance. By excluding the cellular UEs’ communication (surrounding interference problems) that needs to utilize the same with the D2D link resource, the deployed eNB periodically adjusts the transmitted power of the D2D link. This power is estimated after using the channel gains between cellular and D2D UEs. After several different LTE parameter settings and considering various channel models, the authors’ results show that their proposal can increase the average SINR of both cellular and D2D users and therefore enhance the overall network performance. In [20], Xiao et al. propose a heuristic algorithm that performs power optimization with joint resource allocation and mode selection in an OFDMA cellular network that supports D2D connection. They firstly allocate the resources for the cellular users and then retain the same procedure and mode selection for the D2D users. If the required power level for the D2D transmission is higher than a predefined threshold, D2D UEs connect via the base station. Their simulations end up proving that the downlink power consumption is improved at a rate of round 20% in no-D2D communication case.

An algorithm studying power-efficient mode selection and power allocation of D2D communication that is integrated in cellular networks is proposed in [21], based on the following steps: firstly, they calculate the optimum power levels for all possible modes of each deployed device with respect to maximum power efficiency and they infer the highest power efficiencies of all devices from the sub-optimal solution described before. In the next step, the mode selection is based on the maximum power efficiency among all devices’ mode combinations that were already estimated in the step before. They compared their technique with conventional low power selection, forced D2D and forced cellular schemes (only D2D and cellular mode, respectively) and their simulation results show that their algorithm outperforms the conventional schemes in terms of power efficiency, average transmission power and capacity. However, the major disadvantage of this algorithm is the high complexity due to the need of calculating all possible combination of modes for all devices.

In [22], Belleschi et al. developed a model of joint mode selection, resource allocation and power allocation (JOMSRAP) via Mixed Integer Linear Programming (MILP) that is proven to be a NP-hard problem to solve. Even though they firstly focused on a multi-cell scenario to implement their algorithm, they then switched to single-cell due to the complexity of the problem. In this scenario, In order to solve this issue, they come up with a heuristic algorithm and consider a distributed sub-optimal one that jointly performs mode selection and resource allocation and they compare it, as aforementioned, with forced D2D and cellular schemes. After multiple simulations, their results show that in close-distanced D2D users the power efficiency gain over conventional cellular networks is prominent in a percentage of up to 100%.

2.1.2 Resource Allocation

In order to improve scalability in D2D communications and increase the spectral efficiency of the network, reliable resource allocation mechanisms are required. Current research works consider mainly only one D2D link for their investigations, whereas there are some exceptions (e.g. the work of Zulhasnine [23]) that occupy the cellular network with multiple D2D pairs. In this section we will refer to the existing resource allocation-based literature and analyse in chronological sequence some of the most important studies in this subject, starting from the chronologically oldest one.
In [12], Janis et al., as already mentioned in sub-section 2.1.1 (paragraph 2.1.1.4.1), assume a time division duplex (TDD) cellular network, where the base station of each cell uses identical splitting for the uplink and downlink resources that are assigned to N cellular terminals. D2D transmissions are synchronized and integrated in this network and occupy the cellular spectrum by reusing the orthogonal radio resources (either on downlink or uplink phase) that are assigned for each cellular user. Their assumption is based on the existence of the same number of D2D users in each cell (i.e. N) and on the simultaneous communication of the BS with the UE and that of the D2D devices. Their objective is to assign the sub-bands to each D2D link and choose properly the transmitted power of the D2D pair such that its SINR is maximized and there are no interference violations from the cellular link that transmits at the same time. In their scheme, D2D devices calculate the received interference power of cellular devices’ transmission and receive their channel measurements from the regular UL transmissions. Based on these assumptions, they proceeded with a resource allocation mechanism to reduce intra-cell interference levels in concurrent transmissions in both UL and DL phases. Their results, compared to conventional methods, show that they can succeed almost 2.6 times better system performance in terms of cell capacity. Furthermore, they made an additional comparison with random resource allocation and achieved a 4.7 dB gain in the 10th percentile of downlink (DL) SINR. However, one significant disadvantage of their work is that their case seems to work properly for a small number of terminals and has some compatibility issues with more of them.

On the other hand, the authors in [23] include multiple D2D pairs in their network model, integrated in a LTE cellular network. After referring to the D2D architectural injection into the base network, they formulate the problem of spectrum sharing to D2D communications as a mixed integer non-linear programming (MINLP) that is very difficult for it to be optimally solved in fast scheduling period. Their proposal is based on a greedy heuristic algorithm that can mitigate the interference levels in the primary cellular network by exploiting the channel gain information. This algorithm selects the proper radio resource for which channel gain between the cellular UE receiver and D2D transmitter is lower for the downlink phase (same scenario for the uplink procedure). Results show that their algorithm improves the total network performance in terms of D2D throughput without deteriorating the cellular networks operation in a harmful way.

Chae et al. ([24]) is the first research group to introduce the concept of Fractional Frequency Reuse (FFR) in D2D communications integrated into cellular network in order to mitigate inter-cell interference caused by users that transmit in the outer region of each cell (Figure 2.6). Looking at this figure, they partition the cell in two different areas, the inner and outer area, with different frequency reuse factors (FRF) for each one of them (FRF = 1 for inner area and FRF = 3 for outer area). Cellular UEs that are positioned in the inner region and close to their controlling eNB are immune to co-channel interference to from the other cells and use resources that are reused by every cell (F1 sub-band). On the other hand, outer cellular UEs can experience interference issues by one or more neighbour cells if they use the same frequency sub-band and that is the reason why the frequency band is partitioned in four pieces (exclusive allocation).

In their scheme, they allocate different resources for the D2D users depending on their different location. If D2D UEs are positioned into cell’s inner are, they can use the sub-band that cellular UEs don’t use (e.g. in Figure 2.6, in Cell 1, the D2D pair that is located in the inner region cannot use the same frequency band that cellular UE in the same region uses; in this case, D2D pair can utilize F3, F4 and not F1, F2 due to interference issues). Also, the D2D UEs that are located in the outer area of a cell are able to use the sub-bands that are not utilized by cellular UEs.
UEs in identical cell outer region. In summary, their proposal is trying to control interference that is developed between cellular and D2D UEs in two different cases: a) when the latter is located in the inner area of the cell, and b) when D2D can be in the outer cell region, where the interference levels between D2D and cellular users is tolerable. Their simulations resulted in a significant amelioration of the total throughput of LTE cellular networks compared to a random resource allocation scheme.

In [25], the authors proceed with a study that D2D UEs can reuse the same radio resource of one cellular user or more. Their study criterion is the sum-throughput of cellular and D2D communication by applying the capacity—based Shannon formula and the comparisons are made among different resource allocation schemes to find the optimal one. They firstly apply resource reuse of only one cellular UE and calculate the sum-throughput and then with more than one cellular user by guaranteeing at the same time the minimum transmission rate of cellular UEs with higher SINR threshold for both D2D and cellular communication. According to the authors, the optimal resource allocations is the one that maximizes the sum-throughput of cellular and D2D communication. Their simulations are made for both UL and DL. D2D resource reuse and achieve better performance for both cases in terms of sum-throughput, especially when the D2D is very close to the cell edge and the distance between the D2D paired users becomes smaller and the cell radius becomes bigger.

In 2012, Xu et al. ([26]) propose the reverse iterative combinational auction (I-CA) as a resource allocation method for multiple D2D pairs underlaying cellular communication. Their proposal focuses on the improvement of sum-rate for the D2D sub-network by reducing the effect of interference between cellular and D2D users. In this auction concept (auctioneer (i.e. the base station) – bidder (i.e. spectrum resources occupied by cellular UEs) game), the spectrum units are allocated to N user packages, consisting of at least one D2D pair. That means that this spectrum units compete in order to obtain D2D communication for improving the channel rate. They formulated this problem as an integer linear program wherein the objective is to maximize the overall channel gain by guaranteeing at the same time some D2D constraints. Their allocation algorithm is based on collecting the location information for all the D2D pairs and proceed with a non-monotonic price loop process. They conclude that they manage to improve the system sum-rate by increasing both the number of D2D pairs and that of the resource units, compared to the random allocation scheme. In [27], the same authors propose a novel algorithm for spectrum allocation to ameliorate the performance of D2D communication in the downlink (DL) procedure of a cellular network. They proceeded from their latter work by introducing a sequential algorithm (called second price auction) to optimally allocate the available radio resources in a single-cell scenario with multiple users. They formulated the value of each unit (i.e. resource unit) for each deployed D2D pair and implemented the aforementioned algorithm using an N-ary tree. The final allocation step is represented by the leaf node of the defined tree, and their simulation results show somehow the same algorithmic performance with their first proposal, leading to superiority in comparison with the random allocation in terms of sum-rate.

Resource allocation schemes for D2D communications underlaying cellular networks, as mentioned before, support the spatial reuse of resource blocks and can be grouped into two different categories: the centralized resource allocation (e.g. [23] and [25]) and the distributed resource allocation scheme (e.g. [22]). In the first case, the base station fully controls the resource blocks’ allocation for all links (i.e. cellular links and D2D links) provided that it knows the channel qualities of each sub-channel for all the links. In the second category, the resource allocation of each D2D link is operated in a distributed way, whereas cellular links are traditionally controlled by the related cell base station.

Opposed to this static-decision researches, the authors of [28] study a combinational semi-distributed resource allocation mechanism, where the base station allocates resource blocks (RBs) to D2D pairs in a centralized way, but at the same time the modulation and coding scheme (MCS) and the transmission power levels of D2D links are distributively decided by the D2D device of each link. This distributing attribute of their scheme contributes to the BS computational complexity offloading, as part of the overall allocation decisions (i.e. allocations in D2D links) is removed from the BS and transferred to D2D primary devices. Furthermore, the overhead concerning the channel reporting requirements is reduced due to the fact that only the BS and not all user equipments need to know the path loss caused in all involved links. Also, compared to the centralized way of allocating resources, here the control overhead on the BS becomes less because, in the first case, the BS needs to broadcast lot of information such as MCS, transmission power and the resource block assignment results. Their problem formulation aims at maximizing the spatial reuse of the available resources by allowing simultaneous transmissions of the D2D links on the same resource blocks. The provided solution is based on converting the problem from an optimization problem into a set covering problem and, in order to acquire its solution, they apply a greedy algorithm for it.

The multiple set covering problem tries to estimate the minimum number of feasible simultaneous transmission patterns (FSTPs) that include each D2D link for a restricted number of times. In their proposed greedy algorithm
for solving the aforementioned problem, they select at each stage a FSTP covering the largest number of D2D links that is not already covered sufficiently. It is based on a constant loop that generates FSTPs, allocates the smallest number of RBs with a generated FSTP such that at least one D2D link satisfies the resource allocation requirements. While reducing the control overhead, their results after multiple simulations show that this proposed scheme can achieve higher network throughput (i.e. sum of data rates of all receiving nodes in the deployed cellular network). Also, in proportion to the number of D2D deployed links in a single-cell scenario, they manage to improve the reduction percentage of control overhead for the network.

The authors’ proposal for optimal resource sharing for D2D communication in a cellular network concentrates on the uplink case resource allocation. Considering a scenario of N orthogonal cellular users and one D2D link controlled by the cell’s BS, they try to design non-orthogonal resource sharing strategies where the D2D link utilizes in the optimal way all the possible cellular resources without causing any cellular communication degradation and keeping the signalling overhead and the total complexity in controllable standards. Their goal is to maximize the throughput of the D2D link (as aforementioned, in most of the works seen so far, network throughput is the main objective in D2D communications underlaying a cellular network and the optimization problem can be met in many different expressions) and at the same time satisfy the cellular users’ constraints. This could be mathematically formulated as follows:

\[
\begin{align*}
\text{Maximize:} & \quad \sum_{i=1}^{N} R_{i} \sum_{j=1}^{M} d(p_{i,j}, q_{i,j}) \\
\text{Subject to:} & \quad R_{i} \leq \rho_{i}, i = 1, \ldots, N \\
& \quad 0 \leq p_{i,j} \leq P_{i}, \quad 0 \leq q_{i,j} \leq Q_{i}, \quad i = 1, \ldots, N \\
& \quad \sum_{i=1}^{N} N_{i,j} = Q_{i}.
\end{align*}
\]

where \(\rho_{i}\) and \(P_{i}\) is the QoS threshold and the power budget for the \(i_{th}\) cellular user, respectively; \(Q_{i}\) is the transmission power limit of D2D user in the \(i_{th}\) frequency band; \(Q\) is the sum power budget of the D2D UE on all frequency bands. After solving this problem, they proceed with two simple sub-optimal resource allocation schemes. The first scheme keeps the transmission power of cellular users in a fixed level (i.e. maximum transmitted power) and, with some alterations to the equation and the QoS constraints stated above, and reduces complexity and the signalling overhead of the joint optimization problem. On the other hand, in the second case scenario, they study the resource sharing scheme between a D2D link and only one cellular user controlled by the BS and their conclusions show that there are not many differences in terms of complexity compared to the joint scheme.

One of the most spherical research works in the literature is that of [29]. Phunchongharn et al. introduce an innovative joint resource allocation scheme that takes into consideration power control and RB scheduling processes. After providing the research community with an analytical view on the issues and background of D2D communications underlaying a cellular network and the effort that is put on introducing them in LTE-A systems, they proposed a resource allocation algorithm for D2D communication based on a column generation mechanism in order to optimize the spectrum efficiency by minimizing the transmission length (i.e. minimum number of timeslots used), with keeping interference towards cellular users in low levels and satisfying at least the minimum D2D QoS limitations. They formulated their problem into a NP-complete mixed integer programming (MIP) problem that finds the minimum number of timeslots to allocate all users in a feasibly scheduled pattern.

However, due to the high complexity of this problem as it contains both continuous and integer variables, they decided to change the solution method by searching it through a column generation method (Figure 2.7). This method is known for solving in an efficient way large integer programming problems with different type of variables. The concept of it is that the optimization problem is decomposed into two sub-problems: a) the restricted master problem (i.e. considers only one subset of the feasible access patterns (simultaneous transmission of D2D links in the same RB) whereas the original MIP problem takes into account all access patterns), and b) the pricing problem. The restricted master problem considers only traffic demand constraints, whereas the pricing problem considers the feasible access pattern constraints. Firstly, they solve the restricted master problem in order to obtain the dual variables (which provide the optimization problem with useful knowledge) and use them for the pricing problem.
Afterwards, they try to retrieve new column variables, meaning a feasible access pattern, to improve the current objective function in the first problem type. The new price, called reduced price, has to decrease the current objective value and therefore, in order to improve the objective function, the new price must be negative and satisfy the constraint $\lambda q_s > 1$, where $\lambda$: dual variable vector and $q_s$: new column variable vector. This algorithm keeps being executed until $\lambda q_s < 1$, i.e. the pricing problem provides a non-negative new price. This idea with the problem formulations and modifications for better solving results contributed in increasing the spectrum utilization by scheduling multiple D2D links to spatially reuse the same resource. The only significant drawback of this algorithmic implementation was that in a deployed network with large number of users distributed in it, the power consumption increases, not enormously, but affecting somehow the energy efficiency of the system.

In [30], F. Wang et al. take into account UEs’ fairness and network throughput and propose the proper collaboration between scheduling and resource allocation in order to ameliorate as much as possible the performance of the D2D underlay communication. They proceed with a framework based on Stackelberg’s game where a cellular and a D2D user represent a leader-follower pair of users and the follower (i.e. D2D user) “buys” channel resources from the leader (i.e. cellular UE). In this “trade”, a leader owns a channel resource and can sell it to a follower (i.e. D2D UE uses the channel allocated for the cellular UE). Therefore, the leader tends to sell as much as possible, that means to share radio resources with the D2D UE, and he is the one to decide the “price”. The D2D UE’s incentive is to choose the optimal power to maximize its payoff towards the leader. The concepts of leader and follower are analytically described in this paper. For every TTI, a cellular and a D2D UE construct between each other a leader-follower pair and proceed to the game. The pairs that are formed create a priority queue and the eNB schedules the D2D pairs sequentially. Their algorithm is characterized by its low complexity and good performance for both cellular and D2D UEs in terms of throughput.

Continuing with some gaming concepts, Xu et al. ([31]) expand their research work in [26] and propose an auction-based resource allocation scheme to improve overall system throughput and optimize the resource sharing for both cellular and D2D users. In the auction (i.e. this is how they call their allocation mechanism), all the radio resources are considered to be a set of resource units which compete like bidders do to obtain business while the packages of the D2D pairs are auctioned off and get sold as goods. They formulate the valuation of each D2D pair for each resource unit and then continue with the development of a non-monotonic pricing auction algorithm that depends on the utility function that represents the channel gain from D2D and the costs for the system. Finally, their paper results in the improvement of sum-rate compared again to random allocation scheme.

In [32], the authors analyse the advantages that D2D communications can offer to enable better performance on wireless multicast services in a cellular network, as an efficient way to propagate the same content simultaneously to different recipients within the network in different applications (e.g. mobile TV). However, due to the existence of several different channel conditions, the recipients of a CN can hardly ever be all satisfied in terms of rate. Also, there is a possibility that the achievable multicast throughput degrades when most recipients have good channel conditions and received high transmitted rate, something that could cause few recipients to become a bottleneck for the throughput. Thus, Zhou et al. studied the effect of D2D communication concept in such a network to offer...
the ease of retransmitting from a recipient with good channel condition to another one with degraded channel characterization (i.e. intra-cluster retransmissions and multicast – Figure 2.8).

![Figure 2.8: D2D Intra-cluster re-transmissions with two multicast senders](image)

Their problem of research is how to perform reliable D2D retransmissions to achieve maximum network throughput with the least usage in terms of time and frequency. In summary, their work can be divided in two main steps: a) resource utilization efficiency analysis for different numbers of D2D re-transmitters and provision of mathematical calculations for how the minimum resource cost varies in proportion to D2D channel conditions and the number of relays; after analysing this relationship, they derive in closed-form expression a probability density function (PDF) for an optimal number of D2D re-transmitters that can prove the improvement of resource allocation efficiency depending of the aforementioned number, and b) proposal of an optimum intra-cluster D2D relaying scheme to ameliorate the cellular system’s throughput. Their algorithm selects the number of re-transmitters that retransmit to one or more receivers and uses an iterative sub-cluster partition algorithm to improve data rates of these transmissions. By performing their scheme executions, they realize an approximate gain of 40% in resource allocation efficiency in comparison to a scheme that does not include adaptation of the number of re-transmitters.

On the other hand, one of the most representative works for Device-to-Device communications that overlay, though, a cellular network is that of [33]. Unlike the main workflow that investigates D2D as an underlay to cellular networks’ infrastructures, Pei et al. propose a spectrum allocation scheme where D2D users sub-help the cellular users’ transmissions by operating as relays between the BS and a cellular user and they are allowed to communicate directly between each other. This concept can be parallelized with the concept of cognitive networks’ spectrum overlay, where secondary users (SUs) facilitate primary users (PUs) relay their signals and obtain the opportunity to transmit when needed. Then, we presume that D2D users are like SUs and cellular users are like PUs and a proper coordination among them is needed to ensure good channel knowledge. Their proposal is met on allowing D2D users to communicate bi-directionally between each other and, at the same time, help and secure the cellular transmissions either in UL or DL procedure. They firstly performed an information-theoretical study on the achievable rates for both cellular and D2D transmissions. Then, they estimate the so called Pareto boundary (i.e. optimal boundary estimated by the solution of an optimization problem with multiple criteria decision making limitations and where more than one objective function has to be optimized) by optimizing the transmission power at the BS and the cellular users, as well as the power splitting factor at the relay D2D node. Additionally, they consider the case of selecting the proper D2D relay form a potential pool to execute the relaying transmission. By using this protocol, they managed through proper power control at the BS the cellular users to improve, at a first stage, the total sum-rate for both types of users and then achieve better performance through the proper selection of D2D relays.

2.2 Current research workflow

2.2.1 Scenarios and methodology

In order to follow the research flow concerning D2D communications, we represented in MATLAB a single-cell scenario of a centre-deployed base station (BS), some cellular users (4 in our main implementations, communicating with the BS) and a number of uniformly distributed D2D users (16 in our case, directly linked into pairs) in a square area of 1km X 1km area (Figure 2.9).
The criterion for creating the links, either cellular or D2D, is based on the distances among all users. For each one of the 8 D2D pairs, one is the sender and the other one is the receiver located in an inter-distance that is not bigger than 120 m and ranges in [20,120]. The path-loss is modelled according to the following equation: \( PL(d) = PL(d_0) + 10n\log_{10}(d/d_0) \), where \( d \) is the distance between D2D sender and receiver that construct a D2D pair, \( d_0 = 0.1 \text{ km} \) is the reference distance, \( PL(d_0) = 40 \text{ dB} \) is the path-loss estimated for the reference distance and \( n \) is the path-loss exponent (3 in our paradigm).

After calculating the SINRs and distances among all involved linked nodes, we applied S-TDMA (spatial TDMA - [34]) as a collision-free scheduling scheme that allows any timeslot reuse by different radio links (if permitted by the resulting interferences) and tried to minimize the frame length in order to enhance performance by leading the system to enhanced spectrum efficiency.

The assumptions made for the aforementioned simulation are the following:

- Inter-cell interference is not existent (in case of multi-cell network) and could be managed efficiently with the inter-cell interference control (ICIC);
- Each user (network node) can either transmit or receive only towards/by only one different node.

Based on this simulation, we assessed the credibility of an existing methodology – algorithm in [35] that exploits the physical interference model and consists of the following decisions:

1. Sorting of the existing links, depending on a pre-defined criterion (Interference / Power level / Number of Receivers)
2. Greedy link packing into timeslots to achieve feasible (without inter-link interferences) schedules.

We distinguished the first step of the algorithm into three (3) different sub-algorithms, namely Greedy Physical (GP), Packing Heuristic (PH) and Number-of-Receivers-based Packing (NRP) that are based on the same strategy but are characterized by different sorting criteria:

- **Greedy Physical (GP)**: links are sorted in decreasing order according to their related interference number.
- **Packing Heuristic (PH)**: links are sorted in decreasing order based on higher transmitted power requirements.
- **Number-of-Receivers-based Packing (NRP)**: links are sorted in decreasing order, starting from the link that its transmission range encloses the maximum number of receivers.
2.2.2 Performance analysis

Despite the differentiation of the defined criteria (section 2.2.1), the results of the simulations for each timeslot scheduling method in relation to the number of deployed links are similar to each other, as it can be easily observed in the below figure (Figure 2.10).

![Figure 2.10: Timeslot scheduling depending on the number of links](image)

The rationale behind the few differences for these methods is, firstly, the random distribution of the deployed users that does not ensure stability for the simulations (for example, NRF sub-algorithm can achieve better results only in the case where due to randomization and big inter-link distances several links can be packed together) and, secondly, the similar nature of the consequences that the used criteria lead the system’s performance.

2.3 Future research planning

For our short-term objective we have implemented a tier-1 network, constituting of seven (7) cells, that contains several cellular and D2D users, uniformly distributed in it (specifically, the total number of users per cell is 20 – 30% of them are D2D pairs), as seen in Figure 2.11 (red dots represent the cellular users and green and black dots are the D2D Tx and Rx UEs respectively).

![Figure 2.11: 7-cell scenario with total number of 20 D2D and cellular UEs per cell](image)
Focusing on the above figure, we aim at following the below listed steps:

1. Separation of each cell to two main regions (inner region / outer region).
2. Possible sectorization of the outer region to three sectors using Fractional Frequency Reuse with FRF other than 1.
3. Sub-carrier allocation for all users, based on a dynamic sub-carrier allocation algorithm called DSA.
4. Frequency (sub-carrier) packing / scheduling investigation.

Based on the aforementioned list, we will try to investigate also the case of inter-sector and inter-cell D2D pairing cases and how can this links be scheduled both in time and frequency in order to enhance the overall network performance in terms of throughput and resource efficiency. Table 2-1 shows the parameters that are going to be used for our simulations.

<table>
<thead>
<tr>
<th>Simulation Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>eNB transmitting power</td>
<td>43 dBm</td>
</tr>
<tr>
<td>D2D transmitting power</td>
<td>8 dBm</td>
</tr>
<tr>
<td>Antenna direction</td>
<td>Omni</td>
</tr>
<tr>
<td>Noise Power Density</td>
<td>-174 dBm/Hz</td>
</tr>
<tr>
<td>Noise figure</td>
<td>5 dB</td>
</tr>
<tr>
<td>Inter-D2D pair distance</td>
<td>120 m</td>
</tr>
<tr>
<td>Inter-eNB distance</td>
<td>1.5 km</td>
</tr>
<tr>
<td>D2D link distance</td>
<td>[20,55]</td>
</tr>
<tr>
<td>Number of cells</td>
<td>7</td>
</tr>
<tr>
<td>Carrier frequency</td>
<td>2 GHz</td>
</tr>
<tr>
<td>System bandwidth</td>
<td>10 MHz</td>
</tr>
<tr>
<td>Number of sub-channels</td>
<td>50</td>
</tr>
<tr>
<td>Sub-channel bandwidth</td>
<td>180 KHz</td>
</tr>
</tbody>
</table>
3. Architectural aspects for SONs and cognition for handovers optimization

3.1 Handover Procedures - State of the Art

Handover (HO) is a procedure which is considered to be vital for the proper operation of the cellular networks. Specifically, HO is used to ensure that a user who is moving towards the coverage area of a node different from the one that is already camped in, will maintain the same or better level of service. It consists of two parts, the decision and the execution part. During the first part, the serving or source node decides on whether a HO should be triggered to a candidate cell or not in order to satisfy the demands of a user. The decision part is user assisted given that the source cell decides which is the best candidate according to measurements periodically performed and reported by users. Three important metrics related to handover are Reference Signal Received Power (RSRP) [1][37] which is the linear average over the power contributions in W of the Resource Elements (RE) that carry cell-specific reference signals within the considered measurement frequency bandwidth. Reference Signal Strength Indicator (RSSI) [1][37] that indicates the total received wide band power from all cells or users in the area of a user, including thermal noise and noise generated in the receiver. The difference between RSRP and RSSI is that RSSI includes the received power not only from a specific target cell, but also from all interfering nodes in the proximity area of the cell. Finally, the Reference Signal Received Quality (RSRQ) that is a parameter used to describe the quality of the received by the user signal. RSRQ is a fraction where the nominator is the RSRP received from the target cell and denominator is the total signal received by every cell in the proximity of the user (RSSI). As mentioned in [37], RSRQ is used to calculate the Signal to Interference Noise Ratio (SINR) that a user can get if served by a specific cell.

In LTE-A networks, the HO decision is mainly based on the neighboring cells’ level of the received RSRP signal at the UE side (conventional RSRP algorithms). However, other parameters such as Hysteresis Margin (HM) and Time-to-Trigger (TT) which are characterized as HO parameters, are considered of major importance to HO procedure [38]. The integration of these parameters assists network on taking HO decisions based on events observed over a period of time. Specifically in order for the source cell to trigger a HO, the observed RSRP level at the UE side of the best neighbor cell, should be higher than the RSRP of the source cell plus a HM and this condition should be valid for at least a period equal to TTT. Thereby network does not make decisions based on instantaneous measurements, which are not reliable due to the presence of shadow fading and fast fading, and avoids unnecessary HOs that would occur in the absence of the HO parameters and would drive the system to unstable and undesired conditions and performance. The HO event described is the A3 event and will be further analyzed in subsection 3.1.6.

3.1.1 Network Entities and Interfaces

In this subsection the definitions of the entities that participate in the HO procedure according to the architecture defined by 3GPP [38] are presented, along with the interfaces that are used to support interconnection between them.

**Evolved UMTS Terrestrial Radio Access Network (E-UTRAN):** It consists of a single type of nodes (eNB-E Node B/HeNB-Home E Node B) that interfaces with the user equipment (UE). The eNB hosts the PHYsical (PHY), Medium Access Control (MAC), Radio Link Control (RLC) and Packet-Data Control (PDCP) layers that include the functionality of user-plane header compression and encryption. It also offers Radio Resource Control (RRC) functionality corresponding to the control plane. It performs many functions including radio resource management, admission control, scheduling, enforcement of negotiated uplink (UL) Quality of service (QoS), cell information broadcast, ciphering/deciphering of user and control plane data and compression/decompression of Downlink (DL)/UL user plane packet headers [39][40].

**Mobility Management Entity (MME):** The MME is the key control-node for the LTE access-network. It is responsible for idle mode UE tracking and paging procedure including retransmissions. It is involved in the bearer activation/deactivation process and is also responsible for choosing the Serving Gateway (S-GW) for a UE at the initial attach and at time of intra-LTE HO involving Core Network (CN) node relocation. It is also responsible for authenticating the user by interacting with the Home Subscriber Server (HSS). The Non Access Stratum (NAS) signaling terminates at the MME and it is also responsible for generation and allocation of temporary identities to UEs. It checks the authorization of the UE to camp on the service provider’s Public Land Mobile Network (PLMN) and enforces UE roaming restrictions. The MME is the termination point in the network for ciphering/integrity protection for NAS signaling and handles the security key management. Lawful interception of signaling is also supported by the MME. The MME also provides the control plane function for mobility between
LTE and 2G/3G access networks with the S3 interface terminating at the MME from the SGSN. The MME also terminates the S6a interface towards the home HSS for roaming UEs [39][40].

**Serving Gateway (S-GW):** S-GW is used for routing and forwarding user data packets and also acts as mobility anchor for the user plane when inter-eNB HO is performed. Moreover it can be used as the anchor in the cases where UEs are moved between LTE and other 3GPP technologies. If a UE is in RRC_IDLE state, S-GW terminates the DL data path and triggers the paging procedure when DL arrives for the UE. Moreover, it manages and stores UE contexts, such as parameters of the IP bearer service and network routing information, and performs replication of the user traffic in case of lawful interception [39][40].

**Packet Data Network Gateway (PDN Gateway):** The purpose of PGW is to connect the UE with the external PDNs. The fact that a UE might access more than one PDNs means that a UE can have simultaneous connectivity to multiple PGWs. PGW works as the point of exit and entry of traffic for the UE and is responsible among others for policy enforcement, packet filtering for each user, charging support, lawful Interception and packet screening. One important operation performed by PGW is that it works as the anchor for mobility between 5GPP and non-3GPP technologies such as WiMAX, CDMA 1X and EvDO [39][40].

**S1 Interface:** S1 is a logical interface that is specified at the boundary between the Evolved packet Core (EPC) and the E-UTRAN. eNBs are the E-UTRAN access points from the S1 perspective and the EPC access points are MMEs and S-GWs. As a result there are two different S1 interfaces, the one that connects an eNB with an MME (S1-MME interface) and the one that connects an eNB with the S-GW (S1-U interface). The main purpose of the S1 interface is to support the communication between eNBs and EPC. Analyzing more the purpose of S1 we should mention that it supports:

- Procedures to establish, maintain and release E-UTRAN Radio Access Bearers;
- Procedures to perform intra-LTE HO and inter-RAT HO;
- The separation of each UE on the protocol level for user specific signaling management;
- The transfer of NAS signaling messages between UE and EPC;
- Location services by transferring requests from the EPC to E-UTRAN, and location information from E-UTRAN to EPC;
- Mechanisms for resource reservation for packet data streams.

The S1 as it was mentioned provides support for HO of users in RRC_CONNECTED state and by mentioning the term "Handover" we refer to Intra-LTE HO, Inter-RAT HO and mobility to CDMA2000 systems which will described in 3.1.5 [40][44].

![Figure 3.1: S1 Interface Architecture](image)

**X2 Interface:** The X2 interface is a logical interface that enables the interconnection of eNBs in the E-UTRAN [41]. Some of the basic characteristics that X2 has are:

- Supports the exchange of information between two eNBs that might have been supplied by different manufacturers.
- Is a logical point-to-point interface between two eNBs which should be feasible even if a physical direct connection between the two eNBs is absent.
- It can support services offered via the S1 interface
between eNBs.
- It can support radio interface mobility between eNBs (HO) of UEs in RRC_CONNECTED state in the E-UTRAN.

There is an amount of functions that can be served over the X2 interface. The one that we are interested in and study in this section is the support for Intra-LTE Mobility (HO) which can be translated to many individual functions like context transferring between source eNB and target eNB as we will explain later. HO cancellation, UE context release in source eNB and control of user plane transport bearers between source eNB and target eNB. Other functions that can be supported by X2 interface are Load Management, Inter-cell Interference Coordination (for the uplink and the downlink), Error handling functions, Trace functions, Application level data exchange between eNBs and Data exchange for self-optimization.

### 3.1.2 UE States

HO is a procedure controlled by RRC protocol layer, a layer which exists in both UEs' and eNBs' protocol layer stack. RRC is responsible among others for:
- Broadcast of System Information related to the non-access stratum (NAS);
- Broadcast of System Information related to the access stratum (AS);
- Paging;
- Establishment, maintenance and release of an RRC connection between the UE and E-UTRAN;
- Security functions including key management;
- Establishment, configuration, maintenance and release of point to point Radio Bearers;
- Mobility functions;
- QoS management functions;
- UE measurement reporting and control of the reporting;
- NAS direct message transfer to/from NAS from/to UE.

HO occurs when a UE is in a mobility state called RRC_CONNECTED, one of the two RRC states that a user can be, the other state is called RRC_IDLE.

A UE in RRC_IDLE state is not connected to a cell because there is no available radio link. A user might also be switched to RRC_IDLE state when it is connected and there is no traffic, so a transition from RRC_CONNECTED occurs in order to save radio resources and not to consume valuable battery life. Despite being in RRC_IDLE state, a user can be detected by the network in case of an incoming call because network knows every time the area where the user is located with precision of a number of few cells, this area is called Tracking Area (TA). Finally it is important that a UE in this power-conservation mode does not inform the network for every cell change it performs so excessive signaling is avoided [38].

When a UE is in RRC_CONNECTED state has an E-UTRAN - RRC connection and this actually means that the network can transmit or receive data to or from the UE. Moreover, network knows the exact location of the UE at cell level precision and is responsible for handling cell changing procedure (HO) according to measurements reports made by UE as mentioned before. This explains why HO in LTE is considered as network-controlled, UE-assisted hard HO. The term "hard" is used because when a UE in RRC_CONNECTED state is about to change serving cell through the HO process, what happens first is that the old radio links in the UE are released, and the new link is established with the new cell. This explains why this type of HO is also known as break-before-make. The following figure Figure 3.2, shows the basis steps followed when a UE is handing off to a nearby cell [38].
3.1.3 Handover Types

Hard HO can be seamless or lossless (non-seamless) and is related to the type of service that a user demands [42][43]. Seamless HO occurs when the type of service that the UE demands is tolerant to losses but less tolerant to delays (e.g. VoIP, voice services). It is applied for user plane radio bearers mapped on RLC Unacknowledged Mode (UM) and its main purpose is to minimize complexity and delay since no security context is exchanged between the source and the target cell during a HO operation. However, it may result in loss of some SDUs because as it is mentioned in [43], PDCP SDUs that haven’t started transmission by UE will be transmitted to target cell directly and PDCP SDUs for which the transmission has already started but have not been successfully received yet, will be lost. Finally, PDCP SDUs which have not yet been sent from the source cell will be forwarded via X2 interface to the target cell. The above explain why seamless HO is applicable to services which are more tolerant to losses.

Lossless HO is used when delay tolerant but loss sensitive data need to be transferred (e.g. file download). This means that even the loss of one PDCP SDU can result in drastic reduction in the data rate due to the reaction of the Transmission Control Protocol (TCP). In lossless HO, which according to [43] occurs in inter-eNB and intra-eNB transitions, in-sequence delivery of the packets is ensured by performing retransmission of PDCP SDUs for which reception has not yet been acknowledged prior to the HO.

There are many reasons to perform a HO, however the most obvious and usual cases where a HO is necessary are the following:

- Avoid having highly congested macro cells by offloading traffic/load to nearby less congested cells
- Avoid having radio link failures due to bad channel quality from serving cell
- A UE is moving towards a different cell so it’s more likely that the new cell can offer him better QoS.

In LTE there are three types of HO according to whether HO happens within the current LTE nodes (Intra-LTE), between LTE nodes (Inter-LTE) and between nodes that belong to different radio access technologies (Inter-RAT).
It is common to find the first two categories under the same name, Intra-RAT HO, since both source and target nodes use the same radio access technology. Further differentiation occurs for the two first categories due to the fact that the HO can be executed through the use of X2 interface or through the use of S1 interface [44] [45].

- **Intra-LTE**: HO happens within the current LTE nodes (intra-MME and Intra-SGW)
- **Inter-LTE**: HO happens toward other LTE nodes (inter-MME and Inter-SGW)
- **Inter-RAT**: HO between different radio technology networks, e.g. GSM/UMTS and UMTS

Although Intra-RAT mobility seems quite reasonable since it concerns handoff of UEs between nodes that operate on the same radio access technology, further explanations should be given on whether Inter-RAT mobility worth to be studied and in which cases Inter-RAT mobility occurs.

Inter-RAT HO occurs when a UE needs to be served by a cell which supports different technology from LTE. This could happen:

- Because LTE will be mainly available in urban dense areas where the demands are higher in terms of traffic load and less tolerant in terms of time delay. As a result, HO between systems of different generation should be feasible and without causing problems such as call drops which usually happen during HO procedures.
- Moreover, Non-real time data which can be served by earlier cellular systems (e.g. UMTS) are considered of equal importance with real-time data, so the effort to avoid drops while executing HO is the same.

![Figure 3.3: Intra-RAT and Inter-RAT mobility](image)

### 3.1.4 Handover Procedure in detail

In this section an analytical description of the HO procedure between two eNBs, in terms of signaling between the entities that participate each time, is given.
In figure 2 a detailed description of the intra-MME/Serving Gateway HO procedure is displayed and an analysis of the messages and the main functions is followed [59] [44].

The main messages and processes are described below:

Messages 1/2; (Measurement control/report) The source eNB configures and triggers the UE measurement procedure and UE sends periodically measurement reports.

Messages 3/4; Source eNB makes the HO decision based on the information received from measurement reports and sends a HO request to the target eNB.

Messages 5/6; Target eNB performs admission control according to the QoS that each UE demands and afterwards, informs source eNB with the HO Request ACK/NACK which includes HO parameters such as security algorithm identifiers and new Cell Radio Network Temporary Identifier (C-RNTI) for the target cell. It might also contain RNL/TNL information and a dedicated RACH preamble among others.

Message 7; Source eNB transfers the RRCConnectionReconfiguration message (HO command) to the UE in order to perform the HO. The RRCConnectionReconfiguration message contains the aforementioned HO parameters and is generated in the target cell.

Message 8; The source eNB sends the SN STATUS TRANSFER message to the target eNB. This message contains among others the uplink PDCP sequence number, the Uplink Hyper Frame Number, The downlink PDCP sequence number and the Downlink Hyper Frame Number.

Figure 3.4: Message Sequence Diagram of Handover Procedure for Macro to Macro case [59]
Messages 9-11; After receiving the RRCConnectionReconfiguration message, the UE detaches from source cell, performs synchronization to target eNB and accesses the target cell through a RACH contention-free or contention-based procedure which will be described later. The response of the target cell consists of a message that includes UL allocation and timing advance. If the UE has accessed the target cell successfully, it sends the RRCConnectionReconfiguration message, along with the C-RNTI, in order to confirm that the HO is completed.

Message 12; The target cell informs the MME with a PATH SWITCH REQUEST that the UE has changed serving cell.

Messages 13-15; S-GW receives from MME a MODIFY BEARER REQUEST message, which purpose is to switch the path of downlink data to the (new) target side. After completing the switch, informs the MME with a MODIFY BEARER RESPONSE

Message 16; MME informs the target eNB for the successful PATH SWITCH REQUEST with a PATH SWITCH ACKNOWLEDGE message.

Message 17; The UE CONTEXT RELEASE message is sent from the target eNB and informs source eNB about the successful completion of the HO procedure.

Message 18; Upon reception of the UE CONTEXT RELEASE message, the source eNB releases the radio and control resources associated to the UE context.

As mentioned, UE can access the target cell through a RACH contention-free or contention-based procedure [43][44].

The contention-based random access procedure is initiated when a UE is in one of the following states:

- A UE in RRC_CONNECTED state, but not uplink-synchronized, needing to send new uplink data or control information (e.g. an event-triggered measurement report)
- A UE in RRC_CONNECTED state, but not uplink-synchronized, needing to receive new downlink data, and therefore to transmit corresponding ACKnowledgement/Negative ACKnowledgement (ACK/NACK) in the uplink.
- A UE in RRC_CONNECTED state, handing over from its current serving cell to a target cell
- For positioning purposes in RRC_CONNECTED state, when timing advance is needed for UE positioning
- A transition from RRC_IDLE state to RRC_CONNECTED, for example for initial access or tracking area updates
- Recovering from radio link failures.

Briefly, a UE in one of the aforementioned states tries to access the target cell with a contention-based random access procedure by sending a request over the shared medium using a preamble signature which was chosen randomly. The drawback of this procedure is the possibility of having collisions between requests from different UEs - especially when the number of UEs in the same area is high - that might transmit the same signature simultaneously. The adoption of contention-based random access can possibly result in delays compared to the contention-free random access procedure where the contention between UEs is eliminated by allocating a dedicated signature to each UE that tries to get access to a cell.

### 3.1.5 Intra-LTE, Inter-LTE and Inter-RAT Handovers

In this subsection we present in more detail the different types of HOs that might occur and were mentioned in an abstract way in 3.1.3. As we notice further differentiation occurs according to whether the HO procedure is controlled by the S1 interface or the X2 [45].

**Intra-LTE (Intra-MME/SGW) Handover Using the X2 Interface**

This procedure is used when a UE handoffs from a source eNB (S-eNB) to a target eNB (T-eNB) through the use of the X2 interface. During this HO the MME and the S-GW remain unchanged. The main advantage that the X2 interface offers is that it allows direct connectivity between the source (serving) and the target nodes and the involvement of EPC in the HO procedure is not needed. This results to time efficiency since all the messages are exchanged directly between source and target eNB, mainly during the HO preparation time, but also after the successful HO where the target eNB triggers the release of the resources which were allocated at the source side.
Intra-LTE (Intra-MME/SGW) Handover Using the S1 Interface

Intra-LTE HO occurs when a Handoff from a S-eNB to a T-eNB happens with the use of S1-MME interface due to the absence of a X2. The reasons could be the fact that the X2 connectivity to the T-eNB might not exist, the possibility that after a previous unsuccessful X2-based HO the T-eNB informed the S-eNB for errors over the of the X2 interface or the probable identification of any kind of problem on the X2 interface by the S-eNB during STATUS TRANSFER procedure. The Intra-LTE HO is initiated at S-eNB side with a HO request message which is send over the S1 interface to the MME and then passes to T-eNB. The role of MME in this type of HO is to help the communication between S-eNB and T-eNB since no direct connectivity between the first two entities exist and so it doesn't play any role in the decision part which is strictly defined by S-eNB.

Figure 3.5: Intra-LTE (Intra-MME/SGW) Handover Using the X2 interface
Reference source not found

Figure 3.6: Intra-LTE (Intra-MME/SGW) Handover Using the S1 interface
Reference source not found
Inter MME Handover Using the S1 Interface (Without Changing S-GW)
Inter-MME HO happens when a user who is connected to a source-eNB that belongs to a specific MME (source-MME), moves to an area controlled by a target-eNB that belongs to a different MME (target-MME). What remains unchanged in this scenario is the connection of both MMEs to the same S-GW.

Inter-MME/SGW Handover Using the S1 Interface
The difference of this scenario compared to the previous is that source-MME and target-MME are served by different S-GWs, so one more level of complexity is added to the communication between source and target eNB.

Inter-RAT Handover: E-UTRAN to UTRAN Iu Mode
During this LTE-to-UMTS Inter-RAT HO, the serving eNB is connected with its MME (S-MME) and S-GW (S-SGW), and the target node (RNC: Radio Network Controller) with its SGSN and S-GW (S-SGSN, S-SGW). Finally both S-GWs are connected to the same PGW. This HO procedure can be separated to two different parts. The first, the Preparation Part, is the time period needed for the resources to be reserved in the target cell (and/or
network) in order to serve the demands of the UE who is about to change serving cell. The second, the Execution part, is the phase where the HO of the UE from the source network to the target is completed.

![Inter-RAT Handover: UTRAN to E-UTRAN](image)

**Figure 3.9: Inter-RAT Handover (E-UTRAN to UTRAN) preparation and execution phase**

Reference source not found.

**Inter-RAT Handover: UTRAN to E-UTRAN**

During UTRAN to E-UTRAN HO, the procedures and the messaging between the source and target nodes are the same as previous with the only difference that the source node operates on UMTS technology and the target cell on LTE technology.
3.1.6 Handover Triggering Events

Different types of HO are triggered by events during the time a UE is in RRC_CONNECTED state. These events are the following and are categorized according to whether they initiate an Intra-RAT HO or an Inter-RAT HO [44]:

**Intra-RAT HO events in LTE**

1. Event A1: Serving becomes better than threshold
2. Event A2: Serving becomes worse than threshold
3. Event A3: Neighbour becomes offset better than Primary Cell (PCell)
4. Event A4: Neighbor becomes better than threshold
5. Event A5: PCell becomes worse than threshold1 and neighbour becomes better than threshold2
6. Event A6: Neighbour becomes offset better than Secondary Cell (SCell)

**Inter-RAT HO events in LTE**

1. Event B1: Inter-RAT neighbor becomes better than threshold
2. Event B2: PCell becomes worse than threshold1 and inter-RAT neighbor becomes better than threshold2.

Figure 3.10: Inter-RAT Handover (UTRAN to E-UTRAN) preparation and execution phase

Reference source not found.
Primary Cell: The cell, operating on the primary frequency, in which the UE either performs the initial connection establishment procedure or initiates the connection re-establishment procedure, or the cell indicated as the primary cell in the HO procedure.

Secondary Cell: A cell, operating on a secondary frequency, which may be configured once an RRC connection is established and which may be used to provide additional radio resources.

These events are triggered when an entering condition is fulfilled and cease to be valid when a second condition, called leaving condition, is active. The aim of the figures that follow is to explain schematically when these events appear. The Y Axis represents the level of RSRP received by the UE from serving or neighboring cell on the same or different frequency (depends on the triggered event) and the X axis is the time. An explanatory list of the variables used in the following inequalities is given to help the reader:

Variables for Intra-RAT HO events in LTE

Ms (dBm, dB) is the measurement result of the serving cell, not taking into account any cell individual offset.

Hys (dB) is the hysteresis parameter for this event

Thresh (dBm, dB) is the threshold parameter for this event

Mn (dBm, dB) is the measurement result of the neighboring cell.

Ofn is the frequency specific offset of the frequency of the neighbor cell (equals Ofs for intra-frequency measurements and is included in MeasObjectEUTRA corresponding to the inter frequency as offsetFreq for inter-frequency measurements)

Ocn is the cell specific offset of the neighbor cell. If not configured, zero offset shall be applied (included in MeasObjectEUTRA of the serving frequency as parameter cellIndividualOffset for intra-frequency measurements and included in MeasObjectEUTRA corresponding to the inter frequency as parameter cellIndividualOffset for interfrequency measurements).

Ofs (dB) is the frequency specific offset of the serving frequency

Ocs (dB) is the cell specific offset of the serving cell

Off (dB) is the offset parameter for this event

Thresh1 is the threshold parameter for this event

Thresh2 is the threshold parameter for this event [38] [44]

Variables for Inter-RAT HO events in LTE

Mn is the measurement result of the inter-RAT neighbour cell, not taking into account any offsets

Ofn is the frequency specific offset of the frequency of the inter-RAT neighbour cell

Hys is the hysteresis parameter for this event

Threshold is the threshold parameter for this event

Mp is the measurement result of the PCell, not taking into account any offsets.

Threshold1 is the threshold parameter for this event

Threshold2 is the threshold parameter for this event
**Event A1** (Serving becomes better than threshold)

Ms − Hys > Thresh (Entering condition)
Ms + Hys < Thresh (Leaving condition)

**Event A2** (Serving becomes worse than threshold)

Ms + Hys < Thresh (Entering condition)
Ms − Hys > Thresh (Leaving condition)

**Event A3** (Neighbour becomes offset better than PCell)

Mn +Ofn +Ocn − Hys > Ms +Ofs +Ocs +Off (Entering condition)
Mn +Ofn +Ocn + Hys < Ms +Ofs +Ocs +Off (Leaving condition)
Event A4 (Neighbor becomes better than threshold)

\[ \text{Mn} + \text{Ofn} + \text{Ocn} - \text{Hys} > \text{Thresh} \] (Entering condition)
\[ \text{Mn} + \text{Ofn} + \text{Ocn} + \text{Hys} < \text{Thresh} \] (Leaving condition)

Event A5 (PCell becomes worse than threshold1 and neighbour becomes better than threshold2)

\[ \text{Ms} + \text{Hys} < \text{Thresh1} \text{ AND } \text{Mn} + \text{Ofn} + \text{Ocn} - \text{Hys} > \text{Thresh2} \] (Entering condition)
\[ \text{Ms} - \text{Hys} > \text{Thresh1} \text{ OR } \text{Mn} + \text{Ofn} + \text{Ocn} + \text{Hys} < \text{Thresh2} \] (Leaving condition)
Event A6 (Neighbor becomes offset better than SCell)
Mn + Ocn − Hys > Ms + Ocs + Off (Entering condition)
Mn + Ocn + Hys < Ms + Ocs + Off (Leaving condition)

In this case the figure is similar to event A3 with the difference that Neighbor becomes offset better than SCell instead of the serving cell (PCell).

Inter RAT HO events in LTE

The events that could trigger an Inter-RAT HO are the following

Event B1 (Inter RAT neighbour becomes better than threshold)
Mn + Ofn - Hys > Thresh (Entering condition)
Mn + Ofn + Hys < Thresh (Leaving condition)
Event B2: PCell becomes worse than threshold1 and inter RAT neighbour becomes better than threshold2

Mp + Hys < Thresh1 (Entering condition 1)
Mn + Ofn - Hys > Thresh2 (Entering condition 2)
Mp - Hys > Thresh1 (Leaving condition 1)
Mn - Ofn + Hys < Thresh2 (Leaving condition 2)

3.2 HetNets and Standardization

The massive use over the last years of smart phones, tablets and generally mobile devices with high capacity and ability to support demanding services such as video demanding, VoIP services, internet navigation etc forced cellular systems to change in order to be able to compensate with the sharp increase in the mobile traffic.

A major change that is related to the architecture of cellular systems, is the densification of the last with low power nodes also named as small nodes or HeNBs. The main characteristics of these nodes are the higher spectrum efficiency that offer which results to better utilization of the spectrum resources and their small coverage areas compared to the macro. Figure 3.18 shows the relation between the coverage area of a HeNB and the bitrate it offers in comparison with macro cell's capabilities.

The introduction of these nodes transformed the cellular networks into HetNets, a term which is used to justify the existence in the network of nodes with different characteristics (eNBs/HeNBs). In addition it emerged the need for further standardization of procedures and proposal of algorithms as we will describe in this subsection and in 3.3, since algorithms used in macro (eNB) only scenarios could not meet the demands of the new architecture in an efficient way. The term HeNB was initially used for nodes which had only indoor coverage and their radius couldn't exceed 20 meters. Now HeNB could refer even to a metro cell with coverage area of 150m.
HeNBs are entities, with unique equipment identity, similar to eNBs meaning that they are deployed by operators (and by individuals in some cases - indoor HeNBs), they use operator's licensed spectrum and are not activated until they receive configuration and authorization by the operator. As it was mentioned they are used for coverage and capacity enhancement of the network by operating in areas with coverage "holes" from the macro cells or in areas with high mobile traffic. The functions supported by the HeNB are the same as those supported by an eNB and the procedures run between a HeNB and the EPC are the same as those between an eNB and the EPC [44].

Three different access modes are defined for HeNBs

- **Open access HeNBs** are the small cells which allow every client of its operator to have a connection with them. An alternative definition is that an open access HeNB appears as a normal eNB. [56] [47]

- **Closed access HeNBs** are those who provide service only to associated Closed Subscriber Group members (CSG). The term CSG refers to an operator's group of subscribers who are permitted to access one or more cells which have restricted access [56].

- **Hybrid access cells** are cells which can service both CSG and non-CSG members. The difference is that normally CSG are prioritized compared to non-CSG members and the charging policy might be different as well. A similar definition is given in [56] according to which Hybrid access HeNBs are those which use part of their resources in open access policy and the rest in closed access policy.

<table>
<thead>
<tr>
<th>H(e)NB Access Mode</th>
<th>Open</th>
<th>Closed</th>
<th>Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>UE allowed access to CSG</strong></td>
<td>Access</td>
<td>Access</td>
<td>Preferential Access</td>
</tr>
<tr>
<td><strong>UE not allowed access to CSG</strong></td>
<td>Access</td>
<td>No Access</td>
<td>Access</td>
</tr>
</tbody>
</table>

Table 3.1: Clarification of HeNB Access [48]

HeNBs which operate in closed or hybrid access mode, broadcast two categories of parameters

- Parameters to support the UE in the identification of closed/hybrid cells
  - CSG Indicator, CSG Identity (CSG ID), HNB Name
- Parameters to support an efficient search of closed/hybrid cells at the UE
  - Range of Physical Cell-IDs (PCIs) reserved for closed cells

Finally HeNBs support Mobility management for different access modes

- Access Control procedures for establishing a connection and for HO
- Differentiating between a member and a non-member at a hybrid cell
- Automatic (re-)selection in idle mode if the CSG ID broadcast by the closed or hybrid cell is in the UE CSG lists
- Manual user selection of a closed or hybrid cell
The network infrastructure to support the operation of small cells did exist in the previous 3GPP Releases. However, what is new in the architecture defined by 3GPP’s latest standards [44] Figure 3.19 is the introduction of X2 interface between macro and small nodes in order to support direct connectivity of these two entities. X2 interface was initially standardized for connection between eNBs and later even for connection between HeNBs. Now, the existence of X2 between macro and small nodes in the framework of Heterogeneous dense networks is very helpful for mobility management of users in RRC_CONNECTED state. HOs will not be very demanding in terms of signaling overheads in the system despite the extended deployment of small cells and the high possibility that a user might perform many HOs while being active. This is achievable because direct X2 interface between source and target cell avoids the involvement of other entities in the HO process. This is not feasible in the presence of a S1 interface as described in 3.1.5.

X2-based HO between HeNBs is allowed if no access control at the MME is needed, i.e. when the HO is between closed/hybrid access HeNBs having the same CSG ID or when the target HeNB is an open access HeNB.

The following table summarizes the HO cases which can be performed through the use of an X2 interface.

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>eNB</td>
<td>eNB</td>
</tr>
<tr>
<td>Any HeNB</td>
<td>Open access HeNB</td>
</tr>
<tr>
<td>eNB</td>
<td>Hybrid access HeNB</td>
</tr>
<tr>
<td>Any HeNB</td>
<td>Hybrid access HeNB</td>
</tr>
<tr>
<td>Hybrid access HeNB</td>
<td>Closed access HeNB</td>
</tr>
<tr>
<td>Closed access HeNB</td>
<td>Closed access HeNB</td>
</tr>
<tr>
<td>Any HeNB</td>
<td>eNB</td>
</tr>
</tbody>
</table>

Table 3.2: Handover Cases Supported by X2 interface [44]

An entity that is important for the proper operation of small cells is the Home eNB Gateway (HeNB GW) as it can be seen in [44].
HeNB GW according to [44][49] is deployed in the E-UTRAN architecture and its main functionality is to assist HeNBs to access the core network. HeNBs can be also connected directly to EPC with the use of S1 interface but HeNB GW offers an alternative, specifically in cases where there is a cluster of HeNBs which we want to support in a scalable manner.

HeNB GW can provide control plane aggregation meaning that can work as a concentrator for the Control plane of a cluster of HeNBs. It can offer registration and routing of signaling between an area with multiple HeNBs and the MME via the S1-MME interface and enables the MME to view this cluster of HeNBs as a single entity. Moreover HeNB GW is visible to HeNBs as a MME. The HeNB GW can also work as a user plane concentrator in case HeNBs are not directly connected to S-GW. If this is the case, then like previously the cluster of HeNBs is visible to S-GW as a single entity. HeNB GW and S-GW communicate with the use of an interface called S1-U, otherwise the S1-U interface is between HeNBs and S-GW.

HeNB GW offers also some other functionalities like paging optimization for the UEs under HeNB coverage. Finally HeNB GW can secure the communication with HeNBs through the use of an integrated Security Gateway [49][44].

What is interesting about the HeNB GW is that it can connect to the EPC in a way that inbound and outbound mobility to cells served by the HeNB GW will not necessarily require inter MME HOs.

The architecture presented in Figure 3.19 supports direct X2-connectivity between HeNBs, independent of whether any of the involved HeNBs is connected to a HeNB GW.

Since the complete network architecture defined by 3GPP was presented in Figure 3.19, summarizing the cases where S1 interface is used along with we have seen in 3.1.1, S1 is used to interconnect:

- HeNB GW with the CN
- HeNB with the HeNB GW,
- HeNB with the CN
- eNB with the CN

The latest standard releases on small cell enhancements "scenarios and requirements" for E-UTRA and E-UTRAN [50][51] provide us with deployment scenarios of small cells with and without macro coverage. The direction of the research at the moment is focusing on dense small cell deployments as mentioned before. These deployments can take advantage of the special characteristics that small cells have and provide an efficient way for the network to cope with the incremental increase of mobile traffic which can be generated indoor or outdoor. In Figure 3.20 we see scenario cases that were proposed by 3GPP for further study since the concept of HetNet is new and the standardization is yet at early stages.

Figure 3.20: 3GPP HetNet scenarios [51]

On the target scenarios defined by 3GPP, small cells are deployed under the coverage area of one or more overlapped macro cells to drive the network to an increase in capacity. Furthermore, there are cases defined where small cells are not deployed under the coverage area of macro cells. The use cases taken into consideration, are:

- The case of having a UE in the coverage area of both macro and small cells simultaneously
- The case of having a UE in the coverage area of only macro cells or only small cells
Moreover there is a separation of indoor and outdoor deployed small cells in order to provide service to UEs moving inside a building or outside. According to the standards again, what is proposed is that indoor small cells should serve users with low speed (0-3 km/h) and outdoor small cells should be responsible for medium speed users (up to 30 km/h).

However this separation of users according to their speed is flexible and they suggest that outdoor small cells could possibly serve users with higher speed than 30 km/h if there are benefits by adopting such policy from the network. On the contrary 3GPP suggests that UEs with even higher speed (50 - 80 km/h) should be treated in a different way. They point out that it could be probable a good solution to exclude the last users from having access to small cells. The rationale behind this suggestion is to avoid frequent HOs, the signaling overhead that they create and degradations in network's performance that might occur.

What is also stated by 3GPP in [51] is the further separation of deployments to sparse and dense. As examples where dense deployment of small cells is needed, they mention dense urban areas or large shopping malls where the generated traffic is usually high and they refer to these cases with the term "hotspot" areas. On the other hand there are some cases/scenarios where the deployment of few small cells is sufficient to serve the demands of the users.

In Figure 3.20 different colors are used to refer to carrier frequency used by the macro and small cell layer. Although it is suggested that different frequency bands should be separately assigned to macro layer and small cell layer respectively, in [50], scenarios also with co-channel deployment of the macro and small cells are described Figure 3.22.

![Figure 3.21: User Case scenario defined by 3GPP [50]](image)

![Figure 3.22: Co-Channel Deployment of macro and small cells [50]](image)
3.3 HetNEts and Handover Procedures

The densification of cellular networks with small cells and the transformation of the last to Heterogeneous due to the deployment of cells with different characteristics like transmitted power, coverage area, resources allocated and maximum number of serving users posed new challenges to them.

One of the procedures that was affected from these rapid changes in cellular environments was the HO procedure. The short radius of small cells and the dense architecture of the network worsens the HO procedure because higher number of small cells are considered as candidates during the HO decision phase and the probability of performing a HO increases with everything that this implies in terms of signaling overhead, wasting of resources and delays.

The conventional algorithms which were used when operators had only macro cell coverage, could not compensate with the new conditions. HO procedures based on RSRP algorithms, were not able to serve effectively the demands of the users who were located in areas with overlapped macro and small cells. As a result, the network could not exploit the advantages and the positive features of small cells. Parameters like the uneven transmission power of macro and small cells, speed and demands of users, interference level caused by the operation of multiple nodes, energy efficiency and available bandwidth were not taken into consideration and the outcome was poor performance of the network.

The need for more sophisticated HO decision algorithms which would take into account parameters like the ones mentioned was emerged to cope well with the mobility management of UEs with medium to high speed and even with low speed.

This resulted to a series of proposals focused mainly on how to increase the utilization of small cells and at the same time to apply strategies which will not allow the HO of fast UEs to small cells, unless it's necessary for maintaining the connection.

The metrics used commonly in literature in order to evaluate the performance of the proposed schemes and algorithms are assignment probability to small cells, Number of HOs that occur in the system (per user per second), Number of unnecessary HOs, Achievable SINR by the users, Aggregated Throughput of the system, Percentage of traffic served by the different tiers and Throughput of the system for different speed of users among others. Most of the proposals in this area try to optimize one or more of the aforementioned metrics.

From the latest standard [52] on mobility enhancements in heterogeneous networks we see that the metrics which are proposed from 3GPP for the evaluation of HetNet mobility performance are mainly the HO failure rate and the ping pong rate.

HO failure is defined as the loss of radio link connection during the HO process and can occur in three cases. The first is when a Radio Link Failure (RLF) timer (called T310) has been triggered or is running when UE receives the HO command and the HO execution time is about to begin (Figure 3.23). The second is when RLF timer T310 expires after the entering condition event but before the HO command is successfully received by the UE (Figure 3.24). Finally HO failure occurs if after the UE is attached to the target cell, the received signal quality from the new cell is below the Qout threshold at the end of HO execution time.

HO failure rate is defined at the ratio of number of HO failures to the total number of HO attempts. Qout is a threshold for the received by the user signal quality under which a timer T310 begins to run and at its expiration UE suffers from radio link failure.

As ping-pong (PP) is defined the handoff of a UE from a cell B to a cell A and then before specific time duration passes, called Time-of-Stay (ToS), handoff again back to cell B. Therefore, ping-pong rate is a fraction with nominator the number of cases where a UE stays connected in a cell for a time less than a pre-determined minimum-time-of-stay (MTS) and then handoffs to the previous cell and denominator the sum of the successful HOs that happened. The value which is recommended by 3GPP for MTS is 1s (Figure 3.25).
As described before, the introduction of small cells in cellular networks inside buildings or outside, with different radius, with low transmit power, which are deployed in a coordinated (controlled by operators) or uncoordinated (by individuals) way in a dense or a sparse architecture affected HO procedures. Literature in this area is quiet rich and focus on different performance evaluation metrics which were also mentioned. As we have seen until now, the conventional HO algorithms used in macro-only deployments had as basis the RSRP signal level from different cells which measured by UEs. However these algorithms were not designed for the cases where HO from macro cell to small cell, or HO between small cells occur. The uneven transmission power between the macro and small cells was taken into consideration by a number of proposal which target on optimizing the HO procedure in HetNets.
In [53] the authors' intention is to maximize the advantages of the use of HeNBs by making them more appealing to UEs even if HeNBs are not located at the cell edge of the macro coverage area. What they achieve is to increase the assignment probability to small cells and at the same time to maintain the same number of HOs that occur in the network. Their proposal consists of an algorithm that takes into account the received RSRP signal from the target HeNB node and the signal received from the serving macro node. The expression is used along with a combination factor which purpose is to balance the uneven level of transmitted power that the two nodes have, specifically in the cases where the small cells are closed to the inner region of the macro cell. As a result, HeNBs are prioritized more for UEs with the drawback that this strategy increases the number of HOs when the distance between macro and small cells is low (up to 250 meters).

An extension of this idea can be found in [54] where the motivation is the same, meaning that the authors are looking for an improvement in the utilization of the small cells due to the fact that they offer higher data rates with a lower - transmission power and management - cost. Furthermore, they want to go one step further and decrease the number of redundant HOs.

In the proposed solution, the authors compare the RSRP signal received from both macro and HeNBs and they also compare the wireless transmission losses from each of the two nodes. The last is done in order to decrease the possibility of having a HO to a small cell that cannot guarantee a certain level of service required by the UE. The outcome is that they manage to increase the assignment probability to small cells compared to the conventional RSRP algorithm and decrease the number of redundant HOs compared to the combinational algorithm used in [53] by 50%.

Summarizing this proposal, an inbound HO (from eNB to HeNB) is allowed when the RSRP signal of the HeNB is above a certain threshold, is also better than the RSRP of the macro cell plus a HO Hysteresis Margin and finally the path losses between UE and the candidate HeNB are lower than the path losses between UE and the serving macro cell.

The concept of Intracell HO (IHO) is studied in [55] and in [56]. What is proposed, is a way to avoid cross-tier interference in two-tier networks. Cross-tier interference is defined as the decline in signal quality of macro cell UEs on the downlink or the uplink due to the presence of HeNB users sharing the same spectrum and vice versa. The proposed scheme consists of an Intracell HO along with a power control technique if it's necessary.

The aim of this algorithm is to cope with cross-tier interference issues if HeNBs follow a closed access policy, or/and the increased number of HOs that appear if HeNBs operate with open access policy. As Intracell HO they refer to the action of transferring one UE from a channel which may suffer from high interference to a "clearer" one without changing serving cell (UEs are still served by the macro cell).

According to the proposal, the macro cell performs an IHO to a non-subscriber macro user (a UE who connects only to macro tier) who suffers from cross-tier interference due to a nearby HeNB. The IHO in the macro cell will fail if there is not any available sub-channel for the UE to be re-allocated, or if the interference that this available sub-channel suffers from, is higher from the interference that currently assigned sub-channel suffers. The alternative for the macro cell is to initiate an IHO to the interfering HeNBs. If again this in not possible because of lack of free sub-channels, then the last measure to decrease the cross-tier interference in the macro tier is to reduce the transmission power of the specific sub-channel in the small cell tier for a small period of time.

Except from the differentiation between the transmission power of eNBs and HeNBs many proposals focus on how to minimize HOs by restricting the access of fast UEs to small cells. The rationale is to avoid handing off to small cell users who will probably stay for a small period of time in the coverage area of that cell. In [57] an algorithm for reducing the unnecessary macro to small cell HOs is introduced. The basic idea is to block UEs who are characterized as temporary HeNB visitors from performing inbound HOs to small cells. The decision on whether to allow or block the access of a UE in a small cell is made according to an algorithm that checks first if the HO conditions are met and then uses a mobility prediction function to estimate how long the UE is going to spend inside the coverage area of the HeNB. If this time is higher than a specified threshold, then the HO is allowed.

The authors in [59] introduce an optimized HO procedure for different types of HO (hand-in, hand-out, inter-HeNB) based on mobility prediction of the UEs. Moreover they propose a reactive and a proactive strategy in order to mitigate the occurrence of frequent and unnecessary HOs. In their proposal, after the RSRP and speed measurement part, the use of absolute speed thresholds allows or permits the access of fast users in a HeNB. The reactive strategy is used to lower the number of HOs occurred (mainly between small cells) by not triggering HO to a target cell until the RSSI signal from the serving cell is almost closed to the minimum accepted value. The objective of the proactive strategy is to minimize packet loss and high latency, specifically for UEs with real-time
traffic demands, and it achieves that by triggering the HO process even sometimes before the RSSI level of the current serving cell reaches the HO hysteresis threshold.

In [58] the problem of signaling overhead due to frequent HOs from fast UEs is introduced in a macro to small cell scenario. The proposal consists of an algorithm which lets inbound HO only if the RSRP signal that the UE receives from the HeNB is higher than the macro cell RSRP plus a Hysteresis margin, and secondary if the mobility state of the UE is below a specified threshold. Otherwise, the UE stays connected to macro. The outcome of this proposal is that it drives the system to a reduction in the HO signaling overhead which is even more notable in the case where the number of high mobility UEs is increased.

An algorithm for reducing unnecessary HOs is proposed in [59] for a single macro cell - single small cell scenario in hybrid access mode which will result in further reduction in HO failures. The proposal takes into account a wide range of parameters; Interference level at small cells, UE speed, RSRP signal received, Quality of Service. The cases under study in this proposal are the hand-in for Closed Subscriber Group UEs, hand-in for non-Subscriber Groups and finally hand-out to a macro cell.

In more detail, a UE served by a HeNB handoffs to a macro cell if the UE speed is above a defined threshold and the target macro cell can serve the demands of the UE, or if the RSRP signal received from the HeNB is continuously decreasing and the macro can satisfy the bandwidth requirements of the UE.

On the other hand a UE served by a macro cell hands in to a HeNB if the UE is a non-CSG then hands out if the interference level in the HeNB is above a threshold, if its speed is lower than a speed threshold and if the HeNB can serve the demands of the user. If the UE is a CSG then in order to hand out to the macro, an RSRP criterion must be met, the speed should be lower than a high threshold and lower than a second lower threshold. If the last criterion is not met, then the UE should demand a real-time service and finally the macro cell should be able to cover the bandwidth requirements of the UE.

In [60] an idea for cutting down unnecessary HOs and increasing the Throughput of both high speed users and eNBs is presented. The scenario under study consists of a macro cell with overlaid indoor HeNBs located in a hotspot area and UEs passing through this area. The solution that the authors propose is a set of rules that compare the speed of UEs with speed thresholds and also compare the RSRP signal received from macro or HeNBs with the lowest RSRP signal received at UE side each time.

The idea introduced in [61] wants to drive the system in better performance in terms of number of unnecessary HOs and number of HOs that occur. The algorithm proposed in this paper has two parts. The first part is a cost based function that accounts for parameters such as the RSRP signal that the UE receives from target cell, the capacity of the target small cell and the number of UEs served by the target cell at that time plus a Hysteresis margin which is used for further protection. The second part is an algorithm which purpose is restrict inbound HO of UEs with high speed (they have set a speed threshold of 30 km/h) to allow the access of UEs with 1) medium speed and high demands or 2) low speed regardless of their demands.

The authors in [62] propose a HO algorithm based on SINR measurements at the UE side, type of service demanded (real-time/non real-time) and speed. The target of the authors is to increase the handoff probability to the small cells, something which is achieved because the SINR criterion is better than RSRP comparison. This results to an increase in the number of HOs due to higher utilization of small cells. However the speed and type of service criteria they use in the proposed algorithm, protect the system from performing unnecessary HOs which are defined as those that doesn't satisfy users' requirements.

A cost function which is proposed as a solution for minimizing HO failure rate (HFR) can be found in [63]. The solution introduced uses performance factors that are related with the HO process and the HO failure as well. These factors which are the load difference between the target and serving cell, the speed of the UE and the service type, are integrated into a cost function with different weight each. The outcome of this cost-function is used along with the conventional RSRP algorithm as an secondary Hysteresis margin to lead the system to the expected results.

A way to provide better QoS for the UEs and save energy from the small cells is proposed in [64]. The idea presented is a new Autonomic architecture in the framework of self-organizing networks, under which each group of small cells will elect a Femtocell cluster head which will be responsible to manage the operations in the cluster it belongs. The Femtocell cluster head is elected in a way that the distance between the small cells of the cluster and the cluster head is minimized in order to reduce the cost of transmission energy when the entities...
communicate with each other. This Femtocell cluster head is also responsible for managing HO operations of UEs in its cluster restricting the signaling needed for HO procedure inside each cluster.

The authors in [67] motivated by the fact that HO algorithms which tune cell (pair) level HO parameters (TTT, Hysteresis Margin) have restricted gain, propose a HO trigger algorithm called network controlled HO. Their idea which is an extension of the conventional HO trigger mechanism, manages to improve the tradeoff relationship of HO reliability and HO frequency. The way it achieves that, is by not triggering the HO mechanism when the RSRP level of a neighbor cell is higher than the RSRP level of the serving cell but by waiting also for the SINR of the source cell to pass below a threshold value. The latter is done because the better RSRP cannot result with high certainty to better downlink SINR for the UE, since it doesn't account for the level of interference received and which is able to degrade the level of the received service.

In [65] the authors analyze the mobility performance in LTE Heterogeneous Networks with macro and pico cells. The aim of these proposals is to accomplish better performance in terms of radio link failures, number of HOs and offloading to pico cells. They achieve this, through adjustment of the TTT parameter according to the speed of a UE and the coverage area of the target cell. The motivation behind this idea is that in HetNets where cells have different coverage areas and where there are UEs with different mobility performance (3km/h, 30km/h, 60 km/h and 120 km/h), it's not efficient to have a fixed HO policy because this drives the system to undesired performance. Specifically they found out that by using in HetNets criteria, suitable for macro-only scenarios, the link failures that occur in the case of pico to macro HO are excessively high. The outcome of their proposal is that there is not an optimal value for the TTT since it is totally dependent on the cell size and the speed of the UEs. What they observe is that the optimum TTT increases with cell size and decreases with user velocity and this explains why their proposal consists of different values of TTT for different combinations of user speed and type of cell (macro - pico).

In [66] which is an extension of the idea proposed in [65], the authors targets are the ones mentioned before. The catalyst of their proposal is the lack in the literature of algorithms that combine a strategy to keep fast users in the macro cell and simultaneously to maximize the offloading of low speed users to pico cells. Their proposal consists of three different algorithms. The first is an extended Mobility estimator algorithm with weights for different type of HO (macro-to-macro, macro-to-small, small-to-macro, small-to-small). The algorithm counts small cell related HOs with lower weight due to the smaller coverage area of the pico cells. The second algorithm is a Gray-list solution which prohibits the inbound HO of fast users to a list of small cells unless it's necessary for interference reasons. Finally, their last proposal is a cell dependent TTT scaling which applies different TTT value for users with different mobility state and different type of HO.

The results present in [66] are in line with the theoretical analysis in [69]. The authors in [69] manage to characterize the relation between two basic HO metrics, namely HO failure and ping-pong rate, as a function of HO parameters like TTT, speed of UEs and range expansion bias in a macro cell - pico cell HetNet scenario. The motivation for this analysis was the realization that using in HetNets the same set of HO parameters used in macro only scenarios, increases the HO failure rate and the ping-pong rate. Their results which were also verified by simulations, prove that there is an optimum TTT value each time for a given UE velocity and a given cell that minimize HOs failures and ping-pongs.

In [68] the authors present a novel scheme to reduce the signaling overhead by UEs who perform frequent HOs in a HetNet scenario with deployed macro and small cells. The proposal uses Inter-Site carrier Aggregation technique and UE autonomous HO decision for small cell mobility selection and camping. In more detail, they assume that according to 3GPP carrier aggregation technique, a UE can be connected to a macro cell (PCell) and if is in the coverage area of a small cell to configure that cell as SCell in order to achieve a higher data rate due to the multiple choices that the UE has to be served. In their proposal they assume network-controlled UE assisted HO between macro cells, but for the small cell management actions they propose to be taken by terminals which later inform the overlaid macro cell for the SCell HO occurred.

3.4 Work plan and future results

As presented in 3.3, most of the references deal with scenarios which cannot be characterized as dense since they consider deployment of a single macro cell - single small cell. In addition they usually study inbound HO from macro cells to small cells (usually indoor or sometimes outdoor) trying to find a way to minimize the number of the HO events, to reduce unnecessary HOs that occur, or to achieve better performance of the network in terms of failure rate or ping-pong rate. Some of the proposals mentioned before focus on reducing the signalling during the
HO operation by making the decision more UE autonomous and less network-controlled while some others relate the reduction in the number of HOs with a reduction in the signalling burden on the EPC side. Moreover, there are solutions that consider Interference among the tiers and solutions that could be characterized as energy-efficient. However, what is noticed is the lack of proposals - with the exception of [65],[66],[68] - that consider HetNet scenarios with all that this implies according to what was previously described, meaning that a few proposals account for HO algorithms in dense scenarios with multiple number of small cells deployed in the coverage area of one or more macro cells. Additionally, few of these proposals account for the heterogeneity in the behaviour of UEs, for example many of the proposals use fix speed values in order to separate fast UEs from slow which is not totally proper.

A simple example where the aforementioned assumption could driver to incorrect judgements is the case of a fast user with speed of 120 km/h that moves on areas where there is not small cell coverage, and a medium speed user 30 - 60 km/h that undergoes many HOs because its trajectory passes through many small cells. If we want more reliable results, the separation to high, medium and normal mobility users should be made not only according to the speed that the UEs have but also taking into consideration their behaviour (direction, number of cells encountered during movement).

Users with different characteristics in terms of speed [52] and demands, cells with different coverage area and different transmission power and finally traffic in the system are parameters that should be taken into consideration in the design of future HO algorithms for HetNets. Furthermore, the fact that most of these parameters undergo high variations during the operation of a network, increases the need for solutions adaptive to these changes. Solutions should be in line with the demands of the system at any time and with the concept of Self-Organizing Networks (SON) that was recently introduced in FP-7 projects [70],[71].

The aim of SON briefly is to minimize the human involvement in optimization of the network, by implementing solutions which give the network the ability to self-adapt to every change that might occur. According to SON concept, cells mainly but also UEs should continuously self-optimize radio parameters in response to changes in network, traffic and even to environmental conditions. The objective of SON is to remove the human factor to higher levels (Network planning, Network policies) and allow the system run more efficient by integrating in it algorithms and alarms that according to the conditions, trigger different operations.

Our work is focused on the study of the HO procedures in dense Heterogeneous networks. We want to propose mobility management algorithms that will tackle the negative effects of excessive HeNB deployment. These negative effects could be in terms of signalling burden for the core network or the air interface due to frequent HO of UEs, especially in medium to high speed scenarios. Our solutions will take advantage of the beneficial capabilities of HeNBs, such as better serving rates and expansion of the bandwidth through re-utilization of their resources, by aiming at better utilization of HeNBs in every proposed algorithm.

Our intention is to cover the gap in literature with sophisticated and adaptive to the network conditions solutions in order to be consistent with the SON emerging ideas for self adaptability of the system as previously mentioned. The base of our proposals will not only be objective criteria like the ones described by also subjective criteria. We support that the Quality of Experience that UEs undergo should be considered as part of every solution/proposal.

Finalizing, every proposed solution should be in alignment with the latest standards and with the CROSSFIRE reference architecture as can been seen in Figure 3.26, that shows the architecture that we consider in the CROSSFIRE project. In this we assume that there are overlapped macro and small cells (femtocells, picocells). UEs can be connected with different type of nodes that exist and can handoff to the cell which offers the best service. Every solution proposed should take into account this architecture and the architecture defined by 3GPP (Figure 3.19).
Figure 0.1: CROSSFIRE architecture
4. Architectural aspects within SONs for Traffic Offloading with small-cells

The increasing number of mobile users, along with the boost of multimedia services, poses new challenges to increase the capacity of the network, particularly the Radio Access Network (RAN). In this context, one of the most impacting trends concerning the network architecture is the densification of the RAN through intensive deployment of small-cells. Such scenarios, composed of multiple access nodes with different coverage areas (the so-called heterogeneous cellular networks), face the necessity of offloading traffic from the macrocell tier to the small-cell tier to meet a major objective: the increase of the network capacity.

The deployment of the small-cells implies nonetheless important drawbacks. Probably, one of the major obstacles is the need for massive investment. This, far from deterring mobile operators from the deployment of small-cells, is envisaged to foster the emergence of new stakeholders in the cellular network market: the small-cell service providers or small-cell infrastructure operators. Therefore, traffic offloading with small-cells will not solely depend on technological aspects, but also on financial considerations.

The focus of this section is twofold: on the one hand, presenting a general overview of the existing 3GPP proposals for traffic offloading in LTE-A; on the other hand, describing the state of the art of the traffic offloading with small-cells, paying special attention to the financial aspects and the transactions between different participants.

4.1 Traffic offloading in LTE-A

Traffic offloading standardization efforts have resulted in two basic approaches: the Local IP Access (LIPA) and the Selected IP Traffic Offload (SIPTO). Both approaches, as well as the set of scenarios and use cases specified in TR22.852 [74] for network sharing, are discussed in the following subsections.

4.1.1 Local IP Access (LIPA) and Selected IP Traffic Offload (SIPTO)

LIPA and SIPTO approaches are defined in [72]. In general, the objective of LIPA is to provide IP capable UEs connected to a H(e)NB with access to other IP capable entities of the same local (residential/enterprise) IP network. In that sense, LIPA is considered an offloading mechanism since traffic is served by the H(e)NB instead of being served by the macro cellular network. It is worth noting that, whereas data traffic is offloaded to the local network through the H(e)NB, signaling traffic is not offloaded, thereby traversing the mobile operator’s network.

SIPTO is a technique for reducing the load on the system by redirecting optimally selected types of IP traffic to a specific IP network node, which is in the proximity of the UE’s point of attachment to the access network. SIPTO can be applied at either a local network with a HeNB or at the macro cell access network. The selection of the types of IP traffic to be offloaded is decided by the mobile operator in order to ensure the QoS for particular services.

A breakout point is defined as the point in the network where traffic offloading takes place. Based on the location of the breakout point, network architectures can be classified into two different categories: first, architectures in which the breakout point is placed at the local network, and secondly, architectures in which the breakout point is located at the RAN or above. With regard to the aforementioned categories, whereas the former includes the LIPA and SIPTO approaches with small cells, the latter includes the SIPTO approach with both macro-cells and small cells.

The rest of this section provides a detailed description of LIPA and SIPTO. For better comprehension of the two proposals, the same structure was applied, based on their requirements, architecture and procedures of each approach.

Local IP Access (LIPA) requirements:

As stated in the 3GPP specifications ([72][73]), there is a set of general requirements defined for applying LIPA. These requirements are listed below:

- The LIPA function should provide access between the UE and other IP capable devices connected to a local network, without the data flow traversing the cellular network. The UE shall be able to connect to the public internet, if the local network is connected to it.
The UE should be provided with simultaneous access to the mobile operator’s network and to the local network, utilizing LIPA.

According to regulatory requirements, the traffic managed by the LIPA function must be maintained within the residential/enterprise IP network, between the UE, HeNB and the other entities.

A UE must have a valid subscription with the mobile operator for using LIPA.

A UE will be able to utilize LIPA in a visited network where, previously, a roaming agreement has been established between the operators.

LIPA shall not affect services running in parallel for the same UE.

LIPA will be available to pre-release 10 UEs.

LIPA deployment should not bring forth vulnerabilities of the mobile operator’s network security.

Architecture

The LIPA requirements described above are translated into the network architecture for HeNBs, provided Figure 4.1 [73].

In this architecture, a new entity, called Local Gateway (L-GW), is incorporated into the RAN. This entity can be either collocated on the HeNB or as a standalone node, includes partial P-GW function and S-GW downlink data buffering function.

In the core network (CN) side, the S-GW is responsible for serving the CN traffic, while the session for LIPA traffic is handled by the L-S11 interface, which connects the L-GW and the MME. The L-S11 interface is also responsible for establishing the LIPA PDN connection, requiring that the L-GW is selected close to the HeNB. Moreover, two functions have been defined for the L-GW for idle and connected modes. Downlink (DL) packet buffering for the former, and data routing between the L-GW and the HeNB, for the latter.

Apart from the previous requirements, the solutions for LIPA should also fulfil, the following architectural requirements:

- The UE must be able to request a LIPA PDN using
  - A well-defined Access Point name (APN), or
  - A specific indication independent of the APN
- Intranet type access to home based network should be provided by the LIPA supporting HeNBs.

Multiple PDN support

An important fact that affects the LIPA architecture is that multiple PDN support is not available in all the UEs. Thus, in accordance with [73], the LIPA solutions defined should examine the subsequent situations:

- Single PDN support, where only one common PDN connection can be used for both LIPA and non LIPA traffic.
- Multiple PDN support, where multiple PDN connections can be used simultaneously, enabling different PDN connections for LIPA and non LIPA data flows.

Procedures

There are four major procedures in LIPA [73]:

---

Security: Public
1. UE Requested PDN Connectivity to LIPA (Error! Reference source not found). Figure 4.2. This procedure is followed for setup, release and data session between the UE and the L-GW. Once the data path among them is created, the UE can be connected, through the L-GW, to the other connected devices of the local network.

2. S1 Release procedure when LIPA PDN connection for UE exists (Figure 4.3). This procedure is used when the UE is in idle mode, and the S1 UE context management function has to release the context previously established in the HeNB to enable active-to-idle transition.

3. UE Triggered Service request procedure if LIPA PDN connection existed (Figure 4.4). This procedure is initiated by the UE in order to enter connected mode in the local network, for the case that a LIPA PDN connection has been already established.

4. L-GW Triggered Service request procedure if LIPA PDN connection existed (Figure 4.5). This procedure is initiated in order for the MME to signal with a UE that is in idle mode, and has to enter the connected mode.

There are some important aspects that should be noted for each procedure. For the UE Requested PDN Connectivity to LIPA there is a necessity for a well defined APN or a special LIPA indication to display the desire of LIPA. In the S1 release procedure it should be noted that in step 4, the HeNB informs the L-GW that the UE is in Idle mode, so that the L-GW enables the S5 path for paging. Regarding the UE Triggered Service request procedure the MME has to include the S5 PGW TEID for each E-UTRAN Radio Access Bearer (E-RAB) in the “E-RABs to be Setup List” in the S1-AP message. As for the L-GW triggered service request procedure a further explanation should be given about the following steps:

1. Downlink packets arriving at the L-GW are buffered in the L-GW.
2. L-GW sends a packet or a "dummy" packet to the S-GW in order to trigger paging.
7. Once the UE triggered Service Request procedure with LIPA PDN connection is complete, the S-GW forwards the packet on S1-U. If there is a "dummy" packet, it is intercepted by the HeNB and discarded (step 7a). Simultaneously, the downlink data that were buffered in the L-GW may start flowing on the direct path (step 7b).
1. PDN Connectivity Request (Well defines APN or LIPA indication)
2. Create Session Request
3. Create Session Request
4. First Downlink data
5. Create Session Response (S5 PGW TEID)
6. Bearer Setup Request (S5 PGW TEID) / PDN Connectivity
7. RRC Connection Reconfiguration
8. RRC Connection Reconfiguration Complete
9. Bearer Setup Response
10. Direct Transfer
11. PDN Connectivity Complete
12. Modify Bearer Request
13. Modify Bearer Request

Figure 4.2: UE Requested PDN Connectivity to LIPA

1. S1-AP: S1 UE Context Release Request
2. Release Access Bearers Request
3. Release Access Bearers Response
4. S1-AP: S1 UE Context Release Command
5. RRC Connection Release
6. S1-AP: S1 UE Context Release Complete

Figure 4.3: S1 Release procedure when LIPA PDN connection for UE exists
Selected IP Traffic Offload (SIPTO)

Requirements
As mentioned earlier, SIPTO can be applied at both macro and small cell layers. Thus, each case introduces its own requirements; however, some of the requirements are common to both layers.

Common (for both macro cell and small cell) requirements [75]
The requirements presented subsequently are necessary for applying SIPTO for both cases, SIPTO in the Mobile Operator Network, and SIPTO at the Local Network.

- The mobile operator must be able to switch on/off SIPTO for a UE and a defined IP network.
- If a UE is associated with a number of IP networks, its IP traffic from one of them shall be offloaded, while from the other network(s) it shall not.
- SIPTO must be available also for pre-Release 10 UEs.
- Services running simultaneously for a UE should not be affected by SIPTO.
- SIPTO deployment should not bring forth vulnerabilities of the mobile operator’s network security.
- Mobility within the macro or the small cell network or between them, should be supported with Service Continuity of the IP flow for SIPTO.
- SIPTO must be available even in case of the UE roaming, while the Home Public Land Mobile Network (HPLMN) provides the Visited Public Land Mobile Network (VPLMN) with necessary information.

Requirements for SIPTO in the macro cell [75]

The implementation of SIPTO in the mobile operator’s network poses additional requirements. These requirements are presented in this clause:

- The availability and function of SIPTO must be defined by the mobile operator for specific parts of the network.
- Integrity and confidentiality of the offloaded traffic shall be secured during SIPTO.
- SIPTO must provide the mobile operator with statistics of offloaded traffic per user.
- SIPTO must be applied by the network without the interaction of the user.

Requirements for SIPTO in the small cell [72]

A HeNB SubSystem of a Local residential/enterprise is enabled to support SIPTO in order to provide access for a UE connected via that HeNB to a defined network (e.g. the Internet). The necessary requirements for this function are listed below:

- Traffic offload must be done without traversing the mobile operator network.
- The function of SIPTO must be decided per HeNB by the mobile operator, and the HeNB Hosting Party.
- Traffic offload decisions must be done in accordance with the mobile operator SIPTO policies, and the user’s consent per APN.
- The SIPTO policies may be defined per APN or per IP flow.
  - SIPTO policies per APN determine if all the traffic related to a certain APN, can be offloaded.
  - SIPTO policies per IP flow determine which APN will be used for routing a particular IP flow.
- A UE must be able to connect to the mobile operator’s core network, while in parallel, it will connect to a defined IP network (e.g. the internet) via a fixed local IP network using SIPTO.
- A UE must be able to access simultaneously both PLMN and fixed services via a fixed local IP network using SIPTO.
- The mobile operator shall be able to configure the SIPTO policies either statically or dynamically.

Architecture

According to [76], SIPTO above RAN corresponds to a traffic offload through a P-GW located in the mobile operator’s core network. A mobile operator can provide SIPTO above RAN to a UE, if it selects a set of GWs (S-GW and P-GW) that is in the proximity of the UE’s point of attachment to the network. The appropriate selection of the above set is done by using the UE’s current location information. Then, the operator’s decision for allowing or prohibiting SIPTO on per user and per APN basis, depends on the subscription data in the HSS. Using this data, the MME decides on whether to allow or prohibit the offload. However, if this information contradicts with the MME’s configuration for that UE, then SIPTO is not applied.

Figure 4.6 illustrates the architecture defined for SIPTO for the EPC macro network in accordance with the 3GPP specifications in [73]. It should be noted that in this diagram, the L-PGW indicates the P-GW closer to RAN, which serves as the breakout point towards the Internet.
Figure 4.6: SIPTO for EPC macro network with S4 SGSN [73]

In case the operator wants to apply SIPTO for internet traffic, but uses the same single APN for both internet and operator services, the solution is to use an L-GW for both types of traffic. While the internet traffic is offloaded locally, the traffic for operator services is routed within the operator’s network.

Procedures

The SIPTO function includes a set of important procedures, which are described in [76]. The basic call flows are illustrated in this following, along with their respective information. Although the figures presented in the sequel describe the main SIPTO procedures when it is applied to eNBs, it is worth noting that these procedures can also take place in HeNBs. In that case, procedures are alike but eNBs should be replaced by HeNBs.

Following are presented three major SIPTO procedures:

1. E-UTRAN Tracking Area Update with S-GW change (Figure 4.7). This procedure is followed when an E-UTRAN-attached UE experiences any of a particular set of conditions, such as the expiration of a periodic tracking area update timer. The rest of the conditions can be found in [76]. In this case, the procedure is adapted to the situation where SIPTO is allowed for the APN related with a PDN connection. Then, the new MME will re-evaluate whether the current P-GW location will continue to be used (step 8). And, in case the MME decides a P-GW relocation, it will start a PDN deactivation with “reactivation requested” at the end of the TAU procedure.

2. UE or MME requested PDN disconnection (Figure 4.8). This procedure enables the UE to request for disconnection from one PDN. It can be also used in a similar way for the MME. For the SIPTO function, this procedure is used in order for the MME to decide if a GW relocation is needed. As mentioned above, in this case, the MME releases the PDN connection(s) associated to SIPTO-allowed APN(s) using the “reactivation requested” cause value, and the UE should then re-establish those PDN connection(s).

3. MME triggered S-GW relocation (Figure 4.9). This procedure allows the MME to trigger S-GW relocation as a result of events, different than those in mobility scenarios. It is used for establishing SIPTO at local network PDN connection with stand-alone GW or above RAN PDN connection.
Figure 4.7: E-UTRAN Tracking Area Update with S-GW change

Figure 4.8: UE or MME requested PDN disconnection
1. MME determines the Serving GW relocation needs to be performed.

2. Create Session Request

3a. Modify Bearer Request

3b. Modify Bearer Response

4. Create Session Response

5. S-GW Relocation Notification/ACK

6a. Delete Session Request

6b. Delete Session Response

Figure 4.9: MME triggered S-GW relocation
4.1.2 Network Sharing Scenarios and Use Cases

Network sharing has been proposed for the past years as the solution to enable cooperation between mobile network operators, according to the respective regulations issued by organizations such as ITU. Since in the not so distant future, network sharing will become a common act between cellular network stakeholders, 3GPP designed its own framework, providing the basic guidelines for such future agreements. This section, presents the agreements’ actor roles, as well as some of the scenarios and use cases specified by 3GPP.

4.1.2.1 Network sharing use-case actor roles

A network sharing agreement can be developed differently between particular scenarios and use cases, due to a variety of financial, technological, network deployment and regulatory conditions. In order to categorize and describe these scenarios, a set of common roles has been defined in the 3GPP specifications [74]. In the following paragraphs these common roles are presented for better comprehension of the described use cases.

- **Hosting RAN Provider**: The Hosting RAN Provider owns the RAN infrastructure. Additionally, it shares the RAN with one or more Participating Operators after establishing a network sharing agreement. It should have primary operational access to specific licensed spectrum (not necessarily owned by it), which is part of the sharing agreement. Since it owns and operates the RAN it should provide the necessary facilities to the Participating Operators for operating the RAN.

- **Participating Operator**: The Participating Operator uses the Hosting RAN Provider’s shared RAN facilities, and depending on the agreement, possibly with other Participating Operators. According to the agreement, it uses a part of the shared licensed spectrum, as well as a part of the shared RAN, in order to serve its own subscribers in the specified geographical area.

Within the concept of the Hosting RAN Provider and the Participating Operator, other entities can be involved, such as outsourcing, joint ventures or leasing agreements for operating or owning the RAN infrastructure, or managing the sharing agreements.

- **Roaming Operator (HPLMN and VPLMN)**: Two networks participate in during roaming: the Home Public Land Mobile Network (HPLMN) and the Visited Public Land Mobile Network (VPLMN). The former refers to the operator a UE belongs, while the latter refers to the operator a UE subscribes, when it is out of its operator’s coverage, using roaming.

Hence, when roaming is used, outside of the HPLMN’s geographic coverage area, the subscriber will utilize the VPLMN’s coverage area. But, in the situation of a RAN sharing agreement, its participants don’t distinguish their coverage area but provide the same one through the Hosting RAN.

- **Operators with multiple roles**: Every network sharing agreement defines the spectrum, the geographical region, and the RAN components that are object of the agreement. These components, as stated by 3GPP, define a network set. These network sets are regarded as independent, and can be combined with each other in a variety of ways, providing with multiple roles the participants of the agreements.

- **What constitutes a shared RAN**: A Hosting RAN Provider may share E-UTRAN resources with Participating Operators in various ways. In accordance with 3GPP, a sharing agreement should consider, at least, the use of a set of Radio Base Stations by the Participating Operators. However, it is not obligatory for the Hosting RAN Provider to share a part of its own spectrum.

4.1.2.2 Scenarios and use cases

In general, RAN sharing can be divided into two categories. The first one, known as “Greenfield deployment” is referred to the situation where two operators come to an agreement for cooperating in the deployment of a new technology. The second one, defined as “Buy in” is regarded for the case when an operator has already deployed its own network, and afterwards, it is shared with another operator. Scenarios from both categories are described in the following clauses, along with the requirements each case necessitates.

Before presenting the use cases, it should be noted that as stated in the 3GPP specifications, RAN sharing is defined for situations where the shared RAN is operated as a home network (HPLMN) by all the Participating Operators.

**Maximizing RAN sharing revenue**

The use of the shared RAN resources by the Participating Operators is considered to be done in a dynamic manner. Based on this variable consumption of resources, the Hosting RAN Provider can use a dynamic charging and
resource allocation model according to the usage. In such conditions, the Hosting RAN Provider has to confront a trade-off between maximizing its revenues and the efficient use of its RAN.

The Hosting RAN Provider provides network wholesale services to Participating Operators, who use the shared RAN in order to offload traffic from their own network, which is not part of the agreement. The pricing model of the Hosting RAN Provider depends on the time period (e.g. month, week, on demand) of the other operators’ demand. As a requirement, the shared RAN must be able to distribute the capacity according to the choices given by the Hosting party to the participating operators, regarding the resource allocation model.

**Asymmetric RAN Resource Allocation**

The Participating Operators create a Joint Venture (JV), which deploys, operates and maintains the Hosting RAN. However, the Operators have different financial, and hence, controlling percentage of the JV. It should be noted that each partner is responsible for its own EPC infrastructure, separately from each other.

Since in situations of low traffic load the QoS objectives for both parties will be fulfilled, the situations of high traffic load attract more interest. Thus, it is required that each partner established its pre-agreed usage portion of the Hosting RAN. Also, there should be an element in the shared RAN which measures constantly each Operator’s resource usage in order to ensure their arranged rights in the JV. However, in near capacity situations for one of the Operators, there could be a margin of tolerance in the distribution of the resources. Moreover, the agreement may contain an option for changing temporarily the resource percentage of the parties, for such circumstances, in order for them to meet their QoS objectives.

**Dynamic RAN Sharing Enhancements**

This clause refers to the situation where the Participating Operator has variable capacity demands from the shared RAN, during specific time periods of the day or the week. The Participating Operator demands various portion allocations of the shared RAN in order to serve its load in unpredicted traffic variations.

For successful results, the Hosting Operator needs to optimize the shared RAN in order to meet both its operating and financial objectives. As for the Participating Operator’s case, it needs to choose its demands for increased RAN capacity based on its subscribers’ traffic patterns, for both its QoS and economic targets.

It is required by the shared RAN to provide the required flexibility for serving the dynamically changing demands of the Participating Operator as well as the Hosting Operator’s load. Another important requirement is that the Participating Operator should be able to keep the pre-agreed schedule by driving its connected and idle UEs towards or away from the Hosting RAN, at the beginning and at the end of the scheduled leased period, respectively.

**On-demand Automated Capacity Brokering**

As the title implies, there is a Hosting RAN provider that shares a portion of its RAN capacity predefined for sharing with other Operators using automatic means. The use of automatic means is necessary to the shared RAN for handling On-demand Capacity requests by Participating Operators.

In order to use such an agreement efficiently, the Participating Operator, after estimating its needs for additional resources, should determine the amount of necessary shared resources, the time period for which it will have access to them, and any service specific attributes needed. One issue is that, based on the agreement, the Participating Operator may or may not have exclusive access to the allocated resources.

Operating in an environment with dynamic changes in the traffic load, there is the possibility that the Participating Operator will require additional resources even after an agreement has been made. In accordance with that, the Hosting Operator should be able to support such requests, as well as requests for cancellation or withdrawal of granted on-demand requests.

**Load balancing in shared RAN**

Another issue specified in the specifications is that of multiple operators sharing a specific coverage area, which consists of several cells. The agreement in this case specifies the operator’s shares. Load balancing among the shared cells should be applied based on each participant’s predefined RAN partition. The system, also, should be able to impose the agreed load levels when one of the Participant Operators exceeds its maximum load levels in a particular cell, by preferably transferring UEs to adjacent cells, if possible.
RAN Sharing Charging Event Triggering

The situation of more RAN operators that have entirely or partially overlapping coverage areas is regarded; one of them may provide contiguous coverage, while the other covers areas where additional capacity is required. A Hosting RAN Provider decides to share its RAN capacity with Participating Operators in geographical coverage areas where the latters are in need of additional capacity. It should be noted that the Participating Operators have already their own non-shared RAN.

This kind of scenario requires a mechanism for producing wholesale start/stop charging records in order to estimate the usage of the shared RAN. That is, the mechanism needs to take into account every UE’s mobility events towards and from the shared RAN. Also, this accounting mechanism should be able to distinguish the mobility events that do not require charging.

4.2 State of the Art on traffic offloading with small-cells

During the past few years, mobile network operators have given a lot of attention and funding in traffic offloading. Subsequently, the research community has been proposing a lot of interesting ideas regarding its efficient implementations. In this subsection the reader is provided with the SoA in traffic offloading from two different views. One that refers to the utilization of SONs and cognitive radio, while the other targets its economic issues.

4.2.1 SONs and cognition aspects

Following, there is presentation of a number of State of the Art proposals that make use of SON functionalities and cognitive radio for more efficient traffic offloading implementations. A system that enables a cellular operator to utilize TV whitespace (TVWS) spectrum, along with licensed spectrum aiming to increase its network capacity, is found in [77]. In order to do that, the authors propose an architecture for cognitive radio system concepts. Then, this architecture is mapped to the LTE architecture, aiming for traffic offloading in LTE systems with the opportunistic use of TVWS.

Next, an overview of the proposed architecture is given. For the protection of incumbents’ transmissions, there is a necessity for accessing regulatory information through relevant databases. Complementary to them, is the use of spectrum sensing, for enabling resource management in the free spectrum bands. For these functionalities, the proposed architecture includes two separate blocks for regulatory databases (REG), and Spectrum Sensing (SS), respectively. For the responsibilities of spectrum management (SM) and resource management (RM), two entities, addressed as Cognitive Manager (CM), are designed for each one (CM-SM and CM-RM). They are based on decision-making mechanisms and characterised by longer and shorter time-scales, respectively. The CM-SM controls co-existing players by managing and optimising spectrum PortFolios (PF), whereas the CM-RM achieves effective handling of time-varying spectrum resources, including reserve channels. For the cognitive communication and signalling between the above and rest of the system’s entities, an Adaptation Layer (AL) is defined.

Figure 4.10 illustrates a mapping of the proposed architecture to the LTE one, for the offloading scenario, using TVWS opportunistically.

![Figure 4.10: Cognitive functional architecture mapped to the LTE offloading scenario [77]](image-url)
A PF is built by the CM-SM at the Evolved Packet Core (EPC), accessing all the repositories spectrum and geo-location databases (REG), as required by applicable regulations. A need for a PF may be triggered by the CM-RM at the eNB, which is tracking the network load and the corresponding resource requests from UEs, and allocating the resources as usual. PFs are, therefore, pulled from the CM-RM for deployment to the eNB, and pulled from the EPC for revocation as needed. The TVWS can be accessed by UEs for their data exchange by using their flexible transceivers (TRX). In order to support radio context acquisition, the SS at eNB exploits the SS units at the UEs to realise collaborative sensing, thus, exploiting the UEs which act as scattered sensors. In addition for incumbent protection whenever applicable, the SS service is used by the CM-SM for PF repository maintenance, thus augmenting the geo-location data for improved calculations. This repository merges all the relevant regulatory constraints with the augmented information, such as the radio context from the SS and the network load and usage reports provided by the CM-RMs within the scope of that PF.

![Diagram](image)

**Figure 4.11: System overview for offloading of LTE** [77]

Figure 4.11 presents a potential implementation of the previously described system within an LTE architecture. This systems utilizes “improved eNBs” that operate in both licensed and TVWS bands, in order to serve the cognitive UEs. The eNBs are connected to new network entities apart from the LTE ones, with the use of both LTE and proposed interfaces. Although the proposed architecture is mapped at the macro layer of an LTE system, it could be as well applied to HeNBs that are designed to support cognitive radio. An interesting approach for a framework that models and analyses two-tier Heterogeneous Networks with cognitive Femtocell Access Points (FAPs) is proposed in [78]. All FAPs operate in open access mode and there is, also, the assumption of perfect spectrum sensing. The authors make use of spectrum sensing region (SSR) that is defined around each FAP by the spectrum sensing threshold. The latter introduces a trade-off between aggregate interference, and spatial frequency reuse efficiency. The objective is the minimization of the outage probability through the discovery of an optimal spectrum sensing threshold. This is determined by the HetNet parameters such as the number of channels, the relative transmit power, as well as the number of macro BSs and FAPs. An extensive analysis is elaborated, based on these parameters, while stochastic geometry (Figure 4.12) is used for modelling the outage probability of macrocell and femtocell users for downlink transmission in a multi-channel environment.

Their work presents the benefits and importance of cognition in improving a network’s overall performance. Introducing cognition decreases the overall outage probability, and with the proper control of the spectrum sensing threshold will offload users from the macro to the femto network tier, while keeping the SINR performance.
An autonomous cooperative solution for the joint access and backhaul design, named JAB-approach is presented in [79]. The scenario considers two operators. One operating a macrocell network, and the other that operates an overlaid picocell network. An infrastructure co-sharing problem is described, which brings financial incentives to the picocell network operator, and increased network performance to the macrocell one.

The aim is to enable the cooperation of the two operators for utilizing the picoBSs for an in-band and over-the-air backhaul. In order to achieve that, the focus is on decentralized self-organizing algorithms. The macrocell’s subscribers can either make use of the picoBSs as a helping relay to establish or not communication with the BS. These connection methods are labelled as cooperative and non-cooperative approach, and are depicted in Figure 4.13.

The study of the JAB design and the infrastructure co-sharing problem is done using a game theoretic learning perspective, and a decentralized reinforcement learning (RL) algorithm is proposed. The experiments show that this proposed scheme can bring substantial benefits to both operators.
Next, a proposal of a cognitive hybrid division duplex (CHDD) scheme for HetNet deployment is presented [80]. This scheme has two frequency bands that are utilized with frequency division duplex (FDD) in the macrocell, whereas in parallel, a number of underlaid cognitive femtocells operate in the same bands with time division duplex (TDD). The deployed frequency bands of the CHDD scheme are presented in Figure 4.14.

It is assumed that the macrocells operate on FDD mode, with frequency reuse one, while the underlaid femtocells are enabled to operate on TDD mode utilizing both spectrum bands. This results in having no need for supplementary spectrum for the femto tier, in which the cognitive femtocells can use the FDD spectrum in an opportunistic and efficient way. It is shown that the proposed CHDD scheme combined with an open access policy for the femto tier can provide an effective interference management, as well as an efficient approach for deploying cognitive and opportunistic small cell networks.

A scheme that enables the analysis of the trade-off between the energy consumption and the traffic offload of cognitive small cell access points (SAPs) placed in a macrocell [81] is described. In particular, the authors analysed the trade-off between the detection efficiency of UE and the total capacity that can be offloaded from a macro base station (MBS). The proposed model considers channel fading, aggregate network interference, bursty activity, network topology and load. It enables computing the effect of critical system parameters such as the interferer density and the SAP coverage on the detection performance, the aggregate offload capacity and throughput, as well as the total energy consumption.

Their experimental results show that having information of the interference environment can provide a considerable reduction of the SAP energy consumption. The results provide evidence that their model can be used both for the design of the optimal sensing time and sensing probability, as well as for the evaluation of the energy efficiency with respect to network topology and load. Hence, it is claimed that their framework can be used for the energy efficient design and operation of cognitive SAPs in heterogeneous networks.

4.2.2 Financial aspects

In this section there is an extended description of the latest work in traffic offloading with small cells from a financial point of view. It should be noted that some specific works presented here, do not seem to comply with the section’s topic. Nevertheless, their proposals could be adapted to a scenario similar to those of our interest e.g. by changing roles for basic entities.
A case of traffic offloading through third party WiFi or femtocell access points (APs) is considered in [82]. The authors assume a network initiated procedure with the mobile Users (MUs) being unaware of the offloading process. This work is mainly focused on the transactions and the behavior between the network operators and the APs, and in particular, on the incentives the operators need to provide to the APs in order to initiate a cooperative offloading.

Figure 4.15 depicts a particular case of data offloading between macrocellular base stations (BSs) and APs. It is assumed that the offloading can occur between one macrocell BS and multiple APs, and vice versa. A market based solution is introduced where economic incentives are given for offloading traffic. In order to find a balance between the amount of traffic that should be offloaded and its financial compensation, the authors designed a two-stage multi-leader multi-follower game, called data offloading game (DOFF), where BSs act as leaders and APs as followers.

In the first stage of the game, all the BS offer prices to the APs in their coverage areas, and in the next one, the APs decide on the amount of traffic volume they will offload based on the proposed prices. The proposed game manages to reach an equilibrium that exists between two extremes, which are classic market outcomes, the Market Balance and the Monopoly Outcome.

Continuing from their previous work [82], the authors in [83] make use of a market where mobile network operators (MNOs) lease third-party deployed WiFi or femtocell access points (APs) to dynamically offload the traffic of their mobile users. Figure 4.16 represents this general case. Again, as in [82], it is assumed that the offloading can occur between one MNO BS and multiple APs, and vice versa. In this work though, an iterative double auction mechanism is introduced, which is managed by a – independent from the participants– broker.

Considering incomplete information, the broker must incentivize both sides for truthful bidding in order to maximize the market’s efficiency, while it maintains profit. The mechanism consists of two stages during each iteration. In the first stage, the MNOs bid for every AP for their offloading needs, and the APs bid for every MNO for their serving costs. In the second stage, the broker decides on the distribution of the APs’ resources based on the previous bids. The proposed iterative double auction mechanism, satisfies the desirable economic properties, and maximizes the welfare of the market.
The study of a cognitive radio network that consists of a primary spectrum owner (PO), multiple primary users (PU) and multiple secondary users (SU), is conducted in [84]. An example of the studied network is illustrated in Figure 4.17. The proposal of this work is a reserve price auction mechanism for spectrum sharing in cognitive radio networks. This auction scheme enables the SUs to buy spectrum bands from the PO. It is differentiated from similar works by assuming a realistic case of channels with different qualities. Also, based on the latter, the SUs can have different valuations for every channel. Moreover, the PO sets reservation prices according to the channel quality.

The auction mechanism manages to allocate efficiently the available spectrum, and the numerical results indicate performance improvements compared to the case of reserve price auction with identical channels and the case of having no reservation prices.

Although this scenario doesn’t comply with this section’s main subject, its mechanism could be useful in studying relevant subjects, while compared with other mechanisms. If we changed the roles in this scenario, we could design a scenario where macro service providers (in this case the SUs) compete in this auction for resources of a third party small cell service provider (in this case the PO’s available channels).

Following, a Vickrey-Clarke-Groves (VCG) auction based incentive framework for accessing selfish femtocells [85] is presented. There are considered two scenarios according to the number of the macrocell users. The first one is for a single macrocell user that can aggregate the throughput of neighbouring femtocells. For this case a multi-unit reverse auction has been formulated, for femtocell competition. The second one includes multiple macro users, as represented in Figure 4.18, and a double auction scheme is designed for the optimal allocation of the femtocell resources.
For both cases, the VCG auction has been used as a foundation for designing the proposed auction schemes, orchestrated by the macrocell operator. The results show that the framework is individually rational, can conduct truthful auctions, hence, preventing market manipulations, and it improves the network’s performance.

Figure 4.19: Sample cellular sector and its WiFi regions [86]

Continuing with auction schemes, iDEAL [86] is presented, an auction-based incentive framework that enables a cellular service provider to lease on demand resources from third parties. The provider conducts reverse auctions with the third parties, from which it can buy additional, necessary capacity. As shown in Figure 4.19 only a single sector is considered, and is divided into WiFi clusters. In each sector the cellular resources are distributed into different regions. Then, a decision is made for each region on the amount of WiFi resources needed. Afterwards, a local auction is used for the cellular and WiFi resources of the region. This procedure then is used in an urban area of numerous macrocells.

iDEAL consists of two stages. In the first stage, the allocation of cellular traffic is done among the WiFi regions, while the operator minimizes its costs regarding its bids. In the second stage, the payment of the resource owners is decided, while concerning for providing them with enough incentives. The proposal manages to avoid collusion and conduct truthful auctions. It also manages to gain more revenue for the operator by not providing service during the peak traffic hours, and the third-parties earn more revenue by selling their unused capacity, while the subscribers’ QoS is ensured.

Figure 4.20: Hierarchical dynamic game framework for pricing, spectrum sharing, and service selection [87]

A scenario where a macrocell service provider (MSP) needs to incentivize small cell operators (SSPs) to accept its excessive traffic with a monetary compensation, is studied in [87]. The fact that differentiates this work from others, is that apart from the service providers, their subscribers also have a major role. They can dynamically select their service provider based on the provided QoS and the cost, which depend on the pricing and the open access ratio (the SSPs’ resources for the macro users).
As illustrated in Figure 4.20, a hierarchical dynamic game framework is proposed for modelling the interactive decision model. In the lower level, an evolutionary game is designed in order to describe the users’ dynamic service selection. In the upper level, the MSP and SSPs decide successively the pricing strategy and the open access ratio, considering the constantly changing status of the users. The upper level defines the MSPs as leaders and the SSPs as followers in the designed Stackelberg differential game. The game manages to reach an equilibrium for a variety of situations, and the numerical results show the effectiveness and advantages of dynamic control of the open access ratio and pricing.

Another case of traffic offloading to privately owned femtocells and the respective economic framework is described in [88]. The femtocells can be accessed by the public users due to a hybrid access mode. Figure 4.21 depicts the network scenario used. The problem of providing the femtocell owners with incentive for traffic offloading is addressed with the concept of profit sharing, which is handled with a game theoretical analysis.

A two-stage sequential game is modelled, named Femtocell Service Selection Game (FSSG). In the first stage, the operator chooses the ratio of the revenue distribution to femtocell owners in order to maximize its own benefit. In the second stage, the femtocell owners decide on the proportion of spectrum for the public users, while ensuring their subscribers’ service requirements. Simultaneously, all the subscribers select from which layer they will be served. It is shown that the concept of sharing motivates the femtocell owners to share their resources, while providing the operator with more served users, and hence, increased profits. Also, the proposed framework shows the feasibility of femtocell service provision.

A case study different from the rest is one that assumes the cooperation of a mobile operator (MO) with a fixed-line operator (FO), in order to provide femtocell service to indoor users [89]. In order to achieve an agreement regarding a QoS guarantee, the MO offers a share of the profits of the indoor femtocell service, provided that the FO can deploy a low cost infrastructure. When they collaborate for this service, they can operate as virtual integrated operators. Moreover, regardless of the benefits shared, both of them can achieve higher revenues.

The scenario is formulated with a sequential game and Nash bargaining. The sequential game describes the relation between the MO and the users. First, the operator sets the price for the femtocell service and then the users, based on the price, select their spectrum requirements. Nash bargaining is then used between the two operators in order to share fairly the profit. The operators will cooperate only in the case where they can make more revenue. Thus, the evaluation is based on the comparison of the net profits in the cooperation and non-cooperation cases. The results have verified the cooperation expectations, along with a substantial increase in the spectrum efficiency.
Apart from the cases where typical mobile network operators initiate network sharing arrangements, there are ones, where the femtocell operator/service provider, takes the initiative. In [90] such a scenario is presented, where a femtocell enables the hybrid access of macro UEs, provided that they rent its power resources. It is assumed though, that the femtocell’s power is limited. In case its power is totally used and it cannot serve its own femto UEs that return to its coverage area, it selects one macro UE as a relay for control data to the femto UEs. When this happens, the particular macro UE will be served free of charge.

The system model is depicted in Figure 4.22. The optimal power and price values for the above procedure are found with the use of a Stackelberg game. The equilibrium of the game provides these values, which also improve the throughputs for both the macro and femto UEs.

In contrast with the rest SoA, the next work introduces a policy based offloading model, which is based on a cost function approach [91]. The policies studied are user centric, network centric and hybrid. For both user and network centric policies, inter-system measurements are performed by both entities, respectively, and in accordance with them, they make an offloading decision.

In the case of hybrid offloading approaches, the decision is made cooperatively between the user and the network. For this case, a mechanism is also presented, which is based on autonomic networking principles in order for the policies to be chosen dynamically, according to the traffic load and the operator strategies. The mechanism is divided in three phases, as illustrated in Figure 4.23. During the Monitoring phase, the system checks constantly QoS parameters and reports to the Central Network Management Entity (CNME). Then, in the Decision Making phase, the CNME analyses this data and concludes to a set of suitable policies. Finally, in the Execution phase, the
policies are sent to the user, which in turn will make the final decision. The evaluation of the mechanism is done under realistic traffic loads, and it shows substantially better performance compared to the other two policies, in terms of offloading efficiency and blocking ratio.

In [92] the authors propose an ACcess Permission (ACP) transaction framework. This framework enables a macrocell service provider (MSP) to buy ACP from multiple femtocell service providers (FSPs), which will operate in hybrid access mode, in different geographical areas for a specific duration (of T timeslots). In case the FSPs have overlapping coverage areas they compete for selling their ACP.

The scenario’s problem is the dynamic change of the MSP’s decisions in time and space, which leads to information incompleteness, and hence, difficulty in the choice of an appropriate strategy for the FSPs. The proposed solution used for this problem, is an adaptive strategy updating algorithm, which is based on online learning process. A theoretical proof is given that the payoff gap between the proposed algorithm and the compared optimal static strategy is bounded. From the results is shown that the profit of the FSPs depends on the learning speed of the proposed algorithm, as well as on their provision with additional information.

The next work describes a model on user service adoption for choosing between a basic wireless technology and a bundle of it [93]. The supplementary technology is used by the service provider in order to offload traffic from the basic technology’s network. The choice of the technology depends on parameters such as the technology’s qualities, throughput degradation due to congestion, and the services’ access prices offered by the ISP. It is assumed that the ISP runs a monopoly. The model of the dynamics of user adoption is formulated according to the user’s payoff for each of the service provided. The choice of the service is based on the user’s valuation of each option. The choices evolve over time according to fluctuations in the congestion levels. It is shown that user adoption reaches a unique, stable equilibrium, and that population condition in the ISP’s network is a critical factor for the adoption behaviour equilibrium.

Proposal baseline scenario for small cell traffic offloading

The ESR’s current research is focused on the financial aspects and the transactions between different participants that small cell traffic offloading incorporates. In this section a scenario is proposed by the ESR, aiming to bring the attention on a network sharing scheme that will become common in the future. This scheme refers to particular individuals that have in their possession small cell infrastructure, which can be utilized by the current mobile network operators and service providers. Such individuals could be municipal authorities, public services and schools or athletic organizations, which own such infrastructure. Following the guidelines set by the SoA presented in the previous section, the scenario is designed based on the competition that appears during network sharing arrangements, deriving from the selfish behavior of the different participants. Its objective is directed at the provision of satisfactory outcomes for all the participants of the network sharing scheme.

The scenario incorporates three main stakeholders: a Small Cell Operator (SCO) and two Macro Service Providers (MSP). Each MSP has in its possession a licensed band, whereas the SCO is solely an infrastructure provider. The scenario’s topology is depicted in Figure 4.24.

Figure 4.24: A general scenario topology

When the MSPs are unable of serving the amount of their offered load, they demand to use the SCO’s infrastructure. Since the SCO doesn’t possess its own licensed bandwidth, it leases the necessary portion from the MSPs, for its RAN operation.

Since there is a need for introducing a competitive environment for the MSPs, the offloading decisions, that is, the distribution of the SCO’s resources, will be made based on an auction scheme. The auction will be held for the time period the MSPs cannot serve their load. These periods are divided in timeslots (e.g. 1h), at the beginning of which, an auction is conducted. This scheme aims for the maximization of profits of the participants, while the subscriber’s QoS is guaranteed.
The challenge that derives from this implementation is the suitable choice of a bidding strategy by the MSPs, bearing in mind the scenario’s important characteristics. That is, the macro and small cell layer topology, the offered load and each participant’s finances. These variables determine the MSPs’ valuation for additional resources, for the timeslots an auction is held.

To address this challenge a learning mechanism will be used for the MSPs. This mechanism, having as inputs the MSPs’ valuations that change dynamically with every auction, will aid the auction scheme reach an equilibrium. This equilibrium, in turn, should be a state in which the system meets the aforementioned objectives.
5. Architectural aspects for SONs and cognition for decentralized network management methods, policies and algorithms

5.1 SON Architectures

A SON entity for Self-configuration will be created in operation, administration and management system (OAM) which will be responsible for the self-configuration of eNB. Self-optimization functions are flexible and can be located in OAM or eNB or both depending on the scenario and the use case [108]. Therefore, according to the location of optimization algorithms, SON can be classified into three types namely: Centralized, Distributed and Hybrid.

Centralized SON-In this case, optimization algorithms are executed in the OAM system. In such solutions, SON functionality is located in a small number of locations at a high level in the network architecture. All the SON functions are located in OAM systems, making it easy to deploy. But there is low support for optimization cases as the OAM system could be provided by different vendors. Existing Itf-N interface needs to be extended to implement Centralized SON. Figure 5.1 shows a basic representation of Centralized SON.

Distributed SON-Here, SON functionality exists in many locations at a relatively low level in the network architecture. SON optimization algorithms are executed partially or completely in eNB depending on the scenario being considered. Distributed SON takes more time to be deployed in the network and SON functions are located in the eNB. It offers quick optimization responses and faster execution of algorithms enabling real-time adaptation of network based on user requirements but it is also difficult to support complex optimization schemes which require coordination between a large number of eNBs. To efficiently and effectively implement Distributed SON solutions, X2 interface needs to be extended. Figure 5.2 shows a basic representation of Distributed SON.
Hybrid SON—In Hybrid SON, part of optimization algorithms are executed in eNB and others are executed in OAM system. Simple and quick optimization techniques can be implemented in eNB and complex optimization schemes can be implemented in OAM. This type of approach provides flexibility to support different kinds of optimization use cases. Optimization between network elements from different vendors can also be supported via the X2 interface. However, this approach could lead to increased deployment work, costs and considerable interface extension work. A basic representation of Hybrid SON is shown in Figure 5.3.

Figure 5.2: Distributed SON

Figure 5.3: Hybrid SON
5.1.1 SON-Load Balancing and Information Exchange for TD-LTE System

LTE is a standard for high speed data and voice wireless communication and is maintained as a project of the 3rd Generation Partnership Project (3GPP) [114]. LTE has been defined to accommodate both paired spectrum for Frequency Division Duplex (FDD) with sufficient frequency separation to allow simultaneous transmission and reception and unpaired spectrum for Time Division Duplex (TDD) operation as both transmission and reception occur on the same channel. In FDD, the transmission in uplink (UL) and downlink (DL) is simultaneous in different frequency bands. In TDD, the UL and DL data traffic are transmitted under the same frequency band, but within different time domain and a proportion of downlink and uplink regions could be instantaneously changed to support asymmetric traffic conditions. Therefore, the TDD scheme could be used to serve asymmetric traffic in future wireless communication systems. One of the features of TD-LTE system is the asymmetric UL/DL data rates due to the different UL/DL configurations. There are total 7 different UL/DL configurations defined in the standards for TD-LTE systems as shown in Figure 5.4.

<table>
<thead>
<tr>
<th>Sub-frame Number</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>UL/DL Config.</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>1</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>D</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>D</td>
</tr>
<tr>
<td>2</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>U</td>
<td>U</td>
<td>D</td>
</tr>
<tr>
<td>5</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>6</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>D</td>
<td>S</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>D</td>
</tr>
</tbody>
</table>

Figure 5.4: Uplink / Downlink sub-frame configurations for LTE TDD (TD-LTE)

Where:
- D is a sub-frame for downlink transmission
- S is a "special" sub-frame used for a guard time
- U is a sub-frame for uplink transmission

Fixed or static TDD uses one of the configurations and lacks flexibility in terms of resource scheduling in UL and DL, while dynamic/flexible TDD is relatively more flexible and considers switching different configurations, based on different demands of UL/DL data traffic or different applications’ requirements. To satisfy different applications’ requirements, the data rates in TD-LTE for UL and DL can be varied dynamically. There are various time slot allocation strategies and cell specific dynamic UL/DL re-configuration schemes that have been investigated in the past and many schemes are still under research. In contrast to the static TDD, the use of dynamic UL/DL configuration in different cells of TD-LTE system could give rise to a phenomenon termed as cross-slot interference. The cross-slot interference arises when either of the neighboring TDD cells have different time slot configuration. This means that the eNBs are required to be synchronized with respect to the uplink and downlink transmission times. If the neighboring base stations use different UL/DL configuration and share the same channel, then interference may occur between cells. Managing this cross-slot interference is an important and imperative task for supporting cell specific dynamic uplink/downlink time slot allocation and configuration. Figure 5.5 shows a representation of cross-slot interference.
For self-optimizing, load balancing and auto-tuning the network, UE and eNB measurements and performance measurements can be used. The process starts when the network is in operational state. The self-optimization process collects the measurement information from the UE and the eNB. It could then auto-tune the configuration data to optimize the network with the help of a logical or software defined entity. For efficient and effective load balancing and network management, information is required to be exchanged between the eNBs and different network elements. Several interfaces have been identified and standardized to exchange information for self-organization. These interfaces include:

- X2 interface: standardized interface for information exchange between eNBs
- S1 interface: standardized interface for information exchange between eNB and the core network—it is split into S1-MME for control data and S1-U for the user data traffic
- Ift-N interface: standardized OAM interface, between NEM layer and OSS (network management) layer—this interface provides a vendor independent description of the network functions to the OSS layer

The presence of self-organization entity in the network element (i.e., eNB with X2 and S1 interfaces), allows real-time execution of algorithms and does not require a large bandwidth resource for exchanging data as there are local X2 and S1 interfaces present in LTE-Advanced systems. The X2 and S1 interfaces between eNBs have significant impacts on SON performance. To optimize distributed SON performance, the X2 and S1 interfaces should be modified according to certain SON applications in 3GPP LTE-Advanced standards. On the other hand, because of the distributive characterics, it is difficult to modify this architecture, the existing organization strategy and algorithms as too many communication entities would then be required to be modified. Furthermore, the strategies and algorithms related to the cross correlation of parameters and performance indicators from adjacent communication entities are difficult to support because they require information exchange with each other.

The distributed SON has a characteristic feature of failure resilience because the outage or malfunction of a SON communication entity will affect only one node, but not a large number of network elements, as often happens in a centralized SON architecture.

The load balancing algorithms primarily aims to spread the traffic across neighboring cells to reduce the constraints experienced by the system arising from elevated baseband processing load or limited transmission resources (for e.g. resource blocks). Based on the status of the traffic load at the target eNodeB (eNB), the serving eNB needs to estimate the appropriate number of users that can be handed over to the target eNB. It is worth noting that the users will generate different load in the target eNB than they did in the serving eNB before the handover. Therefore before initiating handover to optimize load in the serving eNB region, it is important to know the existing traffic load in the target eNB region and evaluate the additional load that the newly shifted users will cause to the target eNB. A method based on SINR estimation has been proposed in [109] to predict the uplink and downlink load at the target eNB. Furthermore, the load balancing algorithm needs to decide which neighbor cell to
choose as target eNB out of a number of eNBs available in the neighbor list, and needs to know how much the handover offset should be adjusted to achieve the highest network performance gain.

Information exchanged between the network elements can be used to perform optimization of load balancing and handover parameter optimization. The optimization of load balancing and handover parameter optimization should be performed with least human intervention. For load balancing different scenarios can be considered for example overlapping coverage, hierarchical coverage and neighboring coverage. Each scenario is required to include load balancing on intra-frequency, inter-frequency and inter-RAT cases. Certain parameters have been identified which can be tuned to optimize and balance load in the network. However, tuning of some of these parameters may affect one or multiple optimization metrics such as interference, coverage, energy consumption, handover performance etc. This can be avoided by using advanced methods to further differentiate the relevance and interdependencies of the parameters affecting common optimization metrics. Some of these methods have been discussed in [113]. Following parameters have been identified which affects the load balancing optimization metric.

- Transmit Power
- Antenna Parameters
- Handover Parameters
- Cell Reselection Parameters
- Tracking Area Parameters

The aim here is to identify traffic imbalance in the serving cell and its neighboring cells, optimize the traffic load, change the uplink or downlink bandwidth to improve system performance, avoid QoS degradation, minimize the number of handovers needed to achieve efficient load balancing and reduce the number of blocked/dropped calls. If the load is found to be imbalanced, switching configuration in case of TD-LTE can be considered. Handover parameter optimization function is responsible for reducing inefficient use of network resources due to unnecessary handovers. It focuses on reducing number of handover related failures.

5.1.2 SON Interference Management

In mobile communications networks, the arrival rate of the users and the resulting traffic load are random, time-varying and often unbalanced. This leads to imbalanced load at different times in different cells. There might be some cells which are overloaded with a large number of UEs and other not so loaded with adequate number of UEs where resources are underutilized. Optimal network management and planning could mitigate the inefficient resource usage and improve the overall performance of the network. It is also required that the future wireless communication systems support various multimedia services, such as voice over IP (VoIP), video streaming, interactive gaming, peer to peer (P2P) file transfer etc. This could lead to significant traffic asymmetry. In such a scenario duplex schemes such as frequency division duplex (FDD) and time division duplex (TDD) should be able to serve asymmetric traffic without impacting the end user quality of service (QoS). In TDD, a proportion of downlink and uplink regions could be instantaneously changed according to the property of the traffic asymmetry. This property in LTE-TDD systems could be really helpful to serve the asymmetric traffic in the next generation communication networks.

The TDD scheme can be further divided into:

Static-TDD (S-TDD) - In S-TDD scheme, since the proportion of the numbers of downlink and uplink time-slots is fixed, downlink and uplink time-slots among adjacent cells do not cross. Hence, there are only two types of interferences, such as mobile station (MS)-base station (BS) (from MS to BS) interference and BS-MS interference. The MS-BS interference is caused by MSs in their uplink cycle included in adjacent cells when the serving cell is in its uplink cycle, and the BS-MS interference is caused by adjacent BSs in their downlink cycle when a serving cell is in its downlink cycle.

Dynamic-TDD (D-TDD) - the proportion of the numbers of downlink and uplink time-slots is variable, the downlink/uplink crossed regions occur among adjacent cells. Thus, besides MS-BS and BS-MS interferences, there are two additional interferences, such as MS-MS interference and BS-BS interference. The MS-MS interference is caused by MSs in their uplink cycle included in adjacent cells when a serving cell is in its downlink cycle, and the BS-BS interference is caused by adjacent BSs in their downlink cycle when a serving cell is in its
uplink cycle. The BS-BS and MS-MS interference in D-TDD cause severe performance degradation. In case of MS-MS interference, downlink SINR values for cell edge users in the serving cell decreases due to uplink transmissions of the cell edge users in the adjacent cell. Similarly, in case of BS-BS interference, uplink SINR values of all users in serving cell degrades severely due to downlink transmission of the adjacent base stations. Performance degradation of BS-BS interference is severe compared to MS-MS interference as the transmission power of base station is very high as compared to the uplink transmit power of the UE [110]. Therefore, by using just D-TDD schemes to support traffic asymmetry would lead to worst performance compared with S-TDD. Several versions of cell specific D-TDD schemes and schemes to mitigate interference have been proposed and studied, however, they do not take into account resource sharing and they do not address the issue of pseudo congestion. There are a number of techniques in use in LTE networks to address inter-cell interference, including:

- Inter-cell interference coordination (ICIC), which selectively reduces the power for sub channels in the frequency domain
- Enhanced inter-cell interference coordination (eICIC), where macro cells are complemented with Pico cells inside their coverage area (for hotspot in public places such as coffee shops, Airports, etc.)
- Coordinated multi-point transmission/reception (COMP), where interference is decreased on edge users of cells by jointly scheduling several cells with rather strong edge interference, or by joint transmission so that the reception power and service experience of a cell’s edge users can be improved

Interference coordination scheme intend to reduce interference in the network there by enhancing the network performance in terms of throughput and network capacity. With low interference, each user would require less bandwidth due to higher SINR (Signal to Interference and Noise Ratio). Several advanced interference mitigation and coordination techniques are under investigation with the aim to reduce complexity and enhance the system performance.

Interference is one of the major issues to address while serving asymmetric traffic in next generation communication networks. With increasing heterogeneity in either radio access technology, topological layers for example, macro, pico, femto or spectrum (multiple frequency bands, maybe aggregated), the issue of interference becomes even more critical. Several parameters have been identified that directly affects interference. Some of these parameters are:

- Transmit Power
- Resource Block Assignment
- Switching point configuration
- Adjustable beam forming parameters
- Channel Quality Indicator (CQI) thresholds for schemes switching

As in the case of other optimization metrics such as load balancing and handover parameter optimization etc. tuning of some of these parameters may be related or affect, one or multiple optimization metrics and may also have a direct influence on other parameters, hence it is required to further differentiate relevance and interdependencies of these parameters by using advanced methods as discussed in [113]. The choice of parameters and their new values must be carefully examined, requiring some degree of coordination between the parameters associated with different optimization metrics and the search process to find the sub-optimally configured regions in the network. This would lead to improved and efficient network optimization and resource utilization.

It has been observed that increase in traffic causes throughput degradation in UL and DL. There are several technical reasons for this performance degradation. One of the main reasons is that when the traffic becomes more intensive the probability of simultaneous DL and UL signal transmissions increases. This leads to increased cross-slot interference. If the neighboring cells are coupled (i.e. cause strong mutual interference) and DL traffic becomes dominant in the adjacent cell, the UL transmissions in the cell being considered (serving cell) become infeasible due to strong BS-BS inter-cell interference. Moreover, the degradation of the UL performance in turn has a negative impact on the DL throughput performance. In the case of BS-BS interference, UL users will demand larger amount of resources (as they have lower average throughput). Since UL and DL flows compete for available sub-frame resources, total amount of allocated DL resources will also decrease. Hence the impact of BS-BS interference is more severe and cause performance degradation in both UL and DL directions. Similarly interference causes degradation in throughput performance of the DL users but the impact is relatively less severe than the BS-BS interference.
5.2 SON Optimization and Management

5.2.1 Virtual Cell Concept - Overview

The concept of virtual cells in TD-LTE system enables users residing in overlapping coverage regions of different cells to utilize resources from multiple base stations. The benefits of virtual cells are realized by the efficient resource utilization, by matching the traffic demand with the network resource availability taking also into account the impact of cross slot interference. To address the issue of cross-slot interference, conventional power control methods similar to [115] could be employed. The use of advanced power control mechanisms could further improve the efficiency and performance of the virtual cell scheme. Virtual cells offer a novel way of resource management, which allows UEs to utilize sub-frames from multiple eNBs matching better their traffic demands. UEs are no longer restricted to use the frame of a single cell, but can utilize specific sub-frames from different overlapping cells. Hence, eNBs and UEs not only exploit the spatial domain of conventional load balancing but additionally use the time-domain to change the cell-setup.

For associating UEs to virtual cells, the UL and DL resources should be considered separately, allowing UEs to utilize UL and DL sub-frames from different cells. Fig. 5.6 illustrates a simple example of a UE residing within the virtual cell region, utilizing DL resources from eNB B and UL resources from eNB A in order to match its traffic demand. Specifically, a UE with a high UL and relatively low DL resource demand cannot be served by any of the depicted eNBs solely without experiencing and causing congestion. By assigning the UE to both eNBs and utilizing resources from both, pseudo congestion can be avoided. For instance a UE associated with eNB A to take advantage of the UL resources may switch to the DL of eNB in order to fulfill its DL demand, instead of remaining on the UL of eNB A. Such operation creates a customized or virtual frame for these particular UEs, which is composed of the shaded UL slots from eNB A and the shaded DL slots from eNB B as shown in Fig. 5.6.

![Figure 5.6: A simple example of virtual cell with the UE utilizing resources from both eNBs](image)

In real world practical scenarios, the resource requirements are not always equally high in both UL and DL directions at a given time instant, therefore the virtual cell scheme resolves the problem of pseudo congestion by using a hybrid configuration which is derived from the configuration being used by the serving eNB and its neighboring eNBs that work in collaboration with each other. Thus, the virtual cell scheme dynamically allocate adequate resource to serve the users located in overlapping outer regions of different cells referred to as virtual cells in both UL and DL directions based on the real time traffic demands thereby providing increased flexibility and improved resource sharing while maintaining the desired QoS (Quality of Service) for the end users. It is possible to have multiple virtual cell regions in an area with different hybrid configurations being used to serve the users in the virtual cell region. The structure of the hybrid configuration in this case will depend on the configurations in use by the serving cell and its neighboring cells considering also, the status of the interference in the region. The virtual cell scheme takes into account interference issues and it is possible to implement the virtual cell scheme in LTE-TDD systems with existing interference mitigation techniques to better serve the users under...
asymmetric traffic conditions without degrading the QoS. It is anticipated that advanced interference mitigation techniques could further improve the performance of virtual cell scheme in LTE-TDD systems.

5.2.2 Self-Organized Network Optimization and Management

The first step towards self-organized network optimization and management is to assess current network status and performance needs. For efficient and effective network optimization, accurate information is required to be collected from the network in operation. The information can be collected via inputs gathered from procedures such as network counters, measurement reports and drive tests. Network counters provide measurements taken at the eNBs such as level of load, number of users in handover, this information is then transferred to the OAM module. Information such as transmitted power level can be monitored at the eNB and passed to OAM module every 15 minutes or so in the form of average value or statistical distributions. The software tools to manage the collection, processing (e.g. measurement averaging) and transferring the information from network elements to the OAM module are usually provided by the vendors to the network operators. The type of measurements that a particular network element is able to perform is generally defined in the standards [106], [107]. Measurement reports are the observations taken by the mobile terminals or the user equipment (UEs) while they are active or in connected state in the network. These reports include measurements such as received signal power, channel energy over total received power density for the serving and neighboring cells etc. These measurements are then sent to the eNB and serves as an input for Radio Resource Management (RRM) algorithms. It is also possible to use these reports for the optimization process, in that case a logical entity or SDN function is needed which is able to manage, store and forward them to the OAM module or the responsible network entity. Drive tests involve measurements carried out by one or multiple specialized terminals which are able to record a number of parameters while following a pre-defined path in the field. As the drive tests are equipped with Global Positioning Systems (GPS), they are also able to note the position where each measurement has been recorded. If any misbehavior in the network is detected, drive tests can be triggered over the suspected areas to obtain further and precise information about the operating conditions. They can also be performed regularly as a network quality control and assessment procedures and serve as an input to the network optimization process. Drive tests performed with a frequency scanner are able to capture true coverage predictions without the influence of a specific network configuration. This can be useful in a heterogeneous scenario where LTE/LTE-A, 3G/2G, Wi-Fi etc. networks are present and if an issue or a problem arises in one of the network, traffic can be diverted via alternate networks thus masking the problem in the network where it occurred. Drive tests acquire similar information as the measurement reports but with an added advantage that the precise position where each measurement is taken is also recorded. While measurement reports mainly collect the status and information for the downlink, network counters are able to acquire the status and behavior for the uplink. In the case of drive test measurements, it is important to specify a pre-defined path as drive test measurements mainly exhibit the status of the downlink for the particular path followed by the testing vehicle. Therefore, it is very important to specify a path that ensures the acquisition of measurements in the areas where the real traffic is generated.

Observation stage involves collecting information from different available Key Performance Indicators (KPIs). The optimization procedure followed by collecting information in the observation stage contains different processes for each of the optimization targets specified based on individual operator policies for e.g. interference reduction, avoiding coverage holes, reduction of overlapping regions between the cells or in case of TD-LTE could be formation of virtual cells with dynamic cell specific configuration selection. For each optimization target following steps can be followed:

Processing the KPIs obtained during the observation phase which involves combining several KPIs, obtaining statistical measurements like averages or percentiles etc. Next step involves, identifying the sub optimal network operation region, which is the region where the optimization target is not properly configured, based on certain criteria related to the optimization target. After this, optimization search could be performed to find the appropriate parameter setting to solve the sub optimal network operation or serve the users in real time as per their requirements.

The final objective of the self-optimization process would be to autonomously find the appropriate configuration of selected network parameters (e.g. eNB powers, antenna down-tilts, RRM configuration parameters etc.) with least human intervention, affecting either the cell which is found to be sub-optimally configured and requires optimization and/or its neighboring cells. The changes in the configuration parameter(s) becomes the action executed over the network.

In TD-LTE systems, during the optimization process, observations and optimization information gathered can be used to select appropriate dynamic cell specific configuration in real time to serve the asymmetric traffic and adapt
5.2.3 Virtual Cell Distributed SON Resource Management

Once virtual cells are created, self-organized mechanisms to perform maintenance and management are essential to adapt to evolving traffic demands. The scope of such an operation is to assess the current UL/DL configuration in combination with virtual cell formations in order to perform potential alternations on both planes that enhance the system performance. A key feature is to consider the UL and DL load separately in relation to the UE geographical distribution within the serving cell. UL/DL location based information, which is beyond the conventional FDD cell summarized load, may potentially be derived from UE positioning [111]. Based on such information a TDD resource management SON function creates load statistics, which is used to rearrange virtual cell region(s) or re-configure the UL/DL ratio providing alternative virtual cell formation opportunities.

The proposed SON function is envisioned to be hybrid, executed partly on eNBs and OAM. In particular, the process of managing the formation of virtual cells could be handled in a distributed manner by each eNB, while the UL/DL re-configuration could be handled by a centralized OAM/SDN controller responsible for a group of eNBs. Hence, eNBs are required to regularly exchange load information associated to UL and DL resources via the X2 interface. Such UL/DL load information should be related to cell areas adjacent to virtual cell regions and not to the overall cell. Enhancements on the conventional X2 load-representation are also essential in order to specify the transmission direction. Once the distributed virtual cell management cannot ensure adequate resources based on the given traffic demand, it should notify the OAM/SDN controller responsible for the set of eNBs involved in the process. The OAM should then try to identify an alternative UL/DL configuration, which increases the potential virtual cell resources using another SON function.

Such a SON function could be executed both on demand to enhance the performance of virtual cells and/or periodically for maintenance purposes. Its input information contains UL/DL traffic demands and cell planning data, i.e. expected throughput, upper bound capacity and target KPIs (Key Performance Indicators) as well as the set of eNBs in case the function needs to resolve a specific problem on-demand. The core of the SON function consists of an optimization algorithm that aims to match the input constraints to a specific UL/DL configuration mode for a set of eNBs considered in the process. Local optimization solutions, e.g., Simulated Annealing [112] are good candidates. Once an UL/DL configuration mode has been selected, a network performance verification that involves conventional performance management operations is applied.

5.3 Conflict Handling

A conflict may arise when two or more SON algorithm attempt or request to perform contradictory actions in terms of adjusting same parameters or influencing metrics optimized by another algorithm unintentionally. The conflict can be categorized into mainly two types

Parameter Value Conflict- This type of conflict occurs when two or more SON algorithms request different values on the same parameter.

Optimization Metrics Conflict- Occurs when the metric optimized by one SON algorithm is unintentionally affected by one or more other SON algorithms aiming at optimizing other metrics.

It is worth noting that the optimization metrics conflict will only occur when a parameter in the second algorithm is affecting the result of the first algorithm without being present in the first algorithm. If a parameter is a part of both the algorithms, the conflict will be detected as a parameter value conflict.

When conflicts occur, a conflict resolution entity is required to be present to solve the issues. Conflict resolution entity may initiate optimization procedures in order to find an optimal trade-off between the actions of different algorithms or use prioritization. The prioritization could for example be between the goals or targets of different algorithms or between different parameter changes or values. The conflict resolution entity should be able to take inputs from the operator in order to enable operator policy driven behavior in the network.
6. Conclusion

This document showed the advances done by the consortium about the architectural aspects for SONs and cognition within the LTE-A standard, where the work from each ESR has been presented in detail. Each ESR carried out a full state of the art review to identify the current activities in the research community, as well as the open topics to perform research on them. Four topics have been presented within the SON technology: the D2D communication, the handover optimization, the traffic offloading and the virtual cells in TD-LTE system; where also the research target and future contributions have been presented as preliminary works.
References


[36] 3GPP TS 36.214 v11.1.0, “Physical layer; Measurements”, 2012


[40] 3GPP TS 36.410 v11.1.0, ”(E-UTRAN); S1 general aspects and principles”. Sep. 2013

[41] 3GPP TS 36.420 v11.1.0, ”(E-UTRAN); X2 general aspects and principles”. Sep. 2012


[44] 3GPP TS 36.300 v12.0.0, ”E-UTRA and E-UTRAN; Overall Description”. Jan. 2014


[48] 3GPP TS 22.220 v11.6.0, ”Service requirements for Home Node B (HNB) and Home eNode B (HeNB)”. Sep. 2012


[51] 3GPP TR 36.932 v12.1.0, ”Scenarios and requirements for small cell enhancements for E-UTRA and E-UTRAN”, Mar. 2013


[70] SOCRATES (Self-Optimisation and self-ConfiGuRATion in wirelEss networkS), available in http://www.fp7-socrates.eu/

[71] BeFemto (Broadband Evolved FEMTO Networks), available inhttp://www.ict-befemto.eu/

[72] 3GPP, TR22.220 (V11.6.0), “Service Requirements for Home Node B (HNB) and Home eNode B (HeNB)”, Sep. 2012


[88] Li, Yanwei; Zhou, Zhenyu; Nguyen, Nam; Sato, Takuro, "Game Theory Based Hybrid Access for Macrocell-Edge Users in a Macro-Femto Network," Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th, vol., no., pp.1,5, 2-5 June 2013


[90] Yanjiao Chen; Jin Zhang; Qian Zhang, "Incentive mechanism for hybrid access in femtocell network with traffic uncertainty," Communications (ICC), 2013 IEEE International Conference on, vol., no., pp.6333,6337, 9


[99] 3GPP TR 32.816, “Study on Management of Evolved Universal Terrestrial Radio Access Network (E UTRAN) and Evolved Packet Core (EPC)"


[106] 3GPP TS 25.215 “Physical Layer Measurements (FDD) (Release 5)”
[107] 3GPP TS 25.331 “Radio Resource Control (RRC) (Release 5)”