Physical Communication 10 (2014) 31–60

Contents lists available at ScienceDirect

Physical Communication

journal homepage: www.elsevier.com/locate/phycom

Full length article

LTE-Advanced and the evolution to Beyond 4G (B4G) systems

Ian F. Akyildiz *, David M. Gutierrez-Estevez, Ravikumar Balakrishnan, Elias Chavarria-Reyes

Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, United States

ARTICLE INFO

Article history:
Received 28 August 2013
Received in revised form 7 November 2013
Accepted 11 November 2013
Available online 20 November 2013

Keywords:
Carrier aggregation
Enhanced MIMO
CoMP
Relays
Heterogeneous networks
Machine-type communications
Device-to-device communications

ABSTRACT

Cellular networks have been undergoing an extraordinarily fast evolution in the past years. With commercial deployments of Release 8 (Rel-8) Long Term Evolution (LTE) already being carried out worldwide, a significant effort is being put forth by the research and standardization communities on the development and specification of LTE-Advanced. The work started in Rel-10 by the Third Generation Partnership Project (3GPP) had the initial objective of meeting the International Mobile Telecommunications-Advanced (IMT-Advanced) requirements set by the International Telecommunications Union (ITU) which defined fourth generation (4G) systems. However, predictions based on the wireless traffic explosion in recent years indicate a need for more advanced technologies and higher performance. Hence, 3GPP's efforts have continued through Rel-11 and now Rel-12. This paper provides a state-of-the-art comprehensive view on the key enabling technologies for LTE-Advanced systems. Already consolidated technologies developed for Rel-10 and Rel-11 are reviewed while novel approaches and enhancements currently under consideration for Rel-12 are also discussed. Technical challenges for each of the main areas of study are pointed out as an encouragement for the research community to participate in this collective effort.

1. Introduction

The fourth generation (4G) of wireless cellular systems has been a topic under discussion for a long time, probably since the formal definition of third generation (3G) cellular systems was completed by the International Telecommunications Union (ITU) in 1997. Upon completing the development of the 3G family of standards, the Third Generation Partnership Project (3GPP) started working on Long Term Evolution (LTE) systems during the Release 8 (Rel-8) of the standards [1]. Being the first cellular system based on Orthogonal Frequency Division Multiple Access (OFDMA), it represented a major breakthrough in terms of achieving peak data rates of 300 Mbps in the downlink [2,3]. However, both LTE Rel-8 and Rel-9 specifications did not meet the IMT-Advanced requirements established by the ITU for 4G systems [4]. LTE-Advanced, the first accepted 4G system whose standardization was initiated in Rel-10 by 3GPP, was born as the resulting efforts of 3GPP to meet those requirements [5–8]. Major performance goals included peak rates of 1 Gbps in the downlink and 500 Mbps in the uplink. A summary of the relationship among the requirements of LTE, LTE-Advanced and IMT-Advanced is shown in Table 1.

However, current predictions for future systems point out tremendous challenges far beyond what the ITU initially established for 4G. Driven by both the explosion of users' demands for mobile data along with new services and applications, and the need for a ubiquitous and wirelessly accessible cloud platform, the evolution of future mobile traffic is expected to boom. Applications and services demand ever-increasing data rates. A traffic growth of up to 30 times has been predicted to take place

* Corresponding author. Tel.: +1 404 894 5141; fax: +1 404 894 7883.
E-mail addresses: ian@ece.gatech.edu (I.F. Akyildiz), david.gutierrez@ece.gatech.edu (D.M. Gutierrez-Estevez), rbalakrishnan6@ece.gatech.edu (R. Balakrishnan), elias.chavarria@ece.gatech.edu (E. Chavarria-Reyes).

1874-4907/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.phycom.2013.11.009

Table 1
LTE, LTE-Advanced and IMT-Advanced performance targets for downlink (DL) and uplink (UL).

Item	Trans. path	Antenna conf.	Rel. 8 LTE	LTE-Advanced	IMT-Advanced
Peak data rate (Mbps)	DL	8 × 8	300	1000	1000
	UL	4 × 4	75	500	–
Peak spectrum efficiency (bps/Hz)	DL	8 × 8	15	30	15
	UL	4 × 4	3.75	15	6.75
Capacity (bps/Hz/cell)	DL	2 × 2	1.69	2.4	–
		4 × 2	1.87	2.6	2.2
		4 × 4	2.67	3.7	–
	UL	1 × 2	0.74	1.2	–
		2 × 4	–	2.0	1.4
Cell-edge user throughput (bps/Hz/cell/user)	DL	2 × 2	0.05	0.07	–
		4 × 2	0.06	0.09	0.06
		4 × 4	0.08	0.12	–
	UL	1 × 2	0.024	0.04	–
		2 × 4	–	0.07	0.03

between the years 2010 and 2015 [9]. By 2016, more than 10 exabytes of traffic per month will be circulating across cellular networks and more than 4 billion 3GPP wireless subscriptions will be operating in the network [10]. With these forecasts in mind, it becomes critical to provide not only very high broadband capacity, but also efficient support for a variety of traffic types, flexible and cost-efficient deployments, energy efficient communications strategies, robust systems against emergencies, and a balance between backward compatibility and future enhancements.

To progressively tackle these challenging requirements, 3GPP has been organizing its work based on releases [10].[1] Fig. 1 shows a roadmap of all releases spanning LTE-Advanced and the main technologies that were developed in each of them. Rel-10 started early in 2010 and was functionally frozen in March 2011 after its approval by the ITU for having met all the requirements for IMT-Advanced. Technologies introduced during that release include carrier aggregation for transmissions in several frequency bands, enhanced multiple input-multiple output (MIMO) techniques, relays, and self-organizing networks (SON). Then, Rel-11 was started and further enhancements were included to the basic LTE-Advanced technologies developed for Rel-10. The major contribution during Rel-11 was cooperative multipoint transmission and reception (CoMP), which allows different cells to cooperate for serving users. In addition, several important enhancements were introduced for heterogeneous networks (HetNets) such as enhanced inter-cell interference cancellation (eICIC) and mobility management enhancements. After the recent completion of Rel-11, the standardization work has started focusing on Rel-12 LTE-Advanced with a preliminary calendar that lasts through 2013 until its predicted freezing in September 2014. For Rel-12, further enhancements and new technologies are being proposed at the 3GPP meetings [11,12]. Although no actual conclusions have been drawn yet, there is a clear view on which key technologies will be addressed in Rel-12, as shown in Fig. 1. Enhancements to CoMP (inter-site CoMP), carrier aggregation (multi-stream carrier aggregation), and MIMO (Full-Dimension MIMO) are key items in the agenda of 3GPP. Moreover, radically new technologies will also be introduced: Machine-Type Communication (MTC) will enable machines to interact among themselves as part of a large network, and Device-to-device (D2D) communication will allow mobile users to interact with each other without the need to go through the network.

Depending on the progress in Rel-12, future releases will start earlier or later in 2014. Rel-13 is expected to further enhance LTE-Advanced technologies while Rel-14 and Rel-15 could potentially define a new access technology [13]. As of today, these technologies comprise what is generally known as Beyond 4G (B4G). Although no clear definition for B4G has been stated to date, it is understood that the evolution to B4G will be characterized by the advanced technologies of Rel-12 and beyond, paving the way to much higher performance. In addition, higher frequency bands are expected to be used in B4G systems. In a shorter term, these may include the bands between 3 and 5 GHz while in the long-term even bands up to 60 GHz. Bands at these frequencies are characterized by higher path loss, thus leading to systems with base stations of much smaller coverage areas. However, these are still only rough ideas on the possible directions. Nowadays, we can talk about an evolution towards B4G but not about the precise definition of B4G itself.

This survey paper contains a comprehensive overview of all the main key technologies that play an important role in the different releases spanned by LTE-Advanced so far. It also includes new concepts currently under discussion at 3GPP and research challenges that will pave the way towards B4G. The remainder of this paper is organized as follows. First, the core technologies of Rel-10 and Rel-11 are described. Section 2 presents carrier aggregation. Enhanced MIMO techniques are shown in Section 3 followed by CoMP in Section 4. Fundamentals of relays, HetNets and SON are explained in Sections 5–7, respectively. Then, further key technologies of Rel-12 and B4G are introduced. Section 8 presents MTC and Section 9 does the same with D2D. Finally, the conclusions are drawn in Section 10.

2. Carrier aggregation

One of the most effective methods to improve the network performance is to increase the amount of utilized bandwidth. Therefore, in order to meet the requirements of IMT-Advanced as well as those of 3GPP operators, LTE-Advanced considers the use of bandwidths of up to 100 MHz in several frequency bands. These bands are set by the ITU for IMT, and include the following [14]: 450–470,[2] 698–960,[3] 1710–2025 MHz (see footnote 2), 2110–2200 MHz (see footnote 2), 2300–2400 MHz (see footnote 2), 2500–2690 MHz (see footnote 2),

[1] Note that although on this article we focus on LTE-Advanced, a release contains enhancements for all past and present 3GPP systems.

[2] To be used globally for IMT systems.

[3] To be used in specific countries.

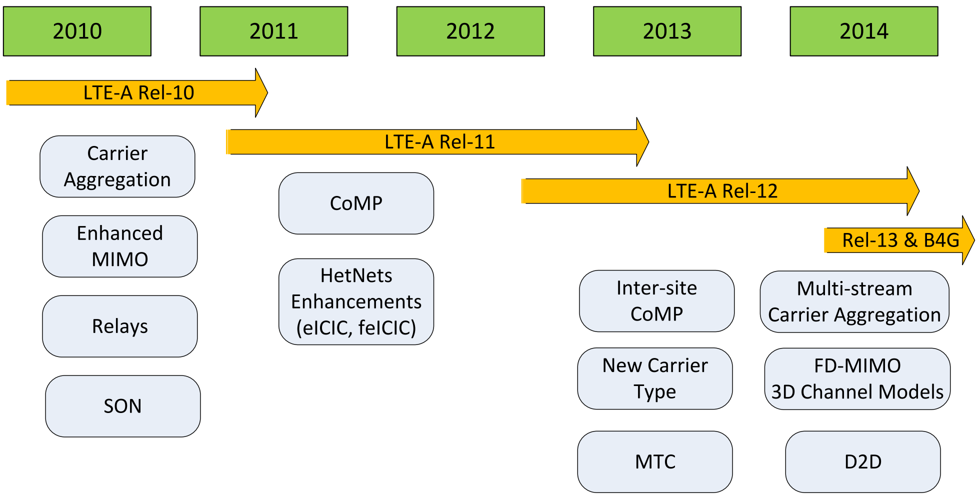


Fig. 1. Releases roadmap.

3400–3600 MHz (see footnote 3). Supporting several frequency bands increases the flexibility of LTE-Advanced. However, it also leads to fragmentation, where different countries utilize different frequency bands to serve users. As a consequence, a user equipment (UE) that works in one country or region may not work in another. This can be tackled from a technical point of view by designing devices that support multiple frequency bands. Nevertheless, this has the drawback of increasing the cost of devices. Regardless of the issues, LTE-Advanced tries to exploit as much as possible the flexibility of supporting multiple frequency bands through the use of carrier aggregation.

Carrier aggregation (CA) consists of grouping several "component carriers" (CC) to achieve wider transmission bandwidths. An LTE-Advanced device can aggregate up to five CCs, each of up to 20 MHz. With the largest configuration, this implies a total bandwidth of 100 MHz. To support backward compatibility with LTE devices, CCs shall be configured as a typical LTE carrier. Therefore, any of the CCs used for CA should also be accessible to LTE UEs. Nevertheless, mechanisms, such as barring [15], already exist in order to prevent LTE UEs from camping on specific CCs. This way, operators have the flexibility of adjusting the characteristics of the CCs to support a mixture of LTE and LTE-Advanced devices.

LTE-Advanced supports three schemes of carrier aggregation. The basic scheme of CA occurs when contiguous CCs are aggregated, as shown in Fig. 2. It depicts the aggregation of five CCs of different bandwidths, where the two rightmost are used exclusively by LTE-Advanced devices and the rest are shared among LTE and LTE-Advanced devices. While this scenario is the easiest to implement from a technical point of view, operators do not always have enough contiguous spectrum to perform this type of deployment.

Therefore, non-contiguous CA is also supported. In this case, the CCs may belong to the same or to different spectrum bands, also called intra-band CA and inter-band CA, respectively. These two scenarios are depicted in Figs. 3 and 4. These two types of CA are extremely useful to operators that have fragmented spectrum along multiple frequency bands, since they can effectively reuse their spectrum fragments to provide improved service to their users.

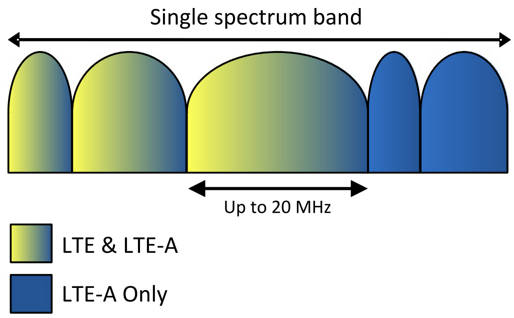


Fig. 2. Intra-band contiguous carrier aggregation.

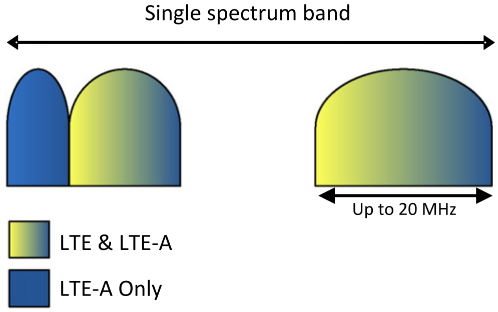


Fig. 3. Intra-band non-contiguous carrier aggregation.

In these three cases, CA can be used for both FDD and TDD systems, with some minor constraints. While the aggregation of CCs that work in TDD implies that the number of uplink and downlink CCs is the same, the aggregation of those that work in FDD allows the use of a different number of CCs for uplink and downlink. The rationale behind this is that users mostly consume rather than produce data. Therefore, most of the time downlink traffic is greater than uplink traffic. The only restriction established by LTE-Advanced is that the number of CCs

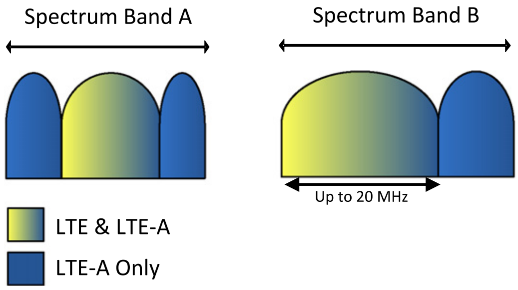


Fig. 4. Inter-band non-contiguous carrier aggregation.

used for downlink must be no less than the number of CCs used for uplink.

Efforts to support CA can be separated into two categories. The first one is the "common part", and involves all the aspects that are independent of the frequency bands that are utilized for CA. We will discuss the most relevant aspects later on in this section. The second one is the "specific band combination part", and involves addressing the issues associated with the specific operating bands that are used for CA.

3GPP defined more than 40 operating bands for LTE and LTE-Advanced [16]. Exploring all possibilities of CA within these 40 operating bands implies a huge number of options. Therefore, 3GPP has focused its efforts on a subset of all possible combinations. Up to Rel-11, 3GPP had studied the use of five (5) bands for contiguous CA [17], and the use of twenty (20) band combinations for inter-band CA [18], leaving intra-band contiguous CA for Rel-12. Within Rel-12, 3GPP is studying three (3) additional bands for intra-band contiguous CA, four (4) bands for intra-band non-contiguous CA, and eleven (11) additional bands for inter-band CA [19], showing its commitment to enabling CA support in as many regions as possible.

2.1. Benefits of CA

In addition to supporting wider bandwidths, carrier aggregation also provides the following benefits:

- Inter-cell interference mitigation.
- Handover improvement.
- Energy savings.
- Load-balancing.

First, CA is an efficient tool for avoiding inter-cell interference. Consider CCs configured to have different cell sizes and coverage areas, as shown in Fig. 5. In this scenario, there are two base stations, also called enhanced NodeBs (eNBs) in 3GPP terms. The eNBs have overlapping coverage areas. However, the overlap occurs with different CCs. Therefore, continuous coverage can be provided with reduced inter-cell interference, which is especially important at the cell edge.

Second, CA improves handover performance. In the same scenario of Fig. 5, consider an UE initially served through CC1 and CC2, and moving from eNB1 to eNB2. As the UE approaches the cell-edge of eNB1, it loses eNB1s CC1 coverage but maintains eNB1s CC2 coverage. Shortly after, it can connect to eNB2s CC1. Once it completely abandons eNB1s total coverage area, the user can also connect to eNB2s CC2. Therefore, a smooth handover among the eNBs is achieved with no service interruption.

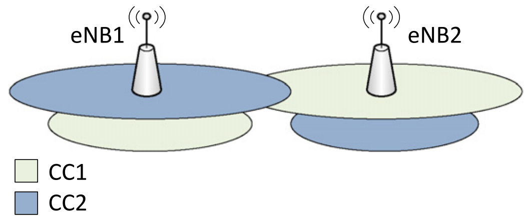


Fig. 5. Benefits of carrier aggregation.

Third, energy savings and load balancing can also be achieved through carrier aggregation. Consider Fig. 5. During the periods of low load, CCs can be turned off in order to save energy in the radio access network (RAN). Although simple, this is an extremely powerful solution since more than 70% of the energy in a cellular network is consumed by the RAN. Similarly, if a UE is served by any two CCs, and the traffic in one of the CCs becomes too high, the UE can be reconfigured to use a third CC instead of the overloaded one, to balance the load across the CCs.

In order to support carrier aggregation, 3GPP introduced modifications to the LTE protocol stack ranging from information elements modifications to functionality modifications [20]. The major modifications will be described below.

2.2. Protocol stack

To enable the support of carrier aggregation, certain modifications were introduced to the radio access protocol stack, as shown in Fig. 6. On the one hand, the packet data convergence protocol (PDCP) [21] and the radio link control (RLC) protocol [22] required no modifications. On the other hand, the radio resource control (RRC) protocol [23], medium access control (MAC) protocol [24] and physical layer (PHY) protocol [25] were modified. At the RRC level, a single connection is maintained independently of the number of CCs used by a UE. The CC that is used for establishing the RRC connection is called the Primary Cell (PCell). It provides Non-Access-Stratum (NAS) [26] mobility information, as well as security input. If the UE supports multiple CCs, then additional Secondary Cells (SCells) can be added to the user's set of serving cells. The addition and removal of SCells is performed through new procedures at the RRC level. One important aspect is that the PCell and SCell of different UEs need not be the same, i.e. the SCell of one UE may be the PCell of another UE.

At the MAC layer a single hybrid-ARQ entity is introduced per CC, and one transport block is sent per transmission time interval (TTI) from each serving cell (when spatial multiplexing is not being used). From the point of view of scheduling, which is also part of the MAC layer, both same-carrier and cross-carrier scheduling are supported. The first corresponds to the scheme used for LTE. The second allows an eNB to use a particular CC to assign resources contained in a different CC.

However, there are certain restrictions. First, the PCell is the only one that schedules the PCell's resources. It

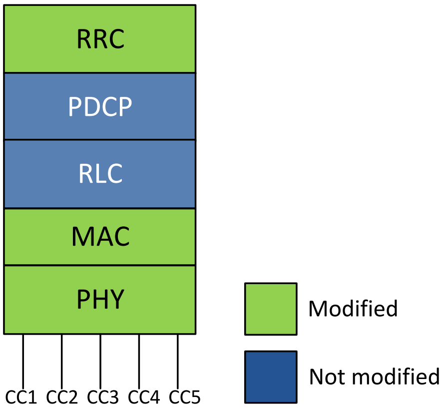


Fig. 6. Radio access protocol stack modifications to support carrier aggregation.

does so through the Physical Downlink Control Channel (PDCCH), at the PHY. Second, cross-carrier scheduling only applies when the PDCCH of a SCell is not configured. Cross-carrier scheduling is especially beneficial when the SCell's bandwidth is small, or when the PDCCH from a SCell cannot be reliably received.

Multiple timing advance was also introduced to improve the performance of CA. With multiple timing advance, UEs are able to independently advance or delay their transmission to each CC in order to compensate for the propagation delay. The concept of timing advance group (TAG) was introduced to refer to a group of CCs with the same timing advance. Therefore, if different CCs belong to different TAGs, then each one will have its own timing advance.

2.3. Research challenges

Carrier aggregation is an active research area, both in industry and in academia. The main issues are related to: the design of efficient transceiver architectures, managing all the available carriers to provide enhanced service to the users, and introduction of new carrier types.

2.3.1. New carrier types

One of the most important items under study for Rel-12 is the introduction of new carrier types (NCT) [27].

The main characteristic of the NCT is that the legacy control signaling and cell-specific reference signals (CRS) will be reduced (or even eliminated) in order to tackle limitations in legacy carriers. First, in legacy carriers, cell-specific reference signals are always transmitted, independently of the actual load in the cell. Therefore, the base station consumes energy and causes interference to other cells even when there is no load. In addition, the first symbols of each subframe contain a reserved set of resources that are used for legacy control channels. These resources cannot be utilized for data, even when no legacy control channels are present (e.g. in the case of cross-carrier scheduling). These limitations have motivated the introduction of NCT, especially for high-density heterogeneous networks.

In order to support a gradual introduction of NCT, two deployment scenarios are being considered. In the first scenario, the NCT acts as a SCell and is synchronized in time and frequency with a legacy carrier that would act as PCell. Work in Rel-12 has focused on defining the procedures to achieve this synchronization as well as the reception of PDSCH for both adjacent and non-adjacent NCT (with respect to the legacy carrier). For this scenario, extensive work has been done in defining the transmission modes that will be supported [28], while trying to keep compatibility with the enhancements that are also being introduced for MIMO and CoMP. In the second scenario, the focus is to have an NCT that is able to work in a stand-alone mode, with special attention to enhancements for small cell scenarios. Being stand-alone, these NCT will need to transmit either the Rel-8 PSS/SSS [29], or a new set of signaling and reference signals. Independently of the scenario, NCT is expected to be beneficial in low and medium loads due to the scalability of the signaling and reference signals in higher loads [30]. Therefore, for scenarios experiencing load variations from low to high, dynamically changing the carrier configuration from legacy to NCT may be required. Schemes that efficiently manage these transitions are needed.

2.3.2. Resource management

By introducing carrier aggregation, the decision on how to use the component carriers immediately follows [31]. In general, the network has to enable/disable the use of a subset of CCs in a subset of all the eNBs in the network [32]. This decision will be based on the different objectives that the network operator wants to achieve in a specific area. These objectives may include: improve the robustness of mobility procedures [33], coverage, capacity [34], and load balancing [35]; reduce energy consumption [36], as well as inter-cell interference.

These objectives also face certain constraints. For example, the QoS requirements (e.g. jitter, bit rate, delay) of the UEs should be satisfied while taking into account the capabilities of each UE and eNB (e.g. type of CA supported, MIMO capabilities [37–39], maximum power). However, the said objectives enter into conflict with each other most of the time. Therefore, designing techniques and algorithms that allow the operators to meet the specific requirements and resolve the aforementioned conflicts is a challenge in need of novel solutions.

2.3.3. Multi-stream aggregation

The initial scope of carrier aggregation was to enable a single eNB to serve its users utilizing two or more CCs. However, cellular networks have been evolving to become heterogeneous, where several small cells are deployed within the coverage region of the larger macrocell, as shown in Fig. 7. In this type of scenario, it is envisioned that an LTE-Advanced device will be capable of aggregating CCs belonging to the small cells and the macrocell. However, this scheme requires increased coordination between the small cells and the macrocell in order to avoid inter-cell interference.

As a further evolution of the multi-stream concept (also called multiflow), it is also being studied to support CA not

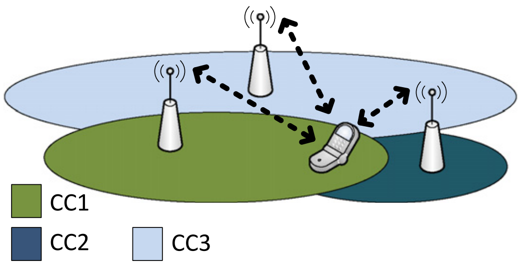


Fig. 7. Multistream CA.

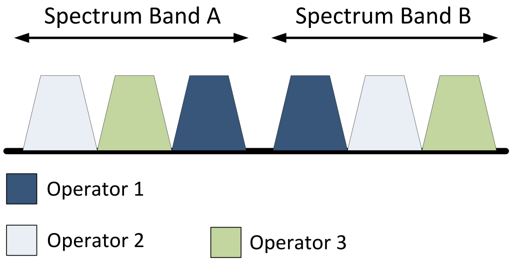


Fig. 8. Network sharing.

only among LTE and LTE-Advanced eNBs working in TDD and FDD [40], but also across base stations belonging to other technologies. This will provide significant improvement in the network, specially when there is an unbalance in the utilization of the radio access technologies (RATs). However, this technique must be carefully utilized since it also increases the amount of energy consumed by the UE, the amount of signaling in the air interface and the core network, and requires the introduction of new data aggregation entities common to the multiple RATs.

2.3.4. Network sharing

Within the context of 3GPP, the concept of network sharing allows multiple operators to share not only their equipments but also the RF frequencies that each one owns [41]. However, this type of cooperation occurs on a per-site basis. The natural evolution of CA is to allow operators to share their CCs in order to support cross-operator multi-stream CA. The first benefit of this scenario is that users belonging to these operators will have access to significantly higher bandwidths; therefore, higher capacity and coverage. A second advantage appears in allowing intra-band CA in scenarios where it was previously not possible. For example, consider the frequency allocation shown in Fig. 8 for three operators. In this scenario, each operator has access to two CCs belonging to two different spectrum bands; therefore, they can only support inter-band CA. However, through multi-operator multi-stream carrier aggregation, they would be able to support contiguous and non-contiguous intra-band CA. Schemes that facilitate this type of sharing dynamically in time and space are required in order to make these scenarios feasible.

2.3.5. Wideband transceivers

In order to support carrier aggregation, transceiver architectures must be designed in order to balance cost, energy consumption, and performance. On one extreme lies the option of processing all the frequency bands by using a single RF chain. On the other extreme lies the option of using a separate RF chain for each frequency band.

The first option is the simplest in terms of the number of hardware components. Therefore, it is expected to have a lower cost compared to the second option, as well as reduced energy consumption. However, it can only be used to support contiguous CA. On the other hand, the second option requires the greatest number of hardware components. Intuitively, this implies higher cost and energy consumption. However, it can be used for contiguous CA as well as intra-band and inter-band non-contiguous CA [42].

Solutions that balance these two approaches are still needed [16]. At the base station, one of the promising solutions for CA and future modifications to the standard is the use of software-defined radios (SDR). SDRs provide a hardware platform whose functionality can be changed by means of software updates. As such, changing a BS from supporting two-carrier contiguous CA to five-carrier non-contiguous CA may become as straightforward as doing a software update to the operating system of the eNB. Several open-source and closed-source projects have already embraced the SDR approach for eNB development, both for academia as well as industry [43–45].

2.3.6. Carrier aggregation enhancements

As was mentioned in Section 2.2, most of the modifications introduced to support CA were done at the MAC layer. Therefore, several procedures are still executed on a per CC basis. As such, suboptimal performance is achieved compared to a configuration where the procedures are done jointly across all CCs. For example, the use of CA increases the peak-to-average power ratio (PAPR), which has a significant impact in the power efficiency of the power amplifier, specially at the UE. Techniques such as codeworking mixing across CCs [46] and partial selective mapping (PSLM) [47] have been proposed to reduce the PAPR by exploiting cross-CC approaches. Similarly, enhancements for MIMO [38], precoding [48], rate matching [49] and discontinuous reception (DRX) [50] are being studied from a cross-CC point of view.

2.3.7. New frequency bands

In addition to the aforementioned research challenges, the evolution of Carrier Aggregation depends on the availability of more spectrum that can be aggregated. Two approaches are possible to acquire such spectrum: exclusive access and shared access. The exclusive access is the licensing approach typically followed by operators who become the sole users of a particular spectrum band. The shared access, also known as Authorized Shared Access (ASA) [51] or Licensed Shared Access (LSA) [52], allows an operator to utilize spectrum that has already being licensed to another incumbent user, but is being underutilized in certain areas and times. For example, in USA the 3.5 GHz band is used for the Naval radar. Therefore, such spectrum is expected to be heavily used near coastal areas, but not in other ones. Under ASA/LSA, the Coast Guard could establish an agreement with multiple operators allowing them to use the 3.5 GHz band in non-coastal areas. In the near future, ASA/LSA will allow operators to have access to highly-desired frequency bands without requiring incumbent users to relinquish their spectrum licenses and vacating the spectrum.

Regardless of the method to gain access to new spectrum bands, it is expected that in the short-term the band between 3 and 5 GHz will be utilized, and in the long-term even bands up to 60 GHz. Bands at these frequency are characterized by higher path loss, leading to smaller coverage areas. Therefore, the number of base stations working under the macrocell can be significantly larger than in any existing deployment, challenging current mechanisms of coordination among base stations.

For these new bands, it is being studied whether new radio interfaces would be more appropriate as well as new network architectures.

3. Enhanced MIMO

The use of multiple antennas both at the transmitter and the receiver is known as Multiple Input-Multiple Output (MIMO) and has become an essential technology for every current or future wireless system attempting to achieve very high data rates. Although MIMO techniques were already introduced in LTE systems [53], significant work has been carried out in the past years to boost the potential gains of this set of techniques for LTE-Advanced in Rel-10, Rel-11 and Rel-12 [54–56]. Therefore, the name of *enhanced MIMO* was selected for the set of LTE-Advanced techniques. After providing a brief introduction to the basic MIMO modes of operation, this section of the paper presents the major novelties in LTE-Advanced over past systems, highlighting both accomplishments and challenges. Furthermore, future MIMO techniques currently under investigation for their consideration in Rel-12 are also discussed.

3.1. Basic modes of operation

Different approaches can be followed for benefiting from the advantages of a MIMO transmission depending on the performance metric to be maximized. In LTE, four basic modes of operation were introduced, namely spatial diversity, spatial multiplexing, beamforming, and spatial division multiple access (SDMA). The first three comprise the set of Single-User MIMO (SU-MIMO) techniques since the antennas are utilized for serving a single user. SDMA is the basic approach for Multi-User MIMO (MU-MIMO), where several users are served on the same frequency and time resources. Fig. 9 shows all these four modes of operation in a simplified scheme.

Rel-8 and Rel-9 LTE allow for a maximum of 4 × 4 configuration in the downlink and 2 × 2 in the uplink. A brief description of each of the basic modes of operation in the context of LTE is provided as follows.

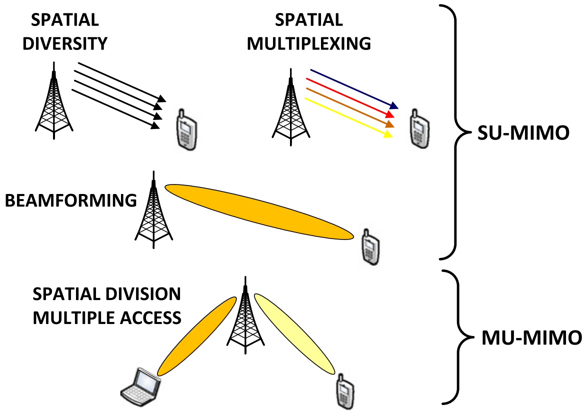


Fig. 9. MIMO basic modes of operation.

- *Spatial diversity*: Multiple antennas can be utilized as a means of improving the reliability of the transmission. To this end, multiple antennas at the transmitter transmit precoded versions of the same data stream, while multiple antennas at the receiver in diversity mode may perform different types of signal combining approaches to improve the received SINR. Diversity approaches supported in Rel-9 LTE include space-frequency block coding (SFBC) and frequency shift time diversity (FSTD).
- *Spatial multiplexing*: The user throughput can be increased via MIMO communication by simultaneously transmitting several data streams (also called layers). This is known as spatial multiplexing and requires precoding at the transmitter to orthogonalize the streams. Rel-8 LTE supports codebook-based precoding, where an indicator called Precoding Matrix Index (PMI) is fed back by the user for the eNB to select the appropriate precoding matrix. Both the basic spatial diversity and multiplexing modes described above rely on cell-specific reference signals used for both channel estimation and demodulation. In addition, the mode of dual-layer beamforming was introduced in Rel-9 LTE as a combination of spatial multiplexing and beamforming. It requires UE-specific reference signals which allow the utilization of precoding weights that are not restricted to a codebook, as reference signals are pre-coded together with the data.
- *Beamforming*: The directivity of the transmission can be improved by applying beamforming techniques to the array of antennas, thus properly reaching users at the edge of the coverage area. Despite utilizing several transmitting antennas, the rank (number of layers) of the transmission can be adapted down to a single layer thus transforming the codebook matrix into a codebook vector. Both codebook-based and non-codebook-based beamforming techniques are supported in Rel-8, and up to two beamformed layers can be transmitted in Rel-9, as explained above.
- *Spatial division multiple access (MU-MIMO)*: The overall system capacity of the network can be greatly increased if several users can be co-scheduled on the same frequency and time resources. This is the objective of the basic form of MU-MIMO supported in Rel-8 and Rel-9 LTE based on codebook precoding. However, this restricted approach does not have enough flexibility to cancel intra-cell interference.

3.2. MIMO features in Rel-10/Rel-11 LTE-Advanced

Most of the major enhancements over LTE MIMO were introduced in Rel-10, although during Rel-11 some further refinements were made. The maximum antenna configuration is increased to 8 × 8 for the downlink and 4 × 4 for the uplink. In addition, an important novel feature is the introduction of a dynamic framework for SU-MIMO/MU-MIMO switching at the eNB, where the mode of operation for each user can be adapted transparently to higher layers in a fast timescale. This feature is important to optimize system performance since the MIMO channel and the UE's SINR conditions, as well as the traffic demands, may vary from one subframe to another. Furthermore, although improved techniques in terms of advanced beamforming and precoding for both SU-MIMO and MU-MIMO have been introduced, the latter mode has been given preference since it is considered to be a key capacity-enabling feature to meet spectral efficiency targets in LTE-Advanced systems. These new features have implications on different system aspects, such as reference signals, feedback design, or precoding codebooks.

3.2.1. Downlink

No major contribution in Rel-10/Rel-11 LTE-Advanced has been introduced to support of SU-MIMO spatial diversity because the expected performance gains of an 8 × 8 diversity setup is marginal. Nevertheless, previous spatial diversity methods for a smaller set of antennas can be still applied by using the so-called antenna virtualization method [55]. This technique allows precoding of the physical array in such a way that the user perceives a transmit array of antennas of smaller size. On the other hand, expanded SU-MIMO spatial multiplexing to up to eight transmission layers has been introduced to increase spectral efficiency and data rates, mainly targeting users around the center of the cell. Extensive and flexible non-codebook-based precoding for multilayer transmission is a key feature, along with MIMO feedback based on PMI, Channel Quality Indicator (CQI), and Rank Indicator (RI).

A critical issue to enable enhanced MIMO techniques in the downlink is the design of suitable reference signals that minimize the overhead while allowing advanced multilayer beamforming techniques. For this reason, reference signals supporting channel estimation have been separated from those supporting demodulation, as Fig. 10 shows. There are two sets of reference signals: Channel state information (CSI-RS) and UE-specific for demodulation (DM-RS). The first set is cell-specific, sparse in frequency and time, and allows flexible configurations of multicell reference signal measurements via a common base pattern to avoid mutual CSI-RS collision. Fig. 10(a) shows the example of multicell four antenna port pattern. The DM-RS set, which is precoded with the same weights as data, has a higher resolution needed for demodulation but is only sent when data is transmitted on the corresponding layer. Up to eight layers can be supported, and those are multiplexed both in frequency and code. Fig. 10(b) shows an example of a 4-port pattern, where the combination of Code Division Multiplexing (CDM) and Frequency Division Multiplexing (FDM) is pointed out.

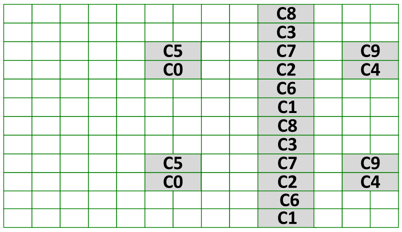
MU-MIMO greatly benefits from the increased number of antennas and the introduction of DM-RS based precoding. This is particularly true for high rank transmission with scattered users' locations since SU-MIMO performance is limited by antenna design constraints at the terminal. Multi-user interference cancellation via advanced beamforming techniques is a critical issue for the design of MU-MIMO. It is well known that Dirty Paper Coding (DPC) [57] offers optimal performance but its high complexity prevents it from being an option in real systems. Other proposed techniques for multi-user interference minimization are the following: (i) Zero Forcing (ZF), which can easily be implemented in practice by choosing the weight vectors as the pseudo-inverse of the composite channel matrix of the users to avoid interference among user streams [58], (ii) Maximum Signal-to-Leakage Ratio (SLR), an approach to design the beamforming vectors that does not impose a restriction on the number of available transmit antennas [59], and (iii) Block Diagonalization (BD), where every user's precoding matrix is restricted to be in the null space of all other users' channels [60]. However, further research with realistic CSI assumptions at the transmitter are still required.

3.2.2. Uplink

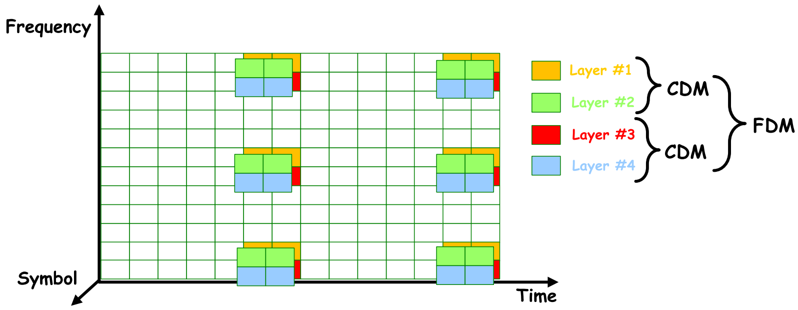
In the uplink, a maximum configuration of 4 × 4 is supported in Rel-10/Rel-11 to achieve spectral efficiency targets. For the control channel, transmit diversity is supported. For the data channel, spatial multiplexing of up to four layers (i.e., four data streams) is supported. The precoder to be utilized is selected at the eNB and notified to the terminal. However, the precoding codebooks are somewhat different from the downlink since a critical element in their design is the cubic metric (CM). The CM is a similar metric to the peak-to-average-power-ratio (PAPR) in the sense that it measures the signal fluctuations around the average value, but it also accounts for adjacent channel interference. In order to avoid signal distortion, the power amplifier at the UE needs to operate far from the saturation point, and therefore the CM should be preserved. Unfortunately, this design constraint comes at the expense of precoding gain. Reference signals are also separated in terms of their purpose: Sounding reference signals (S-RS) are employed for channel estimation purposes while demodulation reference signals (DM-RS) are used for data demodulation. Finally, an advanced MIMO receiver must be capable of eliminating both inter-symbol interference (ISI) and inter-layer interference for spatial multiplexing. Receiver options include minimum mean square estimation (MMSE), successive interference cancellation (SIC), and turbo equalizers [56].

3.3. Full-dimension MIMO for Rel-12 LTE-Advanced

New promising MIMO technologies have been identified in 3GPP for Rel-12 LTE-Advanced and beyond to provide significant network capacity enhancements. A promising concept currently under investigation is *Full Dimension MIMO* (FD-MIMO), where a large two-dimensional array of transmit antenna ports (16, 32, or 64) at the eNB makes use of the so-called active antenna system



(a) Multicell 4-port CSI-RS pattern.



(b) Multilayer 4-port DM-RS pattern.

Fig. 10. Reference signals.

(AAS) technology to provide accurate 3D beamforming to targeted users [61,62]. Full Dimension MIMO allows the transmission beams to be steered by the eNBs in both the azimuth and elevation dimensions, which introduces a higher degree of flexibility than traditional beamforming.

At the core of FD-MIMO lies the AAS. The main idea is to integrate the active transceiver component and the passive antenna array into just one radome [63]. Key benefits include the reduction of cable loss between the RF components and the antenna, reduced site costs of both maintenance and rental, and improved beamforming capabilities in the azimuth and elevation planes for a full utilization of the spatial domain. Rel-10 and Rel-11 MIMO techniques were designed to support antenna configurations at the eNB that are capable of adaptation in azimuth only, which prevents the beams from accurately focusing on the targeted users. Allowing the AAS to have additional control over the elevation dimension enables the eNB to direct beams in different horizontal and vertical directions, thus improving the accuracy of the beams directivity while allowing beams for different operations, dedicated tilts, frequency bands, network standards and link directions. Therefore, a variety of new strategies may be enabled. Examples are sector-specific elevation beamforming, advanced sectorization in the vertical domain, and user-specific elevation beamforming.

In addition, 3D beamforming based on AAS is an efficient way to expand capacity and boost coverage. While traditionally beamformed users located far from the main beam may suffer transmission loss in the vertical domain, 3D beamforming is able to steer the beam in both horizontal and vertical domains for a clear propagation path to the targeted user. Energy can be also saved by focusing the transmission power. Both MU-MIMO and cooperative transmissions can benefit from this technology as well. More accurate nulling of multi-user interference can be performed and cooperative beams can be steered by different eNBs to put the targeted users in the spotlight-paths of multiple beams.

Fig. 11 shows a sample scenario for the FD-MIMO concept. Three different potential uses of FD-MIMO are shown: First, a smoother experience of users benefiting from eNB cooperation can be achieved with well-focused 3D beamforming. Second, users at the same horizontal position in the cell but different vertical positions are simultaneously served by means of vertical beam adaptation. Third, vertical coverage is extended in scenarios such as tall buildings in dense urban areas.

Multiple research challenges need to be tackled before FD-MIMO technology becomes applicable. An overview of those along with other MIMO challenges for performance enhancement is provided in Section 3.4.

3.4. Research challenges

Although MIMO is a well studied communication technique, new research challenges keep arising as the technology moves forward. In this section, we provide an overview of the most relevant current research issues for the case of next generation cellular communications. In particular, we focus on massive MIMO, 3D channel modeling, and feedback enhancements, reference signals and codebook design issues.

3.4.1. Massive MIMO

A current topic of interest in communications is the use of massive MIMO, also known as very large MIMO, where systems make use of antenna arrays with a large number of antenna elements. Although sometimes arrays of more than 64 antenna elements are already considered examples of massive MIMO, it is usually assumed that the number of elements is at least an order of magnitude higher than in systems built today, i.e., 100 or more. These large arrays can accommodate a large number of users, reduce interference, and greatly improve cell edge users' performance and coverage [64]. In addition, they have the potential of radically reducing the amount of transmission power and achieve better performance under the same regulatory power constraints [65].

Nevertheless, massive MIMO systems introduce new challenges that are currently being investigated and require further research. In the first place, the so-called pilot contamination problem has been assessed as the ultimate performance-limiting factor in massive MIMO cellular system [66]. This interference effect appears in multicell scenarios with non-orthogonal pilots in different cells, since channel estimates carried out in a given cell will be impaired by pilots transmitted in other cells. When the number of antennas increases without bounds, uncorrelated noise and fast fading vanish, hence causing

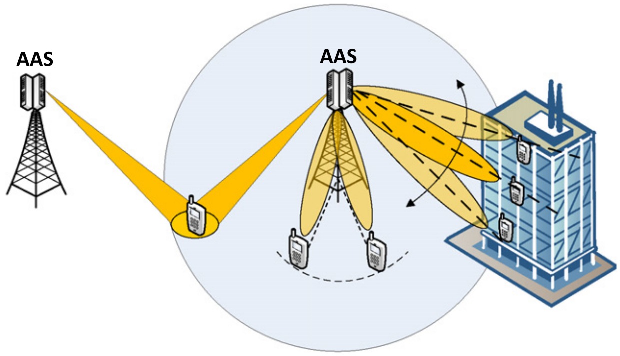


Fig. 11. Full-dimension MIMO concept.

the pilot contamination to appear as the main fundamental challenge of very large MIMO system design [67]. In addition, the complexity of the massive MIMO detector is also a relevant challenge, since the symbol search space is huge. For the particular case of MU-MIMO where the number of single-antenna users and base station antennas are about the same but large, the uplink detection problem becomes extremely challenging as the complexity becomes exponential with the number of users [65]. Finally, MIMO arrays have the ability to perform very narrow beamforming, thus opening the door for compensation of path loss in very high frequencies. However, this type of beams are problematic for cellular users since, due to the narrowness of the beams, they cannot easily track users when these are moving. Hence, mobility solutions for users under very narrow beamforming are needed.

3.4.2. Three-dimensional (3D) channel modeling

Most of the research on MIMO carried out so far has focused on the simplified case of a two-dimensional (2D) channel model. Alternatively, an assumption was usually made stating that most of the channel paths are concentrated on the 2D horizontal plane [68]. However, when the incident power is rich in the elevation domain or the antenna arrays are also vertically distributed the former assumption does not hold. The latter case is particularly important for massive MIMO arrays, where it is needed to accommodate the antenna elements in compact 2D arrays suitable for base stations with physical space limitations. With such layout of the antennas, the base station is capable of forming beams in the vertical domain (also called 3D beamforming as explained in the previous section). However, for those schemes to work there is a need for accurate and reliable 3D channel models that take into account both the azimuth and elevation directions in the signal propagation path from a base station to a user. So far, limited work has been carried out in this respect [69]. 3GPP considered the 3D channel modeling in the form of large-scale 3D sectorization beam patterns, modeled by combining static azimuth and elevation antenna patterns [70]. However, factors such as the multipath fading characteristics in both the azimuth and elevation dimensions must be included, and elevation dimension statistics must be correlated with other large-scale fading parameters.

3.4.3. Feedback enhancements, reference signals, and codebook design

Critical aspects for the performance of any cellular MIMO scheme are the channel information feedback, the reference signals and the precoding codebooks. Further optimizations need to be achieved in the design of all these three elements. For example, a better tradeoff solution for the feedback accuracy versus the incurred overhead needs to be found. Finer spatial-domain as well as frequency-domain feedback granularities are required with support for different antenna polarizations and configurations. Optimized feedback solutions for heterogeneous networks are also missing, and uplink feedback channel enhancements should be further explored. However, besides the need for further optimized solutions, also the introduction of FD-MIMO will have an impact on the design of feedback schemes, reference signals, and precoding codebooks. As an example, support for some elevation beamforming approaches may require modifications to the design of CSI-RS due to the 2D port structure of the transmitter. The methodologies for UE's CSI measurement and feedback may also need to be revised. Furthermore, new codebooks may have to be defined for the base station to support 2D array structures.

4. Cooperative multipoint transmission and reception

A key tool to improve system efficiency, cell-edge throughput and coverage of future 3GPP cellular networks is coordinated multipoint transmission and reception (CoMP) [5,71,72]. Although the concept started to be discussed in Rel-10 LTE-Advanced, it was not included in the specifications until Rel-11 due to initial high complexity concerns. The main principle of CoMP is based on the idea that UEs close to the cell-edge region can benefit from the downlink cooperation of multiple cell sites, thus transforming a problem of interference into an opportunity for improved service. In an analogous way, CoMP can be applied in the uplink in such a way that multiple sites coordinate the reception of signals coming from the UEs. As Sections 4.1 and 4.2 will show, different schemes for cooperation exist that range from a simple scheduling coordination to more complex joint signal transmission.

Different CoMP architectures can be envisioned, depending on whether the cooperation takes place intra-site (i.e., within the same eNB) or inter-site (i.e., among different eNBs). Intra-site CoMP was achievable in earlier releases since communication takes place just within one site and does not involve backhaul. Furthermore, the deployments can be carried out in homogeneous or heterogeneous environments, implying that the cell can have same or different coverage area sizes. According to these possible architectures, 3GPP has considered the utilization of CoMP in four different scenarios, as Fig. 12 depicts. Scenarios 1 and 4 have been included already in Rel-11 while scenarios 2 and 3 are currently being developed within Rel-12 [10]. They are described as follows:

- *Scenario* 1: *Intra-eNB CoMP in homogeneous deployment*. Coordination is performed within the several cells of the same eNB in an homogeneous scenario consisting of regular eNBs. No eNB interconnection links are needed in this scenario.
- *Scenario* 2: *Inter-eNB CoMP in homogeneous deployment*. The coordination area is extended to include other macrocells managed from other sites. The base station entities can be high-power remote radio heads (RRH) managed by an eNB as depicted in Fig. 12 or full eNBs. The decision is network implementation-specific. Low-latency backhaul links are needed to interconnect the sites and perform the coordination.
- *Scenario* 3: *Inter-cell CoMP in heterogeneous deployment*. In this scenario, base stations of different transmission powers coexist interconnected with high-capacity backhaul connection. A typical example of this heterogeneous deployment is a macrocell and several picocells deployed within the macrocell coverage area. Different approaches exist for the implementation of this scenario. Fig. 12 shows an example in which a single eNB controls the low-power RRHs of picocells, but these can be also implemented as separate eNBs and directly cooperate among each other.
- *Scenario* 4: *Distributed antenna system (DAS) with shared cell ID*. The same physical cell identifier is used in this scenario for both macrocell and low-power RRHs, thus effectively creating a DAS where mobility is greatly simplified. However, high-capacity backhaul links are still required to interconnect the sites.

4.1. Downlink CoMP techniques

In the downlink, cooperative communication techniques have been introduced to improve the received signal quality at the UE and reduce the co-channel interference [73]. Downlink CoMP transmission is carried out by means of a set of eNBs which coordinate among themselves to transmit simultaneously to the same user or set of users. Different approaches with different levels of coordination exist. Their requirements in terms of measurements, signaling, and backhaul are different. As usual, the highest performance achieving schemes require the highest system complexity. Fig. 13 shows the three main approaches introduced in Rel-11, which will be further explained in the rest of the section.

4.1.1. Coordinated scheduling/coordinated beamforming

In coordinated scheduling/coordinated beamforming (CS/CB) mode, each UE is served by a single cell known as the "anchor cell". However, scheduling decisions and precoding at each eNB to achieve beamforming may be coordinated among different eNBs to reduce inter-cell interference and improve the sum throughput. As Fig. 13(a) shows, this is accomplished by exchanging control information through a dedicated backhaul connection (X2 interface). An example of CS/CB is dynamic point blanking (DPB), an approach that switches the transmitter on or off depending on whether the interference it causes to the system is beneficial for the system performance or not. The feedback design must be enhanced to give support for this transmission strategy. Furthermore, although the scheduler at each eNB makes its decisions independently, additional information about other users' channel conditions is necessary in order to make more optimal decisions. Since the amount of information that needs to be exchanged is small, this approach is the most suitable one for non-ideal backhaul links, i.e., when the backhaul link incurs a non-negligible latency and its capacity is finite.

4.1.2. Dynamic point selection

Dynamic point selection (DPS), also called Dynamic Cell Selection (DCS) or Transmission Point Selection (TPS), is a scheme where the transmission to the intended UE only takes place from one point at a time, as Fig. 13(b) shows. This point must be drawn from the CoMP cooperating set serving the same UE. The switching is performed according to the wireless resources availability and channel state information at most on a subframe-by-subframe basis, thus the dynamic change in the transmission point that is transparent to the UE. The related radio resource management, packet scheduling, and common channels are tasks always performed by the serving cell. The fact that no more than one eNB transmits at the same time implies that there is no need for the eNBs to have a tight phase synchronization. Hence, DPS can be implemented with more relaxed RF performance requirements. However, a fast backhaul link among the transmitters is needed to share data and coordinate the scheduling in realistic traffic conditions.

4.1.3. Joint transmission

Joint transmission (JT) refers to a CoMP mode of operation where data intended for a particular UE is simultaneously transmitted from multiple sites to coherently or non-coherently improve the signal quality received in a time–frequency resource, as shown in Fig. 13(c). The JT scheme primarily considers that the transmission points correspond to different cell sites and a cluster of base stations interconnected via a fast backhaul link must jointly decide on the transmission scheme to the UE. In the coherent approach, the network must obtain channel state information (CSI) from all the cooperating cell sites. The phase and possible amplitudes of the transmitted signals can be adjusted to the CSI in such a way that the receiver is able to combine them coherently at symbol level. While combining is an implementation issue, it affects the signaling support design for the CoMP downlink. On the other hand, for non-coherent transmission the network does not have information concerning the relationship of the channels among the cooperating cells. Hence, the UE would receive multiple transmissions individually and independently precoded by each transmitter. A final interesting re-

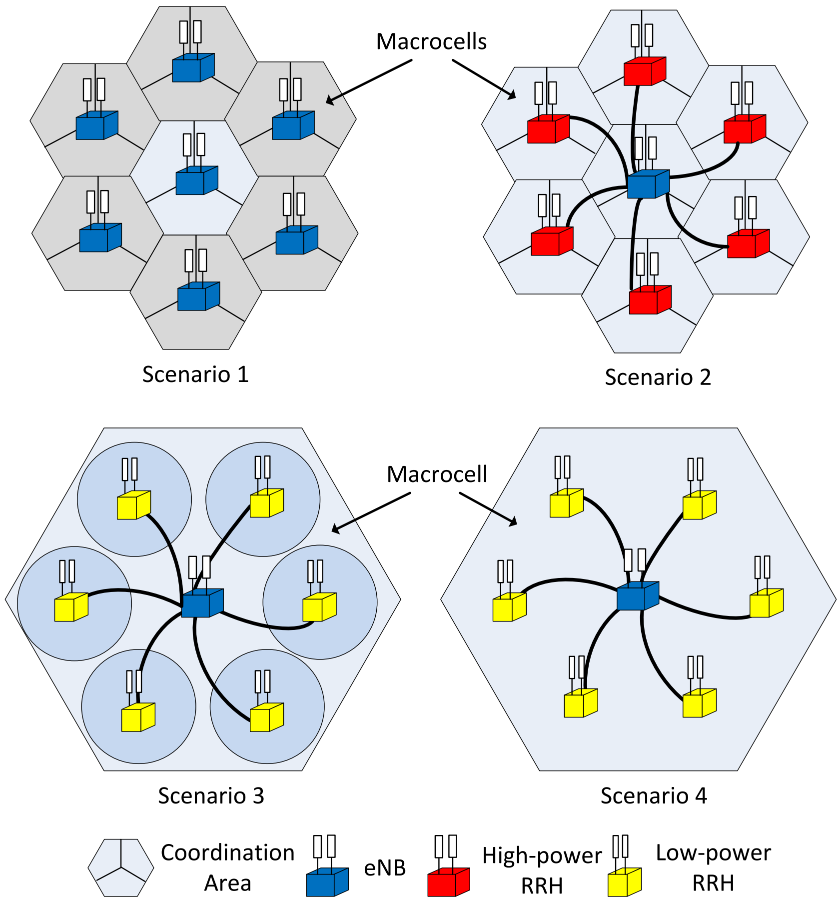
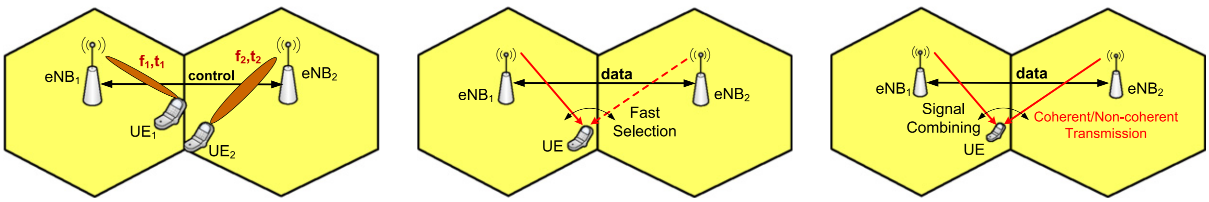


Fig. 12. CoMP scenarios considered within 3GPP.



(a) Coordinated scheduling/coordinated beamforming.

(b) Dynamic point selection.

(c) Joint transmission.

Fig. 13. Downlink CoMP approaches.

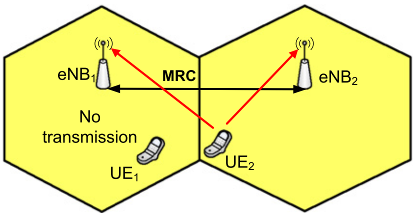
mark is the fact that a combination of DPS and JT is also possible. This means that a set of coordinating sites can be dynamically selected at each time instant.

4.2. Uplink CoMP techniques

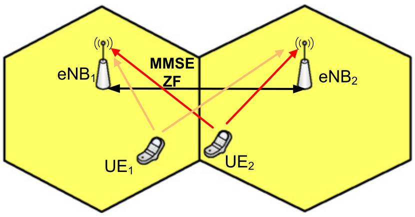
The uplink CoMP scheme is aimed at increasing the cell-edge user throughput in the uplink direction. It considers the reception of the signal transmitted by UEs at multiple and geographically separated points. These points are the set of coordinating eNBs assigned to each UE. Generally speaking, the terminal does not need to be aware of the nodes that are receiving its signal and what processing is carried out at these reception points. It is mostly an implementation issue, so CoMP reception is expected to have limited impact on the specifications, and no major changes in the radio interference should be required. For the previous reasons, the network is not constrained to choose any particular uplink CoMP scheme. However, the main two alternatives are coordinated scheduling (CS) and joint reception (JR), as shown in Fig. 14.

4.2.1. Coordinated scheduling

In a similar way to downlink CS/CB, uplink coordinated scheduling (CS) decisions for UE precoding and scheduling are decided among a set of receiving sites to minimize interference. In general, the load at the backhaul network in CS is reduced as only CSI and resource allocation



(a) Coordinated scheduling.



(b) Joint reception.

Fig. 14. Uplink CoMP approaches.

information needs to be shared among cooperating receivers. Nevertheless, a combination of the received signals should be performed for a coordinated reception benefiting from multiple receiving points. Fig. 14(a) shows an example where uplink CS-based CoMP is applied to avoid multi-user uplink interference. The backhaul link is utilized to perform Maximum Ratio Combining (MRC) and increase the quality of the user's received signal. Other combining methods could also be applied.

4.2.2. Joint reception

Joint reception (JR) unavoidably requires the signals received at different sites to be jointly processed to produce the final output. This means that the received data needs to be exchanged via the backhaul link to eliminate interference and improve performance. The amount of exchanged information varies depending on the amount of processing carried out at the receiving sites: more processing before exchange usually implies less burden on the backhaul. However, the CoMP gain is also reduced in that case. Fig. 14(b) shows a sample scenario where JR is being applied. Inter-site interference cancellation schemes are applied, whereby two typical examples are minimum mean square error (MMSE) and zero-forcing (ZF).

4.3. Research challenges

There is no doubt that CoMP will continue generating attention, as new network paradigms such as heterogeneous networks keep demanding solutions for cooperation strategies and interference cancellation with geographically distributed antennas [74–76]. Some of the most significant issues that need to be tackled by the academic, industrial and standardization communities are outlined in the following.

4.3.1. Backhaul aspects

As stated above, one important backhaul-related issue in 3GPP is to enable support for CoMP with non-ideal backhaul connection. CoMP so far enables only very close dynamic coordination between multiple network nodes connected via very small latency backhaul links. This imposes serious limitations to operators willing to obtain performance benefits from CoMP operation but whose networks are equipped with non-ideal backhaul. Hence, support for non-ideal backhaul CoMP is needed for future systems. Different approaches can be followed to achieve this goal. Smart and more efficient backhaul cooperation techniques are needed. Some schemes for alleviating backhaul requirements have been proposed in the literature. Serving only subsets of UEs with joint transmission [77] or partitioning a cellular network into small subsystems where these schemes can be applied locally [78] are some examples. Furthermore, very efficient protocols for high-speed backhaul communications as well as specific control and signaling in different scenarios are also needed.

4.3.2. Channel information feedback

The issue of efficient channel information feedback is particularly relevant in CoMP since the performance of cooperation schemes such as JT heavily relies on the accuracy of such information. Nevertheless, the challenge resides in the limited uplink capacity to provide such information. How to design efficient schemes that do not occupy excessive resources while enabling the performance boost of CoMP remains an active area of research.

4.3.3. Reference signal design

The driver behind the need for novel reference signal design for CoMP scenarios is the same one as for new feedback schemes. Accurate CSI information required among cooperating sites should be fed back by the UE in the limited amount of available resources. The density of the reference signal transmission is a significant parameter adapted by 3GPP for these purposes. For multi-cell transmission, requirements can be also set on additional signal processing methods to separate the pilots from different cells. Furthermore, the structure of the pilot signals could be improved for better performance. In the uplink, a joint channel estimation must be performed at each antenna head of the base stations for all the users being served, and this may not be an easy task if the users are situated at different distances from the same base station, i.e., the received signal levels vary significantly. Improved uplink reference signals considering these parameters need to be designed.

4.3.4. Other enhancements for Rel-12

Some important enhancements are being considered to improve performance of CoMP schemes in Rel-12 LTE-Advanced [79]. It was mentioned above that two out of the four proposed CoMP scenarios are currently being investigated within Rel-12, namely the inter-eNB CoMP in homogeneous deployment and inter-cell CoMP in heterogeneous deployment. In addition, Rel-11 does not address the support of a network interface for CoMP involving multiple eNBs with non-ideal backhaul. Other current efforts for CoMP are focusing on support for

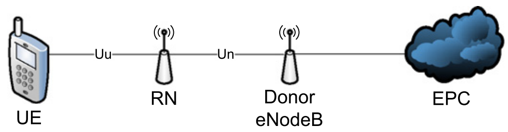


Fig. 15. Relay basic scheme.

multi-carrier CoMP where resources are coordinated and aggregated both in the frequency and time domains. Mobility management for CoMP scenarios is trying to be simplified by coordinating multiple layers of network nodes at different tiers. Moreover, enhanced feedback mechanisms to better support downlink JT as well as enhancing both power control and sounding reference signals in the uplink are active fields of work. All these enhancements pursue the achievement of uniform user experience independently of its location within the network.

5. Relays

Since Rel-10, LTE-Advanced introduced relays to improve the performance of the network. The general idea is shown in Fig. 15. A relay node (RN) receives and transfers information wirelessly to a donor eNB (DeNB) through the new Un interface. Also, the RN receives and transmits data to UEs through the Uu interface, the same one already used by UEs to communicate with eNBs. As such, a RN has both eNB functionality (to serve UEs) and UE functionality (to communicate with the DeNB).

5.1. Benefits

3GPP has identified the following benefits that can be obtained by introducing relays [20]:

- *Provide coverage in new areas*: Relays can be deployed in areas where the usual eNB backhaul solutions (e.g. fiber, microwave links) are not available or are too expensive. In addition, site acquisition should be less of a problem for relays than for eNBs. In Fig. 16, this would correspond to case (a) and (b), where there are areas affected by strong shadowing effects.
- *Temporary network deployments*: Due to their easier deployment, RNs can be deployed and removed significantly faster than eNBs. This makes RNs suitable for temporary deployments.
- *Cell-edge throughput*: By deploying RNs near the cell-edge, the throughput of cell-edge UEs can be improved. This is depicted as case (c) in Fig. 16.
- *Improved data rate*: By deploying RNs in areas with low signal levels, better signal quality can be provided to surrounding UEs, increasing their achievable data rate. This would correspond to case (d), where RNs are deployed to serve underground areas; and case (e) where RNs are used to improve signal strength in indoor environments.
- *Group mobility*: In scenarios where several UEs move in a group (e.g. UEs in a train), a co-located relay can provide improve mobility performance for this group. This scenario is depicted as case (f) in Fig. 16.

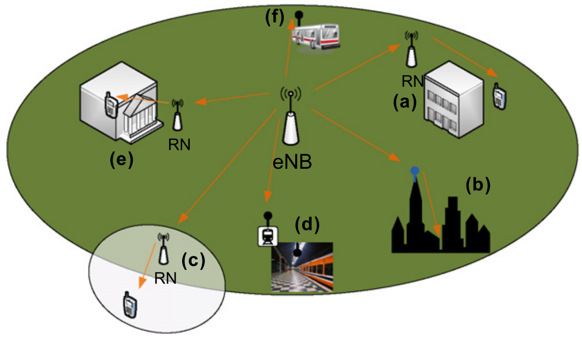


Fig. 16. Relay deployment scenarios.

In addition to the technical benefits provided by relays, operators also see relays as a way to reduce their capital and operational expenses. First, the cost of a relay by itself should be less than the cost of an eNB. This is in general true if we consider that relays should have lower complexity than eNBs. Second, by not needing a wired backhaul, deployments are faster than for eNBs; therefore, less expensive. Third, by deploying the RNs in appropriate locations, the total power needed to serve UEs can be reduced. Therefore, the energy-related operational expenses are lower.

5.2. Relay types

Within the research community relays have been studied for a long time, giving rise to multiple types of relaying schemes. However, introducing relays into LTE-Advanced required a careful selection of the types of relays to be used, as well as how they match the architecture.

The 3GPP architecture to support relays is depicted in Fig. 15. The interface between the RN and the UE is the same one used by the latter to communicate with eNBs, i.e. the Uu interface. Therefore, for an LTE device, any transmission coming from the RN looks like a transmission coming from an eNB. Between the RN and the DeNB, a new interface was defined to support RN-specific functionality, which we will later discuss in Section 5.3.

Within this architecture, 3GPP initially considered two main types of relay nodes [80], namely *type* 1 and *type* 2. A *type* 1 RN is characterized by the following features:

- The backhaul link to the DeNB utilizes the same set of resources as the ones used by the latter to communicate with UEs. This is also called inband relaying.
- It controls cells. Each one has its own Physical ID, synchronization channels, reference symbols, etc. From the UEs' perspective, these cells are different from the ones of the DeNB.
- In single-cell operation, scheduling information, HARQ feedback and control channels are exchanged directly between UEs and RNs.
- For backward compatibility, it appears as a Rel-8 eNB to Rel-8 UEs.
- For further performance enhancements, it may appear differently to LTE-Advanced UEs.

As an inband relay, the backhaul and access links transmissions interfere with each other. To avoid this interference, a *type* 1 RN utilizes different time intervals for each transmission; therefore, it is a half-duplex relay. To avoid the reduced performance associated with half-duplexing, 3GPP introduced two additional types of relays closely related to the *type* 1 RN, namely *type* 1a and *type* 1b, differing from the *type* 1 only in the characteristics of the backhaul. The backhaul of a *type* 1a operates out of band. On the other hand, the backhaul of a *type* 1b operates inband but with enough isolation between the backhaul and access links as to mitigate their mutual interference. Therefore, both *type* 1a and *type* 1b allow full-duplex operation.

A *type* 2 [81] RN is characterized by being an inband relay with the following properties:

- It does not create its own cells; therefore, it does not have a separate Physical Cell ID (PCI).
- It is transparent to LTE UEs; i.e. LTE UEs are not aware of the presence of a RN of *type* 2.
- It can transmit PDSCH, but does not transmit CRS nor PDCCH.

Out of these types of relays, 3GPP selected *type* 1 and *type* 1a to be standardized as part of LTE-Advanced [16]. Therefore, from now on we will focus our discussion on these two types.

In *type* 1a relays there is no interference between access and backhaul links. However, in *type* 1 relays, resources must be shared between both links to avoid interference. 3GPP has followed a general principle to deal with this: uplink transmissions (RN-to-eNB and UE-to-eNB) should be time multiplexed, as well as downlink transmissions (eNB-to-RN and RN-to-UE). To achieve this time multiplexing with minimal impact on the LTE frame structure, it was decided that MBSFN (multicast/broadcast single-frequency network) frames will be utilized in the access link during the time periods where the RN communicates with the DeNB.

Regarding the backhaul link itself, both TDD and FDD configurations are supported. In the FDD case, eNB-to-RN transmissions use the DL frequency bands, while RN-to-eNB transmissions use the UL frequency bands. In the TDD case, eNB-to-RN transmissions use the DL subframes, while RN-to-eNB transmissions use the UL subframes. To facilitate the resource assignment for the backhaul, 3GPP added new control and data channels. The new physical control channel (R-PDCCH) dynamically or semi-persistently assigns resources for the new downlink physical shared channel (R-PDSCH) and uplink physical shared channel (R-PUSCH) backhaul data. The specific subframes that can be utilized for these transmissions have already been specified by 3GPP in [82], as well as the procedures to transmit and receive through these new channels.

5.3. Architecture

Introducing relays in LTE-Advanced involved changes not only in the radio access, but also in the rest of the network. Initially, 3GPP studied two different architectures to support relays at the radio access network and the core network [83], namely *Architecture A* and *Architecture B*. In *Architecture A*, both the user plane and control plane of the S1 interface terminated at the RN. Within this architecture, three alternatives were studied, namely *Alt* 1, *Alt* 2, *Alt* 3. They mainly differed in the amount of additional functionality that was added at the DeNB to support RNs. In *Architecture B* the DeNB terminates the S1 connections towards the EPC, and the RN appears to the rest of the network as another cell under the control of the DeNB. Within this architecture a single alternative was proposed, namely *Alt* 4. Out of all these options, 3GPP selected *Alt* 2 for LTE-Advanced. From now on, we will focus on *Alt* 2.

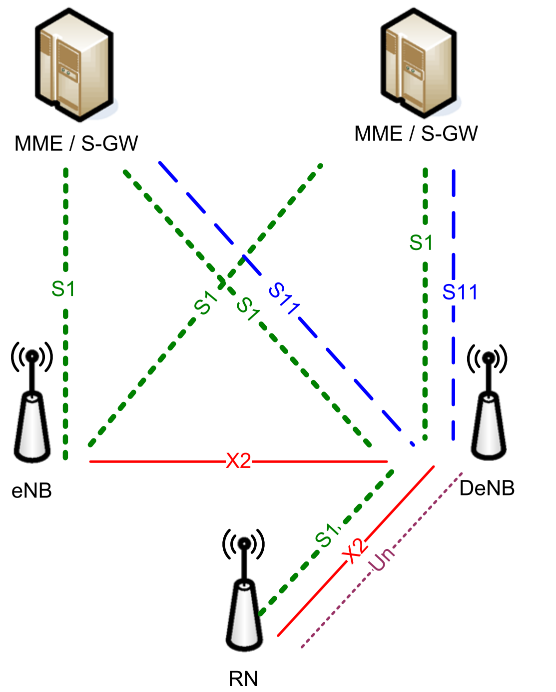


Fig. 17. Relay architecture.

In Fig. 17 the chosen architecture for LTE-Advanced is depicted. Instead of terminating the S1 and X2 interfaces, the DeNB provides S1 and X2 proxy functionality between the RN and the rest of the elements in the core network, leaving to the RN the termination of these interfaces. Therefore, for the S1-MME interface, the DeNB is seen by the RN as an MME. For the X2-interface, the DeNB is seen by the RN as an eNB. Similarly, for the S1-U interface, the DeNB is seen by the RN as an S-GW.

5.3.1. Protocol stack

The general idea behind the framework used for relays is that, besides the DeNB, the rest of the CN and RAN elements consider the DeNB as the eNB to which the UE is attached. As such, any traffic intended for the UE or the "eNB serving the UE" will be sent to the DeNB. The DeNB will map this traffic to the appropriate radio bearer so that it reaches the RN. During these mapping procedures, the DeNB will modify the traffic as to maintain

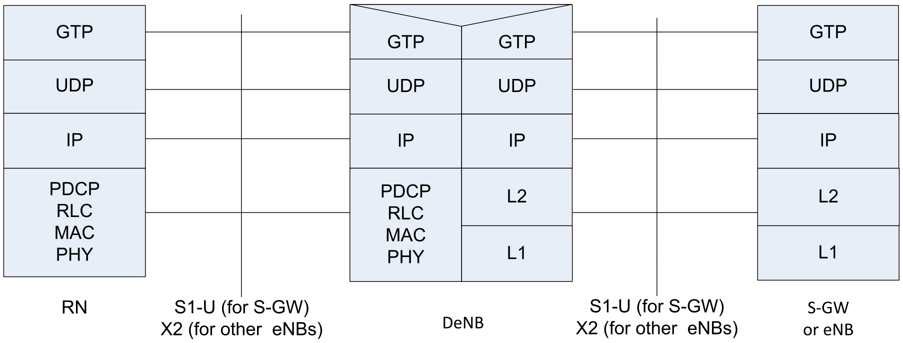


Fig. 18. User plane protocol stack for relay support.

its proxy functionality (i.e. appear as an MME, eNB and S-GW towards the RN). Similar mapping is done by the RN in the uplink.

Fig. 18 depicts the user plane protocol stack supporting relays. For every UE EPS bearer subject to forwarding, there is a unique GTP tunnel between DeNB and the S-GW (for S1-U), and between the DeNB and other eNBs (for X2). The DeNB maps each of these GTP tunnels to a different and unique tunnel that goes from the DeNB to the RN. These tunnels are then mapped to radio bearers over the Un interface.

For the control plane, Fig. 19 depicts the protocol stack supporting relays. Between the RN and the DeNB there is a single S1 interface relation, and a single X2 interface relation. Similarly, there is a unique S1 relation between the DeNB and each MME from the MME pool; and there might be X2 relationships between the DeNB and other eNBs. Regarding forwarding of S1 and X2 messages, all UE-dedicated S1-AP and X2-AP messages are processed and forwarded by the DeNB between the RN and the MME (for S1-AP) and eNBs (for X2-AP). Non-UE dedicated S1-AP and X2-AP messages are terminated at the DeNB. Depending on the specific message, the message can be handled locally between de DeNB and the RN, between the DeNB and the MME (for S1-AP), or between the DeNB and other eNBs (for X2-AP). As with the user plane, control plane packets are mapped to radio bearers over the Un interface.

In order to establish a connection between the RN and the DeNB for the Un interface, the RN utilizes the same protocols and procedures used by UEs to connect with an eNB [84]. This implies that any traffic associated with the UEs (that are served by the RN) is tunneled through data radio bearers that go from the RN to the DeNB. Therefore, 3GPP specifies that this data radio bearers should have functionality to support integrity protection, mainly for protecting signaling traffic associated with the UEs.

The most important relay-related procedures have already been standardized by 3GPP [20]: attachment and detachment procedures for RN, activation and modification of RAB, transfer of neighboring information, mobility to and from RNs, and security [85].

Security mechanisms and procedures are especially important and have also been subject of extensive study [85]. The main concern lies in the fact that any authorized RN has almost unrestricted access to the CN of the operator. If an attacker is able to impersonate a RN, the impact could be extremely high. To prevent this, it was agreed by 3GPP that there should be a one-to-one binding of a RN and a USIM, called USIM–RN [86]. This binding can be achieved through pre-shared key (PSK) or through the use of digital certificates. PSK option does not require a public key infrastructure (PKI) and procedures are easier after pre-establishment. However, the use of certificates allows more flexibility in terms of deployment.

5.4. Research challenges

There is a lot of active work both in industry and in academia to further enhance the capabilities of relays for LTE-Advanced. Within the scope of Rel-12, special attention has been given to the introduction of mobile relays. However, additional topics are also being studied such as site planning for relays, relay selection, backhaul, resource management, intercell-interference, energy and cost, and additional capabilities for relays. We briefly describe them here.

5.4.1. Mobile relay

Mobile Relay (MR) focuses on deploying relays in high-speed transportation systems, such as high-speed trains [87]. These scenarios are characterized by having high-speed, known trajectory, high-penetration loss, and stationary UEs (or moving at pedestrian speeds) with respect to the train. These characteristics create several issues: Severe Doppler effects due to the high speeds, low signal strength due to penetration loss, reduced handover success rates, and increased power consumption.

The idea behind MR is to install the device directly on the train so that it has wireless backhaul connection (through an external antenna) to the eNBs, and provides wireless access to the users within the train (through an internal antenna). UEs continue to receive uninterrupted service from the MR, while the MR takes care of handling all necessary handovers with DeNBs. In WiMAX, this concept has already been studied [88].

In 3GPP, current work is focused on the scenario where the backhaul and access links belong to the same operator (i.e. no multi-operator MR). Nevertheless, it does consider the case where multiple RATs are served internally, independently on the backhaul connection used to communicate with the DeNB.

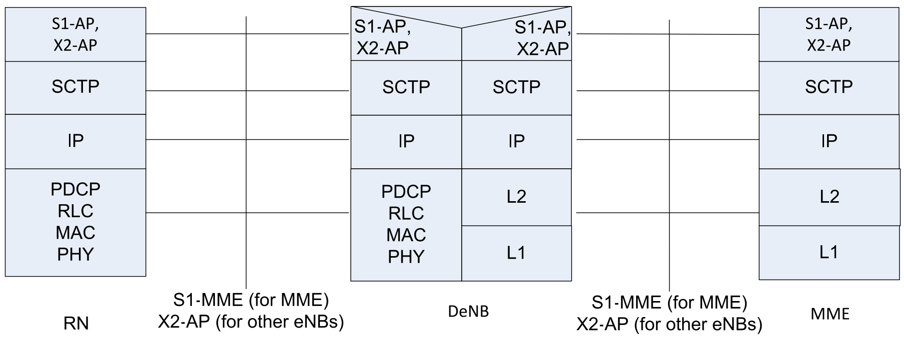


Fig. 19. Control plane protocol stack for relay support.

The functionality associated with MRs is based on the one for fixed relay. The main enhancement in MRs is to support efficient handover procedures between DeNBs. To achieve this goal, three basic architectures [87] have been proposed by 3GPP based on the ones for fixed relays. For MRs, the main difference among the architectures lies on the point where the mobility anchor is located after the handover procedure. However, no final decision has been taken regarding the most appropriate architecture for mobile relays. An important aspect that is taken into account is the compatibility of MRNs with fixed RNs. As such, architectures based on *Alt* 1 and *Alt* 2 (of fixed relays) are being favored [89]. Work is still ongoing, and research is still open to improve the performance and reliability of mobile relays to support seamless group mobility [90].

5.4.2. Site planning

As we mentioned initially, relays can be utilized to achieve multiple objectives in different scenarios. Therefore, an important aspect of relay-enhanced networks is to identify what strategy should be followed to deploy them [91,92]: Where should relays be located? How many relays should be deployed? How much power should they use? Should they be inband or outband? This and many other questions need to be considered according to the specific scenarios. Studies have already been conducted for generic deployments [93], as well as for more specific scenarios such as suburban [94,95] and metropolitan [96] scenarios. However, further studies in additional scenarios are still required.

5.4.3. Relay selection

Once relays are deployed, improvements will not only depend on the site planning, but on how much users rely on the RNs to communicate with the DeNB. In the most simple scenario, a RN will be deployed within the coverage of a DeNB. UEs will be served either through the eNB or through the RN. Among the solutions found in the literature, biasing in cell selection and handover decision [97] has been proposed as a simple and efficient way to encourage users to connect to RNs and improve the load balancing in the network. In scenarios where multiple relays could serve the UEs (e.g. near cell-edge), techniques to select an appropriate set of RNs to serve the users have also been proposed [98], trying to achieve a balance between capacity and fairness.

5.4.4. Backhaul

RNs communicate with the DeNB by utilizing wireless resources. For outband relays, this simply implies that a dedicated frequency is used for the backhaul link. However, for inband relays, the RN uses the same frequency resources to communicate with the DeNB as to communicate with the UEs. This has two implications. The first one is that the RN can only communicate with one of them at a time. Therefore, it needs to achieve a balance between the amount of time it uses to communicate with each of them. The second implication is that the RN will consume DeNB wireless resources (which are also used by the DeNB to serve other users), to serve its own UEs. As such, competing objectives need to be satisfied. On one hand, RNs want to have as much backhaul resources to serve its own UEs; on the other hand, the DeNB has to fairly allocate [99–101] and share [102,103] resources between the RNs and other UEs under its control.

5.4.5. Interference management

The transmissions of a RN can cause interference to nearby UEs that are being served by the DeNB, and vice versa. Therefore, efficient mechanism to coordinate their transmission [104–106] are required to mitigate this type of interference. These effects are potentially larger when multihop relays are deployed [107,108] (currently not under study by 3GPP), where each RN causes interference not only to the eNB but to other RNs. These effects are also present when nearby RNs are served by different eNBs; therefore, the backhaul transmissions from one may interfere with the access transmissions of the other [109].

5.4.6. Other aspects

On top of the research challenges associated with the basic relaying schemes, there are several other aspects that are under study. For example, carrier aggregation [110], enhanced MIMO, and cooperative techniques can be added to relays to further enhance their performance not just at the access link, but also at the backhaul. Similarly, techniques from self-organizing networks [111] are also being introduced in relays to achieve energy savings, improve coverage and capacity [112], as well as self-reconfiguration [113]. All of these options open the door for novel approaches to enhance relaying in LTE-Advanced.

6. Heterogeneous networks

The need for increased system capacity in the order of several orders of magnitude has led to the proliferation of smaller low-powered cellular layers underlaid on the existing macrocell layer. These low-powered small cells include picocells, femtocells, metrocells, etc. While picocells and metrocells are utilized for outdoor deployments, femtocells are proposed for deployment in indoor environments such as residential or enterprise buildings. Such a type of network that includes different cellular layers, each with their unique characteristics such as transmission power, backhaul technology is termed as a heterogeneous network (*HetNet*).

HetNet is not a new concept of LTE-Advanced since it is only a deployment strategy. However, the heterogeneous nature of the different cellular layers raises several important challenges that impact the actual capacity that can be achieved by the overall system. The key challenges due to the introduction of HetNets are listed below:

- Cell association.
- Inter-cell interference management.
- Mobility issues.

We discuss each of these challenges and provide the state-of-the-art work carried out in the literature as well as in the standard bodies to overcome these challenges.

6.1. Cell association

In case of homogeneous cell deployment, the problem of assigning users to a base station is straightforward. This is typically performed by assigning users to the base station with the largest downlink signal strength or signal to interference noise ratio (SINR) at the user's receiver [114,115]. However, in a heterogeneous network, the assignment of users to a particular base station is a non-trivial problem and depends on several system aspects. If the objective is to increase the uplink data rate, then it is better suited for the user to attach to the base station with the minimum path loss. Similarly, the cell association also depends on the load on the system. With the objective of increasing the overall network capacity, it is necessary to assign users to the small cell base station when the system load is high. On the other hand at low loads, the traditional approach of associating to the base stations with the best downlink signal strength or SINR helps achieve downlink higher data rates. Several other factors also play a key role in determining the cell association for users such as the backhaul capacity available at the base station.

These factors motivate the use of cell biasing as a function of the network parameters of load, uplink data rate, backhaul capacity, etc. [116]. This is still under study in 3GPP Rel-12. The impact of utilizing a cell association policy based on maximum biased-received-power (BRP) on the user throughput and outage probability in a heterogeneous network is studied in [117]. The outage probability is shown to increase due to the cell bias while an increase in the user throughput is observed since the user is able to utilize more radio resources as a result of the bias. The cell selection problem is posed as a network-wide

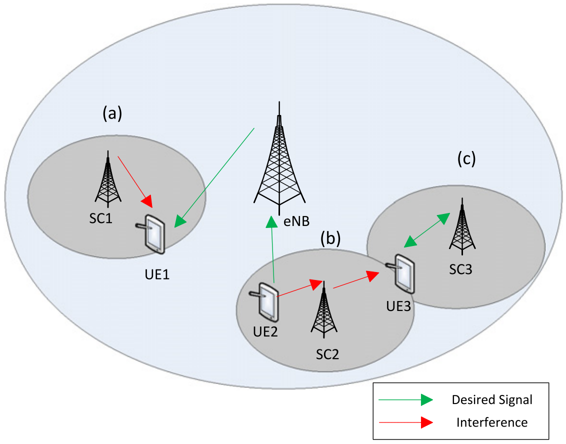


Fig. 20. Inter-cell interference scenarios in a HetNet.

proportional fairness optimization problem where the best cell is chosen by jointly considering the long-term channel condition and load balance in a macro–pico heterogeneous network in [118]. This approach is also shown to improve the normalized user throughput at the expense of the average SINR experienced by the user.

In spite of the gains in terms of user throughput, many of the complex cell association policies for HetNets result in greater outage probability and minimized average SINR for the users. This is expected since any cell association policy other than the one based on the downlink received signal strength will result in increased interference from the base station with the highest signal strength. This phenomenon called as inter-cell interference (ICI) is a major factor affecting the performance of heterogeneous networks. This is discussed in the Section 6.2.

6.2. Inter-cell interference cancellation

Fig. 20 illustrates the different scenarios of inter-cell interference. In scenario (a), the macrocell user (UE1) is unable to connect to SC1 as it is not a part of the closed subscriber group (CSG) of SC1. Therefore, it is attached to the eNB suffering strong downlink interference from SC1. Similarly, in scenario (b), the macrocell user (UE2) is attached to the eNB and results in interference at the receiver of SC2. In scenario (c), the UE3 is attached to SC3 while receiving strong interfering signal from SC2 on the downlink.

In order to overcome this ICI, several inter-cell interference cancellation (ICIC) methods have been proposed as part of LTE-Advanced Rel-8 and Rel-9. One of the common approaches is to use different carrier frequencies for different cell layers [116]. This is an effective means to mitigate the ICI but at the expense of utilizing more frequency resources. In the more common case when the cell layers utilize the same carrier frequency, power control schemes are recommended to reduce the transmission power of at least one of the cell layers to reduce the interference on other layers. Other schemes proposed for ICIC based on

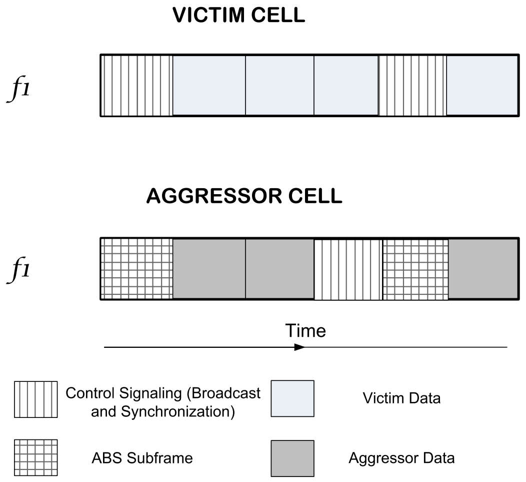


Fig. 21. Frame configuration for time domain-based ICIC scheme.

Rel-8 and Rel-9 capabilities include adaptive fractional frequency reuse, spatial antenna techniques including MIMO and SDMA, and adaptive beamforming [119].

A majority of the solutions proposed in the Rel-8 and Rel-9 do not address the interference at the control signaling channel (PDCCH) since it is transmitted over the entire bandwidth in order to achieve diversity gain in the frequency domain. With the enhancements in the physical layer enabled for Rel-10 and Rel-11 such as Carrier aggregation, more sophisticated methods for performing ICIC are utilized for Rel-11. These enhanced ICIC solutions are basically categorized as (i) time domain-based, and (ii) frequency domain-based ICIC schemes.

6.2.1. Time domain-based ICIC

The time domain-based ICIC schemes basically rely on reducing the transmission activity on certain subframes by each of the cell layers to minimize interference to the victim layers. These subframes are indicated as Almost Blank Subframes (ABS). The knowledge of the subframe timing is assumed to be available at the different base stations. The ABS are achieved by either utilizing the MBSFN (Multicast/Broadcast over single-frequency network) subframes or by deliberately not scheduling (or using power control) over other subframes. Fig. 21 illustrates the frame configuration of a victim cell and an aggressor cell when the above time domain-based ICIC scheme is applied. The aggressor uses the ABS subframe during the transmission of control signaling (broadcast and synchronization signals) by the victim cell. In the ABS subframe, the aggressor applies power control and transmits only common reference signals (CRS), critical control channels or broadcast and paging information for legacy support. Thus the victim's control signals are protected during the ABS subframe duration.

The transmission by the aggressor cell in the ABS subframe, particularly the CRS, cannot be avoided in any of the ABS subframes whereas the rest of the control signaling or broadcast signals can be suppressed for some of the ABS subframes. This transmission of CRS will result in interference experienced at the victim cell. Many of the existing time domain-based ICIC schemes assume that perfect cancellation of CRS interference from the received signal at the victim cells. The performance of time domain ICIC schemes when imperfect CRS interference cancellation is analyzed in [120]. A non-ideal RSRP measurement based CRS cancellation scheme is proposed accompanied by the time domain-based ICIC. The results show that imperfect CRS cancellation can still result in user throughput that is comparable to the ideal cancellation case.

The time domain schemes require that the different cell layers exchange the ABS subframe information. This can be enabled using the X2 interface existing between the base stations. In addition, it is also highly desired that time synchronization is achieved, at least at the subframe boundaries between the cell layers.

6.2.2. Frequency domain-based ICIC

Among the frequency domain-based ICIC schemes, CA with cross-carrier scheduling is widely regarded as a prominent method for performing ICIC. The control channel interference on the downlink can be mitigated by partitioning the component carriers in the cell layers into two different sets, one for data and control while the other mainly for data and perhaps also control signaling with power control applied. Fig. 22 illustrates such a CA-based ICIC scheme. Carrier aggregation is used by both macrocell and small cell layers where both layers enable data communication over $f1$ and $f2$. The macrocell includes the control information only on $f1$ for data transmitted on both $f1$ and $f2$. In a similar manner, the small cell includes the control information only on $f2$. Using this simple mechanism, the control signaling for the different layers are separated. In such a case, the macrocell adopts the carrier $f1$ as the primary carrier while applying power control schemes on carrier $f2$ to minimize interference to the small cell tier. Similarly, the small cell utilizes $f2$ as the primary carrier and applies power control schemes on $f1$. The above scheme, however, has the constraint that the different layers need to be time synchronized.

In Rel-11, further support for performing frequency domain-based ICIC was provisioned in the form of enhanced PDCCH (EPDCCH) [121]. One of the key design aspects of the EPDCCH is that it enables each UE to be configured with up to two EPDCCH sets. Each of these sets consist of 2, 4 or 8 physical resource block (PRB) pairs. The locations of these PRBs in the LTE-A frame indicating the EPDCCH are provided with signaling messages with minimum overhead [122]. Such a flexible allocation of EPDCCH control signaling within the frame enables coordination between different tiers to avoid overlapping EPDCCH and hence the control signal interference.

6.3. Mobility management

Seamless and robust mobility is a bigger challenge in HetNets compared to a homogeneous macrocell deployments. A detailed study of mobility performance for HetNets show that the handover failure rates and ping–pong

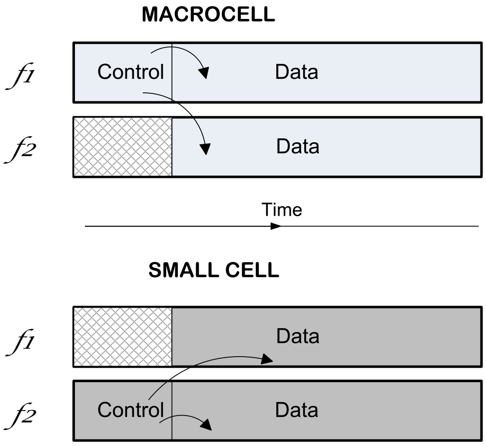


Fig. 22. Frame configuration for CA-based ICIC scheme.

effects are far greater in HetNet environments compared to homogeneous deployments [123]. The poor handover performance of users in HetNets is owing to several factors. One of the major factors is the need to account for the unique characteristics of HetNets when performing user mobility state estimation in addition to accounting for parameters such as the number of cross-overs. Another factor is the need to perform efficient small cell discovery over different carriers at the base stations. This is essential to enable traffic offload from macro to small cells as well as to reduce the battery consumption for the users and the overall system. Furthermore, since the handover failures can be larger in the HetNet case, more robust recovery procedures are required to mitigate the impact of a failed handover.

Due to these factors, the following objectives have been laid out in Rel-11 for mobility enhancements [123].

- Identification and evaluation of advanced small cell discovery/identification schemes.
- Identification of advanced recovery mechanisms from Radio-Link-Failures (RLF) to improve mobility robustness.
- Identification and evaluation of HetNet mobility performance under Rel-10 enhanced ICIC schemes.
- Definition of mobility enhancements for CSG femtocells with carrier aggregation.

6.3.1. Improved small cell discovery/identification

A typical HetNet scenario comprising different small cell layers occurs where the macrocell and the small cell layers are on different carrier frequencies. In this case, the user performs inter-frequency small cell measurements for a carrier that is expected to be used for the small cell layers. An optimal small cell discovery mechanism is defined as one which (i) minimizes the user's power consumption while performing inter-frequency measurements, (ii) minimizes interruptions on the serving cell due to measurements, (iii) does not degrade the inter-frequency mobility performance, (iv) does not degrade the mobility performance of legacy users. Performance evaluations of a baseline small cell discovery scheme showed that continuously performing inter-frequency measurements only increases the battery consumption without significantly increasing the offloading gains.

Consequently, enhancements to the small cell discovery mechanisms were proposed. Many of these solutions can be broadly classified into the following types.

- *Relaxed measurement configuration*: In this case, the performance requirements are relaxed for the inter-frequency measurements with the objective of saving the battery life [124,125].
- *User mobility state estimation-based measurement*: The high-speed users are exempt from performing inter-frequency measurements [126,127].
- *Small cell signal based measurement control*: Whenever the users detect a small cell signal with sufficient strength, they do not perform further measurements [128].
- *Small cell discovery signal based measurement*: Macrocell utilizes the intra-frequency discovery signals of small cells from user measurements to locate the small cells [129].
- *Proximity detection based small cell discovery*: Proximity detection is performed by user or macrocell (or even assisted by network) using techniques including location information from user measurements, etc. [130–134].

6.3.2. Enhanced mobility state estimation

The enhanced user mobility state estimation for Rel-11 and Rel-12 are categorized into user and network based mechanisms with the objective of enhancing mobility performance. This is still in the study item phase and proposed solutions are expected to take into account the unique characteristics of small cells, e.g. the cell sizes and other parameters.

6.3.3. Mobility performance with enhanced ICIC

Performance studies indicate that the enhanced ICIC schemes with ideal ABS pattern coordination between the cell layers accompanied by a cell range expansion (CRE) bias of around 6 dB can improve the mobility performance in HetNets. At the same time, when a non-ideal ABS pattern coordination is employed, the interference from macrocell can be high due to a large CRE bias resulting in mobility performance degradation.

6.4. Research challenges

With the growing need for a tremendous network capacity, HetNets operating on the principle of spatial frequency reuse will play a key role in LTE-Advanced and B4G systems. In spite of the numerous solutions proposed for the key areas of interference management, cell association and mobility management, HetNets still face several issues that need to be tackled by the academic, industrial and standardization communities. Some of the most significant research challenges are discussed in this section.

6.4.1. Cell association and inter-cell interference cancellation

Both the frequency domain-based (CA-based) and time domain-based ICIC schemes are faced with the constraint of requiring time synchronization between the overlapping base stations of different cell layers. Such a tight synchronization calls for high-speed X2 links to be present between the eNBs. For the case of femtocells which utilize the internet as the backhaul technology, achieving the high-speed X2 links with macrocells can be challenging. Therefore, new ICIC schemes proposed must also aim to minimize the amount of signaling exchange between overlapping eNBs. For instance, cooperative beamforming approaches utilized to achieve ICIC require channel state information (CSI) to be exchanged between eNBs. The authors in [135] study the impact of delayed limited feedback information exchanged between eNBs when cooperative beamforming based ICIC scheme is utilized. An upper bound on the mean loss in sum-rate due to delayed limited feedback is derived and closed-form expressions for allocating bandwidth on the backhaul to exchange the desired and the interferer channel feedback are derived. However, a similar analysis for the time domain and frequency domain based ICIC schemes is still an open issue.

In the case of cell association policies, in spite of the gains in terms of user throughput, a number of proposed solutions result in greater outage probability and minimized average SINR for the users since they cause additional ICI. That is why cell association and ICI are coupled problems and should be solved jointly. We envision a joint cell association and ICI cancellation solution that can take interference into account before triggering a user's association decision. This can be accomplished by exploiting the inherent spatio-temporal coverage cross-tier correlation due to ICI as characterized in [136]. By jointly solving the above mentioned coupled problems, we envision that a reduction of the ICI can be obtained while still increasing the network capacity.

6.4.2. Mobility management

Owing to the variable backhaul link quality and increased core network congestion due to frequent handovers, new handover schemes addressing the above challenges need to be proposed. Our initial work on a local anchor-based handover scheme for coordinated small cells showed significant gains in the minimization of handover signaling and session interruption time [137]. However, the case of handovers between macrocells and small cells is still an open issue and new solutions to minimize the handover signaling as well as mitigating the impact of a failed handover will be proposed and evaluated.

Similar to handover management, location management in HetNets including location update and paging is impacted by several factors. First, reduced coverage area of the small cells means frequent location updates are required. This incurs high costs in terms of signaling and network load. Second, backhaul capacity limitations set tighter restrictions on the signaling load. Therefore, intelligently routing the paging messages to a select few small cells for an incoming session to a user must be investigated. In our recent work, we have proposed two schemes for location update and paging [138]. These schemes reduce paging through the use of a local anchor to keep track of users' location inside the TAL.

In addition to the above challenges, other open issues need to be resolved. A mobility management framework taking into account the unique features of HetNets still needs to be developed. Further, small cell discovery mechanisms need to be proposed to enable traffic offload from macro to small cells, and to reduce the battery consumption for the users and the overall system. Very few works address this problem of small cell selection or discovery in the current literature [139,140].

7. Self-organizing networks

The concept of self-organizing networks (SON) was initially introduced in Rel-8 as a framework that supports functions to automate several processes in the network. This automation has a two-fold benefit for operators. First, it reduces operational expenses by reducing the manual involvement in the execution of the processes. Second, it allows the network to react faster to events that trigger the execution of processes. Within the framework of SON [141], the main concepts and requirements are:

- *Self-configuration of eNBs*: Automation of the establishment of new eNBs into the network. This includes the establishment of IP connectivity with the element manager; downloading of software and radio and transport configuration data; and setup of X2 and S1 interfaces.
- *Self-healing of eNBs*: Automation of the detection, mitigation, and solution of problems in the network. This avoids impact on the users and reduces maintenance costs.
- *Automatic neighbor relation (ANR) list management*: Automation of the discovery of neighbor relations.
- *Self-optimization*: Automation of the reconfiguration of system, network and radio related parameters in order to achieve one or more of the following objectives:
  – Coverage and Capacity Optimization.
  – Load balancing.
  – Handover Parameter Optimization.
  – Interference control.
  – RACH Optimization.

The architecture support SON has three variants. In the first one, referred to as centralized SON, the SON algorithms are executed centrally at the Network Management (NM) level or centrally at the Element Management (EM) level. In the second one, referred to as distributed SON, the SON algorithms are distributedly executed at the Network Element level (NE). The third one, referred to as hybrid SON involves the execution in two or more of the following: NM, NE, or NM.

From the OAM system point of view, initial work was focused on self-configuration [142–144], ANR list management [145], and self-optimization management. Initial studies were also performed for self-healing [146], and for SON related OAM interfaces for H(e)NBs [147]. From the RAN point of view, priority was initially given to functions related to early phases of network roll-out and operation: Coverage and capacity optimization, mobility load balancing optimization, mobility robustness optimization, and RACH optimization [148].

In LTE-Advanced, SON functions developed for LTE were enhanced, and study on new functions was initiated. From the OAM system point of view, work on Self-Optimization management was continued. In addition, work was focused on self-healing management, intra-RAT and inter-RAT energy savings, SON coordination management. From the RAN point of view, enhancements were introduced for different scenarios of coverage and capacity optimization, mobility robustness optimization, and mobility load balancing.

7.1. Research challenges

Current work in Rel-12 is focused on enhancing the network management for centralized coverage and capacity optimization, and Multi-vendor Plug and Play eNB connection to the network, both from OAM perspective. From the RAN point of view, there is ongoing work to support LTE-HRPD inter-RAT SON.

By looking at how SON has been developed within 3GPP, we can clearly identify a trend. SON functions have been developed as individual objectives that operators want to achieve in their networks. In most cases, operators are interested in achieving multiple of these objectives at the same time. However, most of these objectives enter into conflict with each other. For example, energy savings can be achieved by turning off eNBs; however, this reduces the overall capacity of the network. As such, coordination among different SON functions has gained a lot of attention and research efforts.

Another important factor is that operators nowadays already have legacy networks (GSM, UMTS, HSPA, WiMAX) on top of which LTE and LTE-Advanced will be deployed. If SON functionality is restricted to LTE-Advanced, limited benefits can be achieved due to rigidity of their existing networks. Therefore, SON functions are being extended in order to enhance not just LTE-Advanced, but also legacy networks. This requires not only individual SON functions in each network, but efficient inter-RAT SON mechanisms that are able to jointly optimize all the resources managed by operators.

An additional aspect is that SON functions may be executed in different time and locations, due to different spatial and temporal objectives of operators. Therefore, it is extremely important to design SON mechanisms and algorithms that can be implemented seamlessly in the network, without causing disruptions on the service offered to users, nor disrupting the stability of the network.

It is important to highlight that SON algorithms will not be specified by 3GPP, but the framework to support those algorithms. Therefore, SON algorithms are left as implementation specific. In addition, SON objectives are not achieved by just a single technology, but by jointly exploiting all the features provided by the technologies that we have described: Carrier Aggregation, MIMO, CoMP, Relays. As such, the research community focused its initial efforts in applying SON for spectrum management and inter-cell interference mitigation [149–151], and the improvement of the mobility robustness across multiple RATs [152,153]. Most recently, efforts have been directed to SON for energy savings [154,155], as well as its extension to relay deployments [112]. However, the most challenging aspect is to achieve a stable interaction among multiple SON objectives [156].

8. Machine-type communication

In addition to the proliferation of human users, there is a fast-growing tendency of automation (Internet of Things, Smartgrids), where machines are able to interact between themselves as part of a large network. This type of communication is referred to as Machine-type communications (MTC) or Machine-to-machine communication (M2M). The number of MTC devices is predicted to grow at a compound annual growth rate of over 25% [157,158]. Several technologies have been proposed to realize MTC communications among which WiFi, Zigbee, Bluetooth are more prominent. However, MTC solutions utilizing mobile cellular networks can offer several benefits. They are easier to install and can provide reliable data delivery to remote MTC servers. Moreover, mobile networks are crucial for supporting a large range of MTC devices including the highly mobile ones. One of the first studies on MTC for 3GPP systems was presented in [159].

The M2M communication scenarios according to 3GPP can be classified as follows: (1) Communication between MTC Devices and one or more MTC Servers, (2) Communication between different MTC Devices. This is illustrated in Fig. 23. The MTC devices can exist in groups and each of the devices have similar capabilities to LTE-Advanced users to communicate with the LTE-Advanced base stations. The core network of LTE-Advanced is connected to an MTC server that is controlled by MTC users (such as an organization or individual). In the first case, an MTC user can control several MTC devices (such as a set of sensors performing measurements) through an MTC server and hence the communication takes place between the MTC devices and the MTC servers. In the second case, different MTC devices either under the same/different operator can communicate between themselves without an intermediate MTC server.

The support of MTC for 3GPP systems raises quite a few challenges, both at the air interface level and the network level. Particularly, for the case of LTE-A, there would be several millions of MTC devices, each one requiring only a fraction of the bandwidth for transmitting data. This could result in the wastage of a significant amount of resource due to the heavy signaling and data overheads involved for communication in LTE-A. In addition, this could also result in a reduction of the resources available for other devices that actually require high data rate communications. A 3GPP work item was organized to identify and propose novel solutions for supporting MTC devices at the LTE-Advanced air interface [160]. Similarly, the service requirements for MTC was laid out in [161]. In this paper, we focus on the efforts that have been initiated for providing air interface support for MTC devices.

8.1. Air interface support for MTC devices

Air interface is considered the first priority to enable MTC communications to fully support the service requirement and optimization categories defined in [161]. The following are some of the key focus areas of Rel-11:

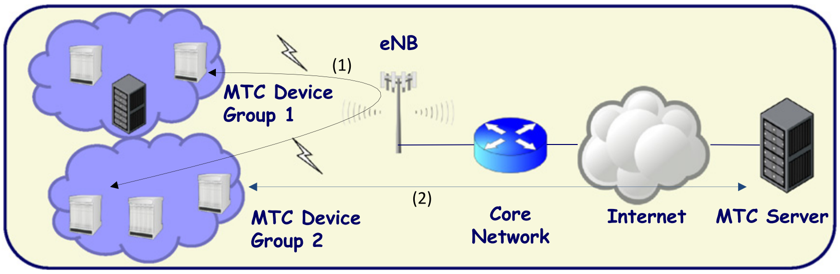


Fig. 23. MTC architecture for LTE-Advanced with communication scenarios.

- RAN Overload Control.
- Efficient Radio Resource Allocation.
- Energy Savings.
- Traffic Models for MTC Applications.

8.1.1. RAN Overload Control

The large number of MTC devices expected to access the radio spectrum in addition to the regular users can result in overload of the radio access network as well as the core network as defined in Rel-10 [161]. This may result in extended packet transmission delay, packet loss or perhaps even service unavailability. Therefore, mechanisms for guaranteed data delivery as well as for providing performance guarantees under huge MTC loads have been investigated. In particular, the load on the Random Access Channel (RACH) that is utilized by the users and MTC devices during cell attach and handover procedures need to be handled without resulting in congestion. The following solutions were proposed under RAN Overload Control.

- *Access Class Barring (ACB) based scheme*: This scheme is utilized for performing access control for the RACH channel for different user types and it is envisioned that this scheme can be applied to perform access control for different MTC device types as well. ACB consists of 16 access classes (AC) where the classes 0–9 represent regular users, class 10 represents an emergency call and classes 11–15 represent high-priority services. The user or MTC device performs random access as follows. First, the access probability p and the AC barring time are broadcasted by the eNB corresponding to the AC 0–9. Following this, the user randomly derives a local value q and performs random access only when $q \leq p$. Otherwise, the device waits for the AC barring time before attempting random access once again. Therefore, by setting appropriate values for p for certain MTC types, RACH collision can be minimized. At the same time, a very low value of p can result in long delays for the devices to access the channel. Therefore, assigning different p values for different MTC types is necessary.
- *Separate RACH resources for MTC devices*: It is also possible to utilize separate RACH resources for MTC devices and the regular users. This is achieved by either splitting the preambles into MTC group(s) and regular UE group(s) or by dividing physical RACH (PRACH) occasions in time or frequency between MTCs and regular UEs.
- *Dynamic allocation of RACH resources*: This scheme proposes that the RACH resources are allocated dynamically based on the load on the access network.
- *Slotted access scheme*: In this scheme, the MTC devices are assigned access cycles/slots. Each MTC device only accesses during its designated access slot. A simple implementation of this access slot would be the paging frame for the MTC device.
- *Pull-based scheme*: According to this scheme, the MTC device accesses the radio network based on a notification from the MTC Server. The notification could be a paging message from the operator's core network. If paging is used, the network could page an individual MTC device or a group of devices.

8.1.2. Efficient radio resource allocation

The variety of MTC applications that need to be supported also require different QoS performance. In addition, MTC devices also have some specific characteristics such as the transmission of a very small amount of data but with a massive number of devices. Furthermore, the hardware limitations of MTC devices also implies that the computational complexity of the radio resource allocation algorithms cannot be very large. These factors indicate the need for new algorithms for radio resource management. Group-based radio resource allocation schemes have been proposed in [162] where MTC devices are grouped into clusters based on their QoS characteristics and requirements. Following this, radio resource allocation is performed on a cluster basis thus reducing the complexity.

8.1.3. Energy savings

MTC devices have strict battery energy constraints, therefore, the LTE-A systems should be capable of enabling energy saving mechanisms for these devices. Currently, LTE-A systems adopt power saving schemes such as discontinuous reception (DRX) and IDLE mode operation in order to enhance the device power saving [163]. However, these approaches are designed with human-to-human communications in mind and are not optimized for the unique features of M2M communication. One of the major characteristics of M2M communication is that these devices generate traffic less frequently, i.e., the traffic inter-arrival time can be extremely large in the order of minutes to hours. Another key characteristic is that a large number of MTC devices can be classified as low-mobile devices and hence do not need to perform mobility-related signaling frequently.

Taking these unique features into account, energy saving mechanisms have been proposed for LTE-A and beyond systems. A simple and effective strategy is to increase the amount of time a M2M user goes to low-power operating state. In LTE-A, the low-power states indicate the IDLE state and the DRX mode within the IDLE state. Power savings can be achieved by increasing the DRX duration beyond the maximum allowed in current LTE-A systems. DRX duration represents the number of radio frames in the paging cycle for which the user can be in sleep mode during an IDLE state before receiving the next paging message from the network. Longer the DRX duration, longer the user can wait before waking up. Hence, energy savings can be achieved by setting a longer DRX duration. Such an approach is proposed in [164]. Similarly, the use of longer paging cycle and its impact on latency degradation is studied in [165]. In addition, a mechanism to reduce the transition time from an RRC CONNECTED state to an IDLE state is proposed in [165] to minimize the power consumption.

8.2. Research challenges

There are several research challenges that still need to be addressed for scalable support of MTC devices for Rel-12 and beyond.

8.2.1. RAN Overload Control

The existing pull-based overload control approaches cannot efficiently resolve RAN overload as the number of MTC devices in the network becomes very large. Furthermore, in most of the proposed overload control approaches, there is an inherent tradeoff between the frequency of collision and delay. In other words, MTC devices need to wait for longer duration to access the RACH channel if they intend to minimize the probability of collision. The authors in [166] highlight that enabling cooperation between eNBs can help overcome extremely long delays to access the RACH channel. To this end, the authors propose a cooperative Access Class Barring (ACB) approach where the eNBs cooperate to compute the optimum access probability vector $\vec{p}$ containing the appropriate access probabilities used for each of the cooperating eNBs. Such an optimum global $\vec{p}$ is shown to achieve significantly improved delay performance compared to existing schemes while minimizing the access collision rate.

Another major issue arises when the user session is mobile originated. In this case, the eNB is unaware of the UE's intention of sending data. As a result, the eNB needs to page the UE in regular intervals but the UE may not have any data to send. This results in long average waiting time for paging especially given the massive number of MTC devices that need to be paged.

Proper system tuning in order to achieve a balance in the collision rate *vs* delay tradeoff is necessary. To this end, a detailed performance evaluation of the RACH overload control methods need to be conducted to determine the suitability of the proposed solutions.

8.2.2. Mobility signaling reduction for low-mobile devices

A significant amount of signaling reduction and energy savings can be achieved by limiting the redundant signaling such as the mobility-related signaling for low-mobile users. Even during IDLE state, there is a significant amount of signaling performed by LTE-A devices for location tracking, neighbor measurements and so on. Neighbor cell measurements are performed periodically by LTE-A devices to determine the strongest cell to attach to. The increase in the cell density will cause the M2M device to perform a number of measurements which increases the battery life. Such an approach is wasteful, especially, for low-mobile and low data-rate M2M devices for which the need for cell re-selection is less likely. Similarly, location area update or tracking area update is performed periodically by LTE-A devices to update their location with the network. This also results in very-high redundancy for low-mobile M2M devices. New solutions are required for future releases of LTE-A that can minimize this overall mobility related signaling for low-mobile devices while maintaining compatibility with legacy and human devices. Initial efforts have seen approaches for introducing intermediate SLEEP states that can retain the key signaling tasks to maintain M2M device context at the network while removing redundant operations. One such approach is provided in [167] where two new states are introduced with varying levels of signaling load in addition to the IDLE and CONNECTED states to minimize mobility related signaling.

8.2.3. Link adaptation techniques

Sophisticated link adaptation techniques are proposed for LTE-Advanced systems in order to achieve very high peak data rates. These techniques include adaptive modulation and coding, hybrid ARQ, advanced MIMO, CoMP among others. However, it is very likely that several of the MTC devices will be used in low data rate applications such as reporting sensor measurements. For such applications, the overhead and the energy consumption associated with these sophisticated link adaptation approaches can be large and unnecessary. Instead, much simpler yet more reliable transmission schemes need to be adopted for the low data rate MTC devices.

8.2.4. MTC traffic modeling and characterization

Traffic modeling of the proposed MTC application types need to be performed. Only two traffic models have been studied as of Rel-11 where the arrival rates are modeled as time-limited Beta distribution or Uniform distribution [168]. More accurate traffic models can help in accurately computing the performance gains of several RAN enhancement methods for M2M communications.

9. Device-to-device communication

The paradigm of device-to-device communication (D2D), also known as proximity services (ProSe), is another key enabling technology for LTE-Advanced that has been included in Rel-12 for consideration [169,170]. It refers to the ability of cellular devices to communicate directly among each other without the need to access the network infrastructure. Although currently existing technologies such as Bluetooth or WiFi-direct already allow this functionality, they do not allow the operator to control the communication process and guarantee certain user experience. Furthermore, they are not designed to be integrated within the cellular network, while D2D should minimize the impact on the cellular network and enable users to transparently switch between D2D modes and regular cellular communication.

The applications driving this new communication technology are already numerous as of today, but it is likely that many others are yet to be found in the near future. We describe some of them in this paragraph: First, operators are becoming increasingly interested in context-aware applications where a plurality of services can be provided depending on the user's location. To make the context-aware functionality of the network as flexible as possible, optimized direct communication of the user with other devices should be enabled. Second, the support for public safety networks in cases where network infrastructure is not available is a major current concern that 3GPP is addressing with the introduction of D2D. Third, the rapid emergence of M2M applications and the Internet of Things makes a use case for D2D, since it would be desirable to control electronic consumer devices directly from the cell phone of the owner without the need to overload the network. Fourth, D2D can be seen as an additional method to offload traffic from the congested macrocell networks, so operators are currently exploring the option of exploiting D2D functionalities with that purpose.

There are some aspects of D2D that deserve special attention. First, some similarities are shared between D2D and mobile ad hoc networks, which have been extensively studied in the literature. However, a significant difference is that D2D can rely on limited assistance from the network for control functions or, if necessary, to carry out the communication. How to select the best network mode for operation is still a research challenges, as it is pointed out later in this section. Second, a question of vital importance is what spectrum to use, and how to distribute it. The direct link between UEs can be established over a variety of air interfaces including both licensed and unlicensed spectrum. The classical advantages and disadvantages of both options apply here, with the particularity that this type of communication must take place in a integrated fashion with the rest of the network. Therefore, network operators may also benefit from less congestion and interference in their network when using D2D in the unlicensed band.

9.1. Research challenges

Device-to-device communication (D2D) is an emerging technology with high disruptive potential in the network. However, a large number of challenges still remain unsolved and constitute an exciting field of research for the next few years. In this section, we summarize the most significant ones.

9.1.1. Mode selection

The D2D devices can operate in different modes, depending on the available resources of the network. First, if no resources are available, D2D nodes may be urged to remain in *silent mode*. Second, D2D users may be able to transmit either in *reuse mode*, where the spectrum is shared with cellular users, or *dedicated mode*, with dedicated resources for the direct communication. Third, D2D users may need to operate in *cellular mode*, with traffic being locally relayed through one or several eNBs. Although the goal of D2D is to establish a direct communication among devices with little to no interaction from the network, this last mode of operation may be often required for control packets, data packets, or both. Also, many relaying options should be investigated, such as the possibility of using access points operating in the unlicensed bands. Establishing mechanisms to determine the level of assistance that is required from the network is also an open challenge.

9.1.2. Group communications

A second important challenge is the ability for D2D to establish efficient group communications, where the number of devices within the group may vary depending on the application and the communication approach (unicast, multicast, broadcast). In this respect, one-to-one, one-to-many, and many-to-many approaches should be enabled. Furthermore, a thorough study of how the throughput scales with the number of devices in the group has to be carried out. In the literature, clustering and user pairing have been both assumed for group communications [171,172], but the fundamental problem of how to set up these communication groups is still open: Which metric should be employed to decide the communication pairs or groups? Ideas in the direction of context-aware or interest-aware selection mechanisms could have enormous potential for this task.

9.1.3. Interference and resource management

An important constraint to keep in mind is the fact that the paradigm of D2D must coexist with regular cellular users in the network. Therefore, while solutions to meet the rate requirements of D2D must be proposed, the QoS of cellular users must remain intact. This leads to the problem of interference and resource management, especially for D2D deployments of low frequency reuse patterns where interference can be a detrimental effect [173,174]. The problem is tightly coupled with the group communications challenge, since interference-related metrics can be employed to establish groups that facilitate the handling of interference. Schemes that effectively mitigate interference in this scenario as well as an adequate management of resources such as power or time are still needed.

9.1.4. D2D-based multi-hop communication

To maximize resource efficiency in D2D communications, it may be beneficial to enable multi-hop communication between two devices, thus transforming every intermediate D2D node into a relay. In this case, problems arising in multi-hop communication strategies must be analyzed in the context of D2D. Examples are the optimum number of relays, the time–frequency reuse pattern, or the end-to-end communication mechanism. Furthermore, network coding has been proposed as an efficient D2D communication strategy to enhance the information flow between two nodes far apart from each other that require intermediate relaying [175].

9.1.5. Other communication aspects

This last research challenge tackles other procedures and services related to D2D for which a solution needs to be found. Probably the most important aspect is the D2D neighbor discovery functionality, focused on identifying nearby D2D-enabled UEs [176,177]. The discovery functionality can be further classified into restricted discovery and open discovery depending on whether permission is needed or not. In addition, procedures for connection establishment and paging are required. Then, guaranteeing connectivity among devices is a further communication challenge, and D2D should incorporate the cellular fallback option in case of disruption in the direct communication. Moreover, solutions for different scenarios that differ in terms of network coverage are required. In particular, 3GPP classifies the scenarios into in-coverage, partial-coverage, and out-of-coverage, differing in the level of SINR that the user receives from the closest eNB. Finally, seamless mobility for users switching from D2D mode to cellular, security, and scalability are other important aspects for further study of this new communication paradigm.

10. Conclusions

The standardization of LTE became one of the most important technology shifts in cellular networks since the introduction of WCDMA. However, it was not until the introduction of LTE-Advanced in Rel-10 that the requirements established by the ITU for 4G technologies were finally achieved. Nevertheless, both industry and academia have continued improving LTE-Advanced through enhancements in the core technologies of carrier aggregation, MIMO, relaying, and cooperative multipoint communications. In addition, in order to cope with the ever increasing demand for ubiquitous and high-speed data access, the efforts have recently focused on improving the support of heterogeneous networks, as well as device-to-device and machine-type communications. In addition to these technologies, the new paradigm of self-organizing networks has the potential of transforming the cellular network into a dynamic entity capable of automatically adjusting itself to guarantee the best possible service by exploiting all the aforementioned technologies. This evolution will continue, not only by improving these technologies, but also by introducing new ones, especially at higher frequency bands capable of satisfying the demand for even faster data access than the ones seen today.

Acknowledgments

The authors would like to thank Eylem Ekici, Linda Xie, Won-Yeol Lee, Xudong Wang, Wenye Wang, and Huseyin Uzunalioglu for their valuable comments that improved the quality of this paper. Special thanks to John R. Krier for his contribution to the initial draft of this paper. This work has been partly supported by Fundacion Caja Madrid and SENACYT.

References

[1] 3GPP, Overview of 3GPP Release 8 V0.1.1, Tech. Rep., 2010.
[2] D. Astely, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, S. Parkvall, LTE: the evolution of mobile broadband, IEEE Commun. Mag. 47 (4) (2009) 44–51.
[3] A. Furuskar, T. Jönsson, M. Lundevall, The LTE radio interface—key characteristics and performance, in: IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC, 2008, pp. 1–5.
[4] ITU-R, Requirements Related to Technical Performance for IMT-Advanced Radio Interface(s), Report M.2134, 2008.
[5] 3GPP, TR 36.912 Feasibility Study for Further Advancements for E-UTRA (LTE-Advanced), Tech. Rep., 2010. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36912.htm.
[6] S. Parkvall, D. Astely, The evolution of LTE towards IMT-advanced, J. Commun. 4 (3) (2009).
[7] A. Ghosh, R. Ratasuk, B. Mondal, N. Mangalvedhe, T. Thomas, LTE-advanced: next-generation wireless broadband technology, IEEE Wirel. Commun. 17 (3) (2010) 10–22.
[8] P. Mogensen, T. Koivisto, K. Pedersen, I. Kovacs, B. Raaf, K. Pajukoski, M. Rinne, LTE-advanced: the path towards gigabit/s in wireless mobile communications, in: International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace Electronic Systems Technology, Wireless VITAE, 2009, pp. 147–151.
[9] ITU-R, Assessment of the Global Mobile Bradband Deployments and Forecasts for International Mobile Telecommunications, Report M.2134, 2011.
[10] 4G Americas, 4G Mobile Broadband Evolution: Release 10, Release 11 and Beyond, Tech. Rep., 2012.
[11] 3GPP, RP-13xxxx, Draft Report of 3GPP TSG RAN Meeting #58, Tech. Rep., 2012.
[12] 3GPP, RWS-120052, Report of 3GPP TSG RAN Workshop on Release 12 and onwards, Tech. Rep., 2012.
[13] E. Seidel, 3GPP LTE-A Standardization in Release 12 and Beyond, Tech. Rep., 2013.
[14] Radio Regulations, International Telecommunication Union Std., November 2012. [Online] Available: http://www.itu.int/pub/R-REG-RR/en.
[15] 3GPP, TS 22.011 Service Accessibility, Tech. Rep., 2011. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/22011.htm.
[16] 3GPP, TR 36.912 Feasibility Study for Further Advancements for E-UTRA (LTE-Advanced), Tech. Rep., 2012. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36912.htm.
[17] 3GPP, TR 36.823 Evolved Universal Terrestrial Radio Access (E-UTRA); Carrier aggregation enhancements; UE and BS Radio Transmission and Reception, Tech. Rep., 2013. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36823.htm.
[18] 3GPP, TR 36.850 Inter-band Carrier Aggregation Technical Report (Release 11), Tech. Rep., 2013. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36850.htm.
[19] 3GPP, TR 36.851 Inter-band Carrier Aggregation Technical Report (Release 12), Tech. Rep., 2012. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36851.htm.
[20] 3GPP, TS 36.300 Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall Description, Tech. Rep., 2012. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36300.htm.
[21] 3GPP, TS 36.323 Packet Data Convergence Protocol (PDCP) Specification, Tech. Rep., 2013. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36323.htm.
[22] 3GPP, TS 36.322 Radio Link Control (RLC) protocol specification, Tech. Rep., 2010. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36322.htm.
[23] 3GPP, TS 36.331 Radio Resource Control (RRC), Tech. Rep., 2013. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36331.htm.
[24] 3GPP, TS 36.321 Medium Access Control (MAC) Protocol Specification, Tech. Rep., 2013. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36321.htm.
[25] 3GPP, TS 36.201 Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Physical Layer; General Description, Tech. Rep., 2011. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36201.htm.
[26] 3GPP, TS 24.301 Non-Access-Stratum (NAS) Protocol for Evolved Packet System (EPS); Stage 3, Tech. Rep., 2013. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/24301.htm.

[27] R1-124776 On new carrier type, Ericsson, TSG RAN WG1 Meeting #71. November 2012.
[28] R1-122471 On the transmission mode(s) for the new carrier, Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, TSG RAN WG1 Meeting #69. May 2012.
[29] R1-121268 PSS/SSS and unsynchronized new carrier type, Ericsson, TSG RAN WG1 Meeting #68bis. March 2012.
[30] C. Hoymann, D. Larsson, H. Koorapaty, J.-F. Cheng, A lean carrier for LTE, 51 (2) (2013) 74–80.
[31] L. Liu, M. Li, J. Zhou, X. She, L. Chen, Y. Sagae, M. Iwamura, Component carrier management for carrier aggregation in LTE-advanced system, in: Proc. IEEE Vehicular Technology Conference, VTC, May 2011, pp. 1–6.
[32] C. Li, B. Wang, W. Wang, Y. Zhang, X. Chang, Component carrier selection for LTE-a systems in diverse coverage carrier aggregation scenario, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC, September 2012, pp. 1004–1008.
[33] K. Yagyu, T. Nakamori, H. Ishii, M. Iwamura, N. Miki, T. Asai, J. Hagiwara, Investigation on mobility management for carrier aggregation in LTE-advanced, in: Proc. IEEE Vehicular Technology Conference, VTC, September 2011, pp. 1–5.
[34] S. Gao, H. Tian, J. Zhu, L. Chen, X. She, A throughput-optimized component carrier selection algorithm for LTE-advanced systems, in: Proc. IET International Conference on Communication Technology and Application, October 2011, pp. 218–222.
[35] Z. Zou, C. Guo, Z. Zeng, S. Shi, A dynamic adjusting carrier coverage-based load balancing scheme in LTE-advanced system based on carrier aggregation, in: Proc. International Conference on Wireless Communications, Networking and Mobile Computing, September 2011, pp. 1–4.
[36] F. Liu, K. Zheng, W. Xiang, H. Zhao, Design and performance analysis of an energy-efficient uplink carrier aggregation scheme, IEEE J. Sel. Areas Commun. PP (99), pp. 1, 11, 0.
[37] H. Wang, H. Nguyen, C. Rosa, K. Pedersen, Uplink multi-cluster scheduling with MU-MIMO for LTE-advanced with carrier aggregation, in: Proc. IEEE Wireless Communications and Networking Conference, WCNC, April 2012, pp. 1202–1206.
[38] H.T. Nguyen, I. Kovacs, Downlink radio resource management for LTE-advanced system with combined MU-MIMO and carrier aggregation features, in: Proc. IEEE Vehicular Technology Conference, VTC, May 2012, pp. 1–5.
[39] A. Cattoni, H. Nguyen, J. Duplicy, D. Tandur, B. Badic, R. Balraj, F. Kaltenberger, I. Latif, A. Bhamri, G. Vivier, I. Kovacs, P. Horvath, Multi-user MIMO and carrier aggregation in 4G systems: The SAMURAI approach, in: Proc. IEEE Wireless Communications and Networking Conference Workshops, WCNCW, April 2012, pp. 288–293.
[40] Y. Li, Q. Mu, L. Liu, L. Chen, M. Peng, W. Wang, Control channel design for carrier aggregation between LTE FDD and LTE TDD systems, in: Proc. IEEE Vehicular Technology Conference, VTC, May 2012, pp. 1–5.
[41] 3GPP, TS 23.251 Network Sharing; Architecture and Functional Description, Tech. Rep., 2013. [Online] Available: http://ftp.3gpp.org/Specs/html-info/23251.htm.
[42] D. Bai, J. Lee, H. Nguyen, J. Singh, A. Gupta, Z. Pi, T. Kim, C. Lim, M.-G. Kim, I. Kang, LTE-advanced modem design: challenges and perspectives, 50 (2) (2012) 178–186. February.
[43] Amari LTE 100—softsoft LTE Base Station on a PC. Amarisoft. http://www.amarisoft.com/.
[44] Open Source Long-Term Evolution (LTE) deployment. OSLD Project. [Online]. Available: https://sites.google.com/site/osldproject/.
[45] NVIDIA, NVIDIA SDR (Software Defined Radio) technology. The modem innovation inside NVIDIA i500 and Tegra 4i, Whitepaper, 2013. [Online] Available: http://www.nvidia.com/docs/IO/116757/NVIDIA_i500_whitepaper_FINALv3.pdf.
[46] Y. Meng, G. Shen, W. Zheng, Q. Jiang, Y. Tang, H. Sun, J. Liu, PAPR reduction with multiple antennas transmission for carrier aggregation, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC, September 2012, pp. 615–619.
[47] P. Yen, H. Minn, C.-C. Chong, PAPR reduction for bandwidth-aggregated OFDM and SC-FDMA systems, in: Proc. IEEE Wireless Communications and Networking Conference, WCNC, September 2011, pp. 1363–1368.
[48] Y. Ikeda, I. Kanno, T. Kitayabu, H. Ishikawa, Proposal of cross-carrier precoding for carrier aggregation enhancement, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC, September 2011, pp. 1738–1742.
[49] K. Takeda, Y. Sagae, N. Ohkubo, H. Ishii, Investigation on rate matching and soft buffer splitting for LTE-advanced carrier aggregation, in: Proc. IEEE Vehicular Technology Conference, VTC, May 2012, pp. 1–5.
[50] C. Zhong, T. Yang, L. Zhang, J. Wang, A new discontinuous reception (DRX) scheme for LTE-advanced carrier aggregation systems with multiple services, in: Proc. IEEE Vehicular Technology Conference, VTC, September 2011, pp. 1–5.
[51] Qualcomm and Nokia, RSPG Consultation on Cognitive Technologies; Qualcomm and Nokia Corporation Joint Response, Tech. Rep., 2011. [Online] Available: http://rspg.groups.eu.int/consultations/consultation_cognitiv_2010/qualcomm_nokia_0114.pdf.
[52] Electronic Communications Committee, Description of Practices Relative to Trading of Spectrum Rights of Use, ECC Report 169, 2011, [Online] Available: http://www.erodocdb.dk/docs/doc98/official/pdf/ECCRep169.pdf.
[53] L. Juho, H. Jin-Kyu, J. Zhang, MIMO technologies in 3GPP LTE and LTE-advanced, EURASIP J. Wirel. Commun. (2009).
[54] L. Liu, R. Chen, S. Geirhofer, K. Sayana, Z. Shi, Y. Zhou, Downlink MIMO in LTE-advanced: SU-MIMO vs. MU-MIMO, IEEE Commun. Mag. 50 (2) (2012) 140–147.
[55] F. Boccardi, B. Clerckx, A. Ghosh, E. Hardouin, G. Joňgren, K. Kusume, E. Onggosanusi, Y. Tang, Multiple-antenna techniques in LTE-advanced, IEEE Commun. Mag. 50 (3) (2012) 114–121.
[56] C. Park, Y.-P. Wang, G. Joňgren, D. Hammarwall, Evolution of uplink MIMO for LTE-advanced, 49 (2) (2011) 112–121. February.
[57] M. Costa, Writing on dirty paper, 29 (3) (1983) 439–441.
[58] H. Viswanathan, S. Venkatesan, H. Huang, Downlink capacity evaluation of cellular networks with known-interference cancellation, 21 (5) (2003) 802–811.
[59] A. Tarighat, M. Sadek, A.H. Sayed, A multi user beamforming scheme for downlink MIMO channels based on maximizing signal-to-leakage ratios, in: Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, ICASSP, Vol. 3, March 2005.
[60] Q.H. Spencer, A.L. Swindlehurst, M. Haardt, Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels, 52 (2) (2004) 461–471.
[61] RP-121805 work plan for MIMO-related topics in Rel-12, Samsung, AT&T, CMCC, Nokia, Nokia Siemens Networks, NTT DOCOMO, TSG RAN WG1 Meeting #58, December 2012.
[62] RP-122015 study on full dimension MIMO for LTE, Samsung, TSG RAN WG1 Meeting #58, December 2012.
[63] W.Y. Zhang Wurong, Z. Rong, Active antenna system: utilizing the full potential of radio sources in the spatial domain, Huawei, Whitepaper, November 2012.
[64] J. Hoydis, S. ten Brink, M. Debbah, Massive MIMO in the UL/DL of cellular networks: how many antennas do we need? IEEE J. Sel. Areas Commun. 31 (2) (2013) 160–171.
[65] F. Rusek, D. Persson, B.K. Lau, E. Larsson, T. Marzetta, O. Edfors, F. Tufvesson, Scaling up MIMO: Opportunities and challenges with very large arrays, IEEE Signal Process. Mag. 30 (1) (2013) 40–60.
[66] H.Q. Ngo, T. Marzetta, E. Larsson, Analysis of the pilot contamination effect in very large multicell multiuser mimo systems for physical channel models, in: IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, May 2011, pp. 3464–3467.
[67] T.L. Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas, IEEE Trans. Wirel. Commun. 9 (11) (2010) 3590–3600.
[68] X. Lu, A. Tolli, O. Piirainen, M. Juntti, W. Li, Comparison of antenna arrays in a 3-D multiuser multicell network, in: IEEE International Conference on Communications, ICC, June 2011, pp. 1–6.
[69] D5.3: WINNER+ final channel models, Whitepaper, June 2010.
[70] 3GPP, TR 25.996 Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations, Tech. Rep., 2012. [Online] Available: http://www.3gpp.org/FTP/Specs/archive/25_series/25.996/.
[71] J. Lee, Y. Kim, H. Lee, B.L. Ng, D. Mazzarese, J. Liu, W. Xiao, Y. Zhou, Coordinated multipoint transmission and reception in LTE-advanced systems, IEEE Commun. Mag. 50 (11) (2012) 44–50.
[72] Q. Li, R. Hu, Y. Qian, G. Wu, Cooperative communications for wireless networks: techniques and applications in LTE-advanced systems, IEEE Wirel. Commun. 19 (2) (2012) 22–29.
[73] S. Sun, Q. Gao, Y. Peng, Y. Wang, L. Song, Interference management through CoMP in 3GPP LTE-advanced networks, 20 (1) (2013) 59–66. February.
[74] W. Liu, S. Han, C. Yang, Hybrid cooperative transmission in heterogeneous networks, in: IEEE International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC, September 2012, pp. 921–925.

[75] K. Huang, S. Lu, J. Guo, An optimized cooperative transmission scheme for interference mitigation in heterogeneous downlink network, in: IEEE Vehicular Technology Conference, VTC Fall, September 2012, pp. 1–5.
[76] B. Clerckx, Y. Kim, H. Lee, J. Cho, J. Lee, Coordinated multi-point transmission in heterogeneous networks: a distributed antenna system approach, in: IEEE International Midwest Symposium on Circuits and Systems, MWSCAS, August 2011, pp. 1–4.
[77] P. Marsch, G. Fettweis, A framework for optimizing the uplink performance of distributed antenna systems under a constrained backhaul, in: Proc. IEEE Int. Conf. Communications ICC, 2007, pp. 975–979.
[78] P. Marsch, G. Fettweis, A decentralized optimization approach to backhaul-constrained distributed antenna systems, in: Proc. 16th IST Mobile and Wireless Communications Summit, 2007, pp. 1–5.
[79] RP-121803 enhanced CoMP for LTE, Samsung, TSG RAN WG1 Meeting #58, December 2012.
[80] 3GPP, TR 36.814 Evolved Universal Terrestrial Radio Access (E-UTRA); Further Advancements for E-UTRA Physical Layer Aspects, (Release 9), Tech. Rep., 2010. [Online] Available: http://ftp.3gpp.org/Specs/html-info/36814.htm.
[81] R1-100951 Type 2 Relay Summary, Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, CHTTL, TSG RAN WG1 Meeting #60, February 2010.
[82] 3GPP, TS 36.216 Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer for Relaying Operation, Tech. Rep., 2012. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36216.htm.
[83] 3GPP, TR 36.806 Relay Architectures for E-UTRA (LTE-Advanced), Tech. Rep., 2010. [Online] Available: http://www.3gpp.org/ftp/specs/html-info/36806.htm.
[84] 3GPP, TS 23.401 General Packet Radio Service (GPRS) Enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access, Tech. Rep., 2013. [Online] Available: http://www.3gpp.org/ftp/specs/html-info/23401.htm.
[85] 3GPP, TS 33.401 3GPP System Architecture Evolution (SAE); Security Architecture, Tech. Rep., 2013. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/33401.htm.
[86] 3GPP, TS 31.102 Characteristics of the Universal Subscriber Identity Module (USIM) Application, Tech. Rep., 2012. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/31102.htm.
[87] 3GPP, TR 36.836 Mobile Relay for Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced), Tech. Rep., 2012. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36836.htm.
[88] R. Balakrishnan, X. Yang, M. Venkatachalam, I. Akyildiz, Mobile relay and group mobility for 4G WiMAX networks, in: Proc. of IEEE Wireless Communications and Networking Conference, WCNC, March 2011, pp. 1224–1229.
[89] R3-121350 proposed RAN3 way forward on mobile relays, InterDigital Communications, Orange, CATT, CMCC, ZTE, Qualcomm Incorporated, LGE, TSG-RAN WG3 Meeting RAN3#76, May 2012.
[90] W. Li, C. Zhang, X. Duan, S. Jia, Y. Liu, L. Zhang, Performance evaluation and analysis on group mobility of mobile relay for LTE advanced system, in: Proc. IEEE Vehicular Technology Conference, VTC, 2012, pp. 1–5.
[91] O. Bulakci, S. Redana, B. Raaf, J. Hamalainen, Performance enhancement in LTE-advanced relay networks via relay site planning, in: Proc. IEEE Vehicular Technology Conference, VTC, May 2010, pp. 1–5.
[92] A. Saleh, O. Bulakci, J. Hamalainen, S. Redana, B. Raaf, Analysis of the impact of site planning on the performance of relay deployments, IEEE Trans. Veh. Technol. 61 (7) (2012) 3139–3150.
[93] L. Rong, S. Elayoubi, O. Haddada, Impact of relays on LTE-advanced performance, in: Proc. IEEE International Conference on Communications, ICC, May 2010, pp. 1–6.
[94] O. Yilmaz, E. Mutafungwa, J. Hamalainen, Performance of relay enhanced LTE-Advanced networks for selected suburban scenarios in emerging market environments, in: International Symposium on Wireless Communication Systems, August 2012, pp. 914–918.
[95] C. Coletti, P. Mogensen, R. Irmer, Performance analysis of relays in LTE for a realistic suburban deployment scenario, in: Proc. IEEE Vehicular Technology Conference, VTC, May 2011, pp. 1–5.
[96] C. Coletti, P. Mogensen, R. Irmer, Deployment of LTE in-band relay and micro base stations in a realistic metropolitan Scenario, in: Proc. IEEE Vehicular Technology Conference, VTC, September 2011, pp. 1–5.
[97] A. Saleh, O. Bulakci, S. Redana, B. Raaf, J. Hamalainen, Enhancing LTE-advanced relay deployments via biasing in cell selection and handover decision, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC, September 2010, pp. 2277–2281.
[98] Y. Hu, L. Qiu, A novel multiple relay selection strategy for LTE-advanced relay systems, in: Proc. IEEE Vehicular Technology Conference, VTC, May 2011, pp. 1–5.
[99] G. Liebl, T. de Moraes, A. Soysal, E. Seidel, Fair resource allocation for the relay backhaul link in LTE-advanced, in: Proc. IEEE Wireless Communications and Networking Conference, WCNC, April 2012, pp. 1196–1201.
[100] S. Yi, M. Lei, Backhaul resource allocation in LTE-advanced relaying systems, in: Proc. IEEE Wireless Communications and Networking Conference, WCNC, April 2012, pp. 1207–1211.
[101] D.W. Kifle, O. Bulakci, A.B. Saleh, S. Redana, F. Granelli, Joint backhaul co-scheduling and relay cell extension in LTE-advanced networks uplink performance evaluation, in: European Wireless Conference, April 2012, pp. 1–8.
[102] O. Bulakci, A. Saleh, S. Redana, B. Raaf, J. Hamalainen, Flexible backhaul resource sharing and uplink power control optimization in LTE-advanced relay networks, in: Proc. IEEE Vehicular Technology Conference, VTC, September 2011, pp. 1–6.
[103] N. Krishnan, R.D. Yates, Narayan B. Mandayam, J.S. Panchal, Bandwidth sharing for relaying in cellular systems, 11 (1) (2012) 117–129. January.
[104] D. Lopez-Perez, I. Guvenc, G. De la Roche, M. Kountouris, T. Quek, J. Zhang, Enhanced intercell interference coordination challenges in heterogeneous networks, IEEE Wirel. Commun. 18 (3) (2011) 22–30.
[105] Z. Ren, A. Saleh, O. Bulakci, S. Redana, B. Raaf, J. Hamalainen, Joint interference coordination and relay cell expansion in LTE-advanced networks, in: Proc. IEEE Wireless Communications and Networking Conference, WCNC, April 2012, pp. 2874–2878.
[106] Z. Zhao, J. Wang, S. Redana, B. Raaf, Downlink resource allocation for LTE-advanced networks with Type1 relay nodes, in: Proc. IEEE Vehicular Technology Conference, VTC, September 2012, pp. 1–5.
[107] K. Zheng, B. Fan, J. Liu, Y. Lin, W. Wang, Interference coordination for OFDM-based multihop LTE-advanced networks, 18 (1) (2011) 54–63. February.
[108] B. Fan, A. Saadani, Full reuse resource partition for multihop relay networks with tree and mesh structures, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC, September 2011, pp. 1860–1864.
[109] O. Bulakci, A. Nedelcu, A. Saleh, S. Redana, J. Hmlinen, Impact of backhaul subframe misalignment on uplink system performance of LTE-advanced relay networks, in: Proc. IEEE Vehicular Technology Conference, VTC, September 2012, pp. 1–6.
[110] J. Gora, S. Redana, In-band and out-band relaying configurations for dual-carrier LTE-advanced system, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC, September 2011, pp. 1820–1824.
[111] K. Samdanis, M. Brunner, Self-organized network management functions for relay enhanced LTE-advanced systems, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC, September 2011, pp. 227–231.
[112] R. Combes, Z. Altman, E. Altman, Self-organizing relays: dimensioning, self-optimization, and learning, 9 (4) (2012) 487–500. December.
[113] W. Saad, Z. Han, T. Basar, M. Debbah, A. Hjorungnes, Network formation games among relay stations in next generation wireless networks, IEEE Trans. Commun. 59 (9) (2011) 2528–2542.
[114] S. Hanly, An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity, IEEE J. Sel. Areas Commun. 13 (7) (1995) 1332–1340.
[115] A. Sang, X. Wang, M. Madihian, R.D. Gitlin, Coordinated load balancing, handoff/cell-site selection, and scheduling in multi-cell packet data systems, in: Proc. 10th Annual International Conference on Mobile Computing and Networking, ser. MobiCom'04, ACM, New York, NY, USA, 2004, pp. 302–314.
[116] 3GPP, TS 36.814 Evolved Universal Terrestrial Radio Access (E-UTRA); Further Advancements for E-UTRA Physical Layer Aspects (Release 9), Tech. Rep., 2010. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36814.htm.
[117] H.-S. Jo, Y.J. Sang, P. Xia, J. Andrews, Outage probability for heterogeneous cellular networks with biased cell association, in: Proc. IEEE Global Telecommunications Conference, GLOBECOM, December 2011, pp. 1–5.
[118] J. Wang, J. Liu, D. Wang, J. Pang, G. Shen, Optimized fairness cell selection for 3GPP LTE-A macro-pico HetNets, in: Proc. IEEE Vehicular Technology Conference, VTC Fall, September 2011, pp. 1–5.
[119] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, S. Vrzic, Interference coordination and cancellation for 4G networks, IEEE Commun. Mag. 47 (4) (2009) 74–81.

[120] B. Soret, Y. Wang, K. Pedersen, CRS interference cancellation in heterogeneous networks for LTE-advanced downlink, in: Proc. IEEE International Conference on Communications, ICC, June 2012, pp. 6797–6801.
[121] S. Ye, S.H. Wong, C. Worrall, Enhanced physical downlink control channel in LTE advanced release 11, 51 (2) (2013) 82–89. February.
[122] R1-125335 PRB indication for EPDCCH, Huawei, TSG RAN WG1 Meeting #71, November 2012.
[123] 3GPP, TS 36.839 Evolved Universal Terrestrial Radio Access (E-UTRA); Mobility Enhancements in Heterogeneous Networks (Release 11), Tech. Rep., 2012. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/36839.htm.
[124] R2-123102: background search for small cell detection, Nokia Siemens Networks, Nokia Corporation, NTT DOCOMO, INC, TSG-RAN WG2 Meeting #78, 2012.
[125] R2-120277: enhanced cell identification and measurements for CA, NTT DOCOMO INC, TSG-RAN WG2 Meeting #77, 2012.
[126] R2-122368: enhanced MSE based small cell detection, Nokia Siemens Networks, Nokia Corporation, TSG-RAN WG2 Meeting #78, 2012.
[127] R2-120652: on UE-speed-based methods for improving the mobility performance in HetNets, Alcatel-Lucent, TSG-RAN WG2 Meeting #77, 2012.
[128] R2-121621: small cell signal based control of inter-frequency measurements, Nokia Siemens Networks, Nokia Corporation, TSG-RAN WG2 Meeting #77bis, 2012.
[129] R2-120654: Inter-frequency Small cell identification, Qualcomm Incorporated, TSG-RAN WG2 Meeting #77, 2012.
[130] R2-121665: using proximity indication for small-cell discovery, Renesas Mobile Europe Ltd., TSG-RAN WG2 Meeting #77bis, 2012.
[131] R2-121417: small cell detection, Ericsson, ST-Ericsson, TSG-RAN WG2 Meeting #77bis, 2012.
[132] R2-121248: small cell discovery in HetNet, Huawei, HiSilicon, TSG-RAN WG2 Meeting #77bis, March 2012.
[133] R2-121538: pico cell detection issues, Samsung, TSG-RAN WG2 Meeting #77bis, March 2012.
[134] R2-114362: On network-assisted pico cell discovery in LTE HetNets, Alcatel-Lucent, Tech. Rep., 2011.
[135] R. Bhagavatula, R. Heath, Adaptive bit partitioning for multicell intercell interference nulling with delayed limited feedback, 59 (8) (2011) 3824–3836.
[136] D.M. Gutierrez-Estevez, I.F. Akyildiz, Spatial coverage cross-tier correlation analysis for heterogeneous cellular networks, IEEE Trans. Veh. Technol. PP (99), pp. 1, 11, 0.
[137] R. Balakrishnan, I.F. Akyildiz, Local mobility anchoring for seamless handover in coordinated small cells, in: Proc. of IEEE GLOBECOM, December 2013.
[138] D. Pacheco-Paramo, I. Akyildiz, X2-interface-based location management for small cell networks, in: Proc. of IEEE GLOBECOM, December 2013.
[139] Y.-J. Kwon, D.-H. Cho, Load based cell selection algorithm for faulted handover in indoor femtocell network, in: Proc. IEEE Vehicular Technology Conference, VTC, May 2011, pp. 1–5.
[140] H.-Y. Lee, Y.-B. Lin, A cache scheme for femtocell reselection, IEEE Commun. Lett. 14 (1) (2010) 27–29.
[141] 3GPP, TS 32.500 Self-Organizing Networks (SON); Concepts and Requirements, Tech. Rep., 2011. [Online] Available: http://ftp.3gpp.org/Specs/html-info/32500.htm.
[142] 3GPP, TS 32.501 Self-configuration of Network Elements; Concepts and Requirements, Tech. Rep., 2012. [Online] Available: http://ftp.3gpp.org/Specs/html-info/32501.htm.
[143] 3GPP, TS 32.502 Self-Configuration of Network Elements Integration Reference Point (IRP); Information Service (IS), Tech. Rep., 2012. [Online] Available: http://ftp.3gpp.org/Specs/html-info/32502.htm.
[144] 3GPP, TS 32.506 Self-configuration of Network Elements Integration Reference Point (IRP); Solution Set (SS) Definitions, Tech. Rep., 2012. [Online] Available: http://ftp.3gpp.org/Specs/html-info/32506.htm.
[145] 3GPP, TS 32.511 Automatic Neighbour Relation (ANR) Management; Concepts and Requirements, Tech. Rep., 2012. [Online] Available: http://ftp.3gpp.org/Specs/html-info/32511.htm.
[146] 3GPP, TS 32.823 Self-Organizing Networks (SON); Study on Self-healing, Tech. Rep., 2009. [Online] Available: http://ftp.3gpp.org/Specs/html-info/32823.htm.
[147] 3GPP, TS 32.821 Study of Self-Organizing Networks (SON) Related Operations, Administration and Maintenance (OAM) for Home Node B (HNB), Tech. Rep., 2009. [Online] Available: http://ftp.3gpp.org/Specs/html-info/32821.htm.
[148] 3GPP, TS 36.902 Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Self-configuring and Self-optimizing Network (SON) use Cases and Solutions, Tech. Rep., 2011. [Online] Available: http://ftp.3gpp.org/Specs/html-info/36902.htm.
[149] H. Zhuang, Z. Luo, M. Pikhletsky, S. Dmitri, F. Khafizov, Traffic-based dynamic spectrum management for inter-cell interference coordination in LTE networks, in: Proc. of IEEE GLOBECOM, December 2012.
[150] M. Maso, M. Debbah, L. Vangelista, A distributed approach to interference alignment in OFDM-based two-tiered networks, IEEE Trans. Veh. Technol. 62 (5) (2013) 1935–1949.
[151] D. Lopez-Perez, X. Chu, A.V. Vasilakos, H. Claussen, On distributed and coordinated resource allocation for interference mitigation in self-organizing LTE networks, IEEE/ACM Trans. Netw. 21 (4) (2013) 1145–1158.
[152] A. Awada, B. Wegmann, I. Viering, A. Klein, Cell-pair specific optimization of the inter-RAT handover parameters in SON, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC, September 2012, pp. 1168–1173.
[153] K. Kitagawa, T. Komine, T. Yamamoto, S. Konishi, A handover optimization algorithm with mobility robustness for LTE systems, in: Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communications, PIMRC, September 2011, pp. 1647–1651.
[154] R. Kwan, Distributed energy-saving mechanism for self-organizing femto lte networks, in: Proc. IEEE Vehicular Technology Conference, VTC, Ubiquisys Ltd., Swindon, 2012, pp. 1–5.
[155] E. Chavarria-Reyes, I.F. Akyildiz, Radio Access Network energy minimization in multi-layer heterogeneous wireless systems, in: Proc. of IEEE PIMRC, September 2013.
[156] I. Karla, Resolving son interactions via self-learning prediction in cellular wireless networks, in: Proc. Int. Conf. Wireless Communications, Networking and Mobile Computing, WICOM, September 2012, pp. 1–6.
[157] Worldwide cellular M2M Modules Forecast Market Brief, Beecham Research, Tech. Rep., 2010.
[158] S. Lucero, Maximizing mobile operator opportunities in M2M: the benefits of an M2M-optimized network, ABI Research, Tech. Rep., 2010.
[159] 3GPP, Study on Facilitating Machine to Machine Communication in 3GPP systems; (Release 8), TR 22.868, Tech. Rep., 2007. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/22868.htm.
[160] E. MCC, R2-101881: Report of 3GPP TSG RAN WG2 Meeting 68bis, 3GPP TSG RAN WG2 Meeting #68bis, Tech. Rep., 2010.
[161] 3GPP, TS 22.368 Service requirements for Machine-Type Communications (MTC); Stage 1 (Release 10), Tech. Rep., 2010. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/22368.htm.
[162] S.-Y. Lien, K.-C. Chen, Y. Lin, Toward ubiquitous massive accesses in 3GPP machine-to-machine communications, IEEE Commun. Mag. 49 (4) (2011) 66–74.
[163] C. Bontu, E. Illidge, DRX mechanism for power saving in LTE, IEEE Commun. Mag. 47 (6) (2009) 48–55.
[164] T. Tirronen, A. Larmo, J. Sachs, B. Lindoff, N. Wiberg, Reducing energy consumption of LTE devices for machine-to-machine communication, in: IEEE Globecom Workshops, GC Wkshps, 2012, pp. 1650–1656.
[165] S.C. Jha, A.T. Koc, M. Gupta, R. Vannithamby, Power saving mechanisms for M2M communication over LTE networks, in: First International Black Sea Conference on Communications and Networking, BlackSeaCom, 2013, pp. 102–106.
[166] S.-Y. Lien, T.-H. Liau, C.-Y. Kao, K.-C. Chen, Cooperative access class barring for machine-to-machine communications, IEEE Trans. Wirel. Commun. 11 (1) (2012) 27–32.
[167] H. Chao, Y. Chen, J. Wu, Power saving for machine to machine communications in cellular networks, in: IEEE GLOBECOM Workshops, GC Wkshps, 2011, pp. 389–393.
[168] 3GPP, TR 37.868 Study on RAN Improvements for Machine-type Communications; (Release 11), Tech. Rep., 2011. [Online] Available: http://www.3gpp.org/ftp/Specs/html-info/37868.htm.
[169] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, K. Hugl, Device-to-device communication as an underlay to LTE-advanced networks, IEEE Commun. Mag. 47 (12) (2009) 42–49.
[170] L. Lei, Z. Zhong, C. Lin, X. Shen, Operator controlled device-to-device communications in LTE-advanced networks, IEEE Wirel. Commun. 19 (3) (2012) 96–104.
[171] N. Golrezaei, M. Ji, A. Molisch, A. Dimakis, G. Caire, Device-to-device communications for wireless video delivery, in: Conference Record of the 46th Asilomar Conference on Signals, Systems and Computers, ASILOMAR, 2012, pp. 930–933.

[172] P. Phunchongharn, E. Hossain, D. Kim, Resource allocation for device-to-device communications underlaying LTE-advanced networks, IEEE Wirel. Commun. 20 (4) (2013) 91–100.
[173] H. Min, J. Lee, S. Park, D. Hong, Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks, IEEE Trans. Wirel. Commun. 10 (12) (2011) 3995–4000.
[174] D. Feng, L. Lu, Y. Yuan-Wu, G. Li, G. Feng, S. Li, Device-to-device communications underlaying cellular networks, IEEE Trans. Commun. 61 (8) (2013) 3541–3551.
[175] B. Zhou, H. Hu, S.-Q. Huang, H.-H. Chen, Intracluster device-to-device relay algorithm with optimal resource utilization, IEEE Trans. Veh. Technol. 62 (5) (2013) 2315–2326.
[176] K. Doppler, C.B. Ribeiro, J. Kneckt, Advances in D2D communications: Energy efficient service and device discovery radio, in: Proc. 2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, Wireless VITAE, IEEE, 2011, pp. 1–6.
[177] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, Z. Turányi, Design aspects of network assisted device-to-device communications, IEEE Commun. Mag. 50 (3) (2012) 170–177.

Ian F. Akyildiz received the B.S., M.S., and Ph.D. degrees in Computer Engineering from the University of Erlangen-Nürnberg, Germany, in 1978, 1981 and 1984, respectively. Currently, he is the Ken Byers Chair Professor in Telecommunications with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, the Director of the Broadband Wireless Networking Laboratory and Chair of the Telecommunication Group at Georgia Tech. Dr. Akyildiz is an honorary professor with the School of Electrical Engineering at Universitat Politècnica de Catalunya (UPC) in Barcelona, Catalunya, Spain, and the founder of N3Cat (NaNoNetworking Center in Catalunya). He is also an honorary professor with the Department of Electrical, Electronic and Computer Engineering at the University of Pretoria, South Africa and the founder of the Advanced Sensor Networks Lab. Since 2011, he is a Consulting Chair Professor at the Department of Information Technology, King Abdulaziz University (KAU) in Jeddah, Saudi Arabia. Since January 2013, Dr. Akyildiz is also a FiDiPro Professor (Finland Distinguished Professor Program (FiDiPro) supported by the Academy of Finland) at Tampere University of Technology, Department of Communications Engineering, Finland. He is the Editor-in-Chief of Computer Networks (Elsevier) Journal, and the founding Editor-in-Chief of the Ad Hoc Networks (Elsevier) Journal, the Physical Communication (Elsevier) Journal and the Nano Communication Networks (Elsevier) Journal. He is an IEEE Fellow (1996) and an ACM Fellow (1997). He received numerous awards from IEEE and ACM. His current research interests are in nanonetworks, Long Term Evolution (LTE) advanced.

David M. Gutierrez-Estevez obtained his Engineering Degree in Telecommunications from Universidad de Granada, Spain, in July 2009. During the summer of 2007 he held an internship position at the Audio Department of Fraunhofer Institute for Integrated Circuits in Erlangen, Germany. From April 2008 to June 2009 he was a research assistant within the Broadband Mobile Communications Networks Department of the Fraunhofer Heinrich Hertz Institute in Berlin, where he worked on projects linked to Long Term Evolution (LTE) cellular systems. He received his M.S. degree in Electrical and Computer Engineering from Georgia Tech in 2011 with a fellowship from Fundación la Caixa. Currently, he is pursuing his Ph.D. degree at the Broadband Wireless Networking Laboratory at Georgia Tech under the supervision of Dr. Ian. F. Akyildiz with a fellowship from Fundación Caja Madrid.

Ravikumar Balakrishnan obtained the B.E. Degree in Electronics and Communication Engineering from Anna University, India, in 2009 and the M.S. Degree from the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, in 2011. Currently, he is pursuing his Ph.D. degree in the Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta. Mr. Balakrishnan held an internship position at the Wireless Technology Division, Intel Corporation focusing on next generation cellular network architecture and protocols. His areas of research include Next Generation Cellular Networks and Cognitive Radio Networks. He is a student member of IEEE.

Elias Chavarria Reyes received the B.E. degree in Electronics and Communication Engineering from Universidad de Panamá, Ciudad de Panamá, Panamá, in 2007. He received his M.S. degree in Electrical and Computer Engineering from the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, in 2010. Currently, he is pursuing his Ph.D. degree in the Broadband Wireless Networking Laboratory, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, with a fellowship from "SENACYT". His current research is focused on Heterogeneous Networks and Self-Organizing Networks. He is a student member of IEEE.