Uplink multi-cluster scheduling with MU-MIMO for LTE-advanced with carrier aggregation
Wang, Hua; Nguyen, Hung Tuan; Rosa, Claudio; Pedersen, Klaus

Published in:
Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)

DOI (link to publication from Publisher):
10.1109/WCNC.2012.6213960

Publication date:
2012

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):
Uplink Multi-Cluster Scheduling with MU-MIMO for LTE-Advanced with Carrier Aggregation

Hua Wang†, Hung Nguyen†, Claudio Rosa‡, and Klaus Pedersen⋄
†Radio Access Technology, Aalborg University, Aalborg, Denmark
‡Nokia Siemens Networks - Research, Aalborg, Denmark
Email: †huw@es.aau.dk

Abstract—LTE-Advanced is the evolutionary path from LTE Release 8. It is designed to significantly enhance the performance of LTE Release 8 in terms of higher peak data rates, improved system capacity and coverage, and lower latency. These enhancements allow LTE-Advanced to meet or exceed the IMT-Advanced requirements and are being considered as part of LTE Release 10. In this paper, some of the physical layer enhancement techniques for LTE-Advanced have been studied including carrier aggregation (CA), uplink multi-cluster scheduling, and uplink multi-user multiple-input multiple-output (MU-MIMO) with non-overlapping allocations. A system-level simulation was conducted to investigate the performance gains that can be achieved in uplink CA with multi-cluster scheduling and MU-MIMO. Simulation results show that with proper differentiation between power-limited and non-power-limited LTE-A users, multi-cluster scheduling with CA has similar coverage performance as in Rel’8, but can achieve substantial gains in average user throughput compared with Rel’8. MU-MIMO can further improve the throughput performance, especially when MU-MIMO is combined with multi-cluster scheduling.

I. INTRODUCTION

Long Term Evolution (LTE) is one of the primary broadband technologies based on Orthogonal Frequency Division Multiplexing (OFDM) for next-generation mobile communication systems. LTE Release 8 (Rel’8) was finalized in March 2009 providing peak data rates of 300 Mbps in downlink and 75 Mbps in uplink with 20 MHz bandwidth, and allowing flexible bandwidth operation of up to 20 MHz [1]. To further enhance the performance, the 3rd Generation Partnership Project (3GPP) started a new study item in March 2008 on evolving from LTE towards LTE-Advanced (also known as LTE Release 10), targeting to meet or exceed the IMT-Advanced requirements defined by the International Telecommunication Union (ITU), i.e., peak data rates up to 1 Gbps in downlink and 500 Mbps in uplink. The study item was closed in March 2010 with a set of new radio features. These enhancements are being considered as part of LTE-Advanced in 3GPP Rel’10.

Carrierg aggregation (CA) is one of the key features for LTE-Advanced to support a much wider transmission bandwidth up to 100 MHz compared with legacy LTE Rel’8, enabling peak date rates requirement to be satisfied. This is achieved by aggregating two or more individual component carriers (CCs) of the same or different bandwidth belonging to contiguous or non-contiguous frequency bands, subject to spectrum availability and the user equipment (UE)’s capability [2]. In addition to bandwidth extension, multi-cluster transmission has been proposed to improve the spectral efficiency in uplink. With multi-cluster transmission, a UE can be allocated to maximum of 2 clusters not adjacent to each other, so it has higher scheduling flexibility compared with Single Carrier Frequency Division Multiple Access (SC-FDMA), which is the multiple access scheme in LTE uplink. With CA or multi-cluster transmission, the single carrier property in the uplink is no longer preserved. As a result, the Peak-to-Average Power Ratio (PAPR) increases, which will cause an effective reduction of the maximum UE transmission power [1]. This effect together with channel estimation loss experienced with multi-CC multi-cluster transmission might reduce the throughput for cell-edge users, since they usually experience unfavorable channel conditions and are limited by the transmission power. Therefore, the assignment of CCs and clusters to users has to be carefully designed. LTE-Advanced supports enhanced multi-user multiple-input multiple-output (MU-MIMO) with non-overlapping allocations. Up to four layers of quasi-orthogonal UE-specific reference signals is available for MU-MIMO enabling co-scheduling of up to four UEs in the same time-frequency resource, which can provide substantial gains in sector throughput. Each of these features has been studied individually in [3]-[5], but to the authors’ knowledge, there has been no study to evaluate the overall performance by taking all of them into considerations. The objective of this paper is to perform a system-level study to investigate the performance gains that can be achieved in uplink CA with multi-cluster scheduling and MU-MIMO, both the overall gain and individual gains under different traffic models.

The rest of the paper is organized as follows. Section II gives a general overview of CA in LTE-Advanced, as well as multi-cluster scheduling, power back-off model, coverage improvement, and MU-MIMO. Section III outlines the simulation methodology and main assumptions. Simulation results and performance analysis are presented in Section IV. Finally, some conclusions are drawn in Section V.

II. RADIO RESOURCE MANAGEMENT

The radio resource management (RRM) framework for multi-CC LTE-Advanced system is illustrated in Fig. 1. Separate RRM blocks operate independently on each CC. It has been agreed within 3GPP working group to adopt independent Link Adaption (LA) and Hybrid ARQ (HARQ) per CC in coherence with LTE Rel’8 assumptions [2]. Such strategy
maintains the backward compatibility so that an LTE Rel’8 terminal can work in an LTE-Advanced system. The admission control module in base station decides whether an incoming connection should be accepted or not. Then the CC selection module allocates one or multiple CCs to the incoming user based on the Quality-of-Service (QoS) requirements, terminal capability, etc. For the layer 2 packet scheduler, since a user may be allocated on multiple CCs, the per-CC time and frequency domain packet scheduler could support joint scheduling across multiple assigned CCs [4] to achieve better performance in terms of fairness and coverage. Since the UEs are limited by the transmission power, power control is also an important issue in uplink. It is worth mentioning that admission control, CC-selection, and packet scheduling are not part of the standard, but are vendor specific.

A. Multi-Cluster Scheduling in Uplink

In downlink, 3GPP uses Orthogonal Frequency Division Multiple Access (OFDMA) due to its high spectral efficiency and good performance in link adaptation and frequency domain scheduling. But one of the main challenges in OFDMA is the high PAPR of the transmitted signal [6]. Therefore in uplink, 3GPP uses SC-FDMA for multiple access. SC-FDMA uses many of the OFDMA principles to achieve high spectral efficiency. In SC-FDMA, UEs can only be scheduled on one set of continuous sub-carriers. In the time domain, only a single modulation symbol is sent at a time. This allows SC-FDMA to reach a very low signal PAPR, facilitating efficient power amplifiers in mobile devices. Since UEs can only be scheduled on a continuous part of the spectrum, SC-FDMA has less frequency diversity gain compared to OFDMA.

For LTE-A uplink, additional multiple access schemes have been proposed which include $N \times$ SC-FDMA, also known as multi-cluster transmission. In multi-cluster transmission, the minimum resource allocation unit is a sub-band, which consists of integer number of physical resource blocks (PRBs). Several contiguous sub-bands can be seen as a cluster, and a UE can be allocated to multiple clusters not adjacent to each other. Multi-cluster transmission has low signal PAPR compared to OFDMA, and higher scheduling flexibility compared to SC-FDMA. It can be considered as a compromise of OFDMA and SC-FDMA. It is worth mentioning that 3GPP only supports up to 2 clusters per CC for multi-cluster transmission.

B. Resource Allocation and Scheduling

In [7], the authors proposed an adaptive transmission bandwidth (ATB) based packet scheduling algorithm for SC-FDMA in uplink, which tightly couples the bandwidth allocation and the packet scheduling together to exploit the bandwidth flexibility. The basic idea behind ATB is to produce an allocation table which closely follows the envelope of the UEs’ scheduling metrics. In this study, we extended the ATB algorithm so that it is catered to multi-cluster transmission. At each scheduling instance, the algorithm generates a scheduling matrix of each UE on each sub-band. The algorithm first selects a UE with the highest scheduling metric and then checks the power constraint of that UE. If it is out of power with new resource allocation, remove that UE from the candidate list and restart from the beginning. After the power constraint check, the algorithm checks the number of allocated clusters. If the new allocation causes the cluster number to exceed the maximum number limitation, the current scheduling metric grid is disabled and the algorithm restarts from the beginning. Otherwise, the scheduler allocates the current sub-band to that UE and expands its transmission bandwidth until either another UE has a higher scheduling metric on the adjacent sub-band or the maximum transmission power of that UE is exceeded. Then the algorithm restarts from the beginning and continues the loop until either all UEs have been scheduled or there are no resources left. A detailed description of the proposed multi-cluster scheduling is illustrated in Fig. 2.

C. Maximum Power Reduction for Non-contiguous Allocations

An important issue in non-contiguous resource allocation (e.g., multi-CC or multi-cluster transmission) in uplink is the increased PAPR and other RF related issues. Studies have
shown that the PAPR increases when a UE is transmitting over non-contiguous allocations simultaneously, which requires a larger power back-off in the power amplifier together with other RF transmission requirements, thereby reducing the maximum transmission power at the UE [8]. Maximum Power Reduction (MPR) proposed in [9] is used to model the power reduction at UE due to increased PAPR and other RF imperfections. Defining a MPR scheme for non-contiguous resource allocation is challenging because there are many dimensions in the signal that affect the required back off, such as the number of clusters, size of clusters, frequency separation between clusters, etc. In this study, we adopt the MPR scheme proposed in [9] that can be used for single and dual CC cases no matter how many clusters are being allocated. The MPR value (in dB) is calculated based on the ratio between the allocated PRBs and the aggregated system transmission bandwidth, specified as follows:

\[
P_{\text{MPR}} = \begin{cases} 
6.2 & \text{if } 0 < A \leq 0.05 \\
7 - 16A & \text{if } 0.05 < A \leq 0.25 \\
3.83 - 3.33A & \text{if } 0.25 < A \leq 0.4 \\
2.83 - 0.83A & \text{if } 0.4 < A \leq 1 
\end{cases} 
\]

where \( A = N_{\text{PRB,alloc}}/N_{\text{PRB,agg}} \) is the ratio between the allocated PRBs and total available PRBs.

Therefore, if a UE is scheduled for transmission only on one cluster, there is no additional power back-off. Otherwise, the UE’s maximum power is reduced by \( P_{\text{MPR}} \) in dB.

**D. CC Selection and Coverage Improvement**

The main difference of LTE-A RRM framework compared to Rel’8 is the CC-selection functionality which is responsible for configuring a CC set for each UE based on their QoS requirements, UE capability, etc. The legacy Rel’8 users can only be assigned on one CC, while LTE-A users can be assigned on multiple CCs. The CC-selection functionality is important to perform load balancing among CCs, as well as to optimize the system performance.

In downlink, allocating more CCs to an LTE-A user generally results in a higher throughput thanks to the larger transmission bandwidth and higher transmission power. However, this is not always the case in uplink. The main difference between uplink and downlink is the transmission power constraint of a UE. For power limited cell edge users, even if they are assigned on multiple CCs, they do not have sufficient power to exploit the increased transmission bandwidth. Furthermore, the impact of increased PAPR will introduce additional reduction of maximum UE transmission power as mentioned previously. For power limited cell edge users transmitting at (or close to) maximum transmission power, such cost might counterbalance the gain brought by multi-CC multi-cluster transmission, and even results in a coverage loss compared to the case where the SC-FDMA properties of the transmitted signals are maintained (single-CC single-cluster assignment). Therefore, multi-CC multi-cluster transmission shall generally be restricted to users with good channel conditions. An effective pathloss-threshold based CC-selection algorithm was proposed in [5] to distinguish between power-limited and non-power-limited LTE-A users. The derived path loss threshold is:

\[
L_{\text{threshold}} = L_{95\%} - \frac{10 \log_{10}(K) + P_{\text{MPR}}}{\alpha} 
\]

where \( L_{95\%} \) is the estimated 95-percentile user path loss, \( K \) is the total number of CCs, \( \alpha \) is the path loss compensation factor, and \( P_{\text{MPR}} \) is the average MPR.

LTE-A users whose path loss is lower than the threshold are considered to be power-limited and are assigned on single-CC and single-cluster, otherwise they are considered to be non-power-limited and can be assigned on multi-CC and multi-cluster. By doing so, cell-edge users will not experience any performance loss from being scheduled over multiple CCs and clusters, while non-power-limited LTE-A users can benefit from the advantages of CA and multi-cluster scheduling. For Rel’8 users, they are assigned on one CC with least load for load balancing and single-cluster.

**E. Multi-user MIMO**

MU UEs selection and scheduling are based on the SU UE scheduled list described in Section II-C and II-B. Only scheduled UEs which have the estimated SINR larger than a pairing threshold are selected into a primary MU-MIMO UE list. This condition is applied to avoid the scheduling of UEs at the cell-edge in MU-MIMO mode. For each primary MU-MIMO UE, a pairing MU-MIMO UE is found from the candidate list \( \Omega_c \). The primary UEs and the UEs scheduled for retransmission are excluded from the candidate list. The pairing UE which has the highest MU-MIMO metric will be scheduled for MU-MIMO transmission together with the primary UE.

\[
\text{UE}_i^{\text{pair}} = \arg\max_{k \in \Omega_c} \beta_k \cdot (1 - \gamma_{k,i}) 
\]

where \( \beta_k \) is the scheduling metric of pairing UE \( k \), and \( \gamma_{k,i} \) is the cross correlation between the precoder of the primary UE \( i \) and the pairing UE \( k \). It should be noted here that, both the user priority and the UE mutual information are considered in this MU-MIMO scheduling metric.

Since the MU-MIMO transmissions are carried out from two different UEs, no power splitting or power sharing as in DL MU-MIMO transmission is required. The selection and scheduling of MU-MIMO UEs are processed until all resource is used or all UEs are scheduled.

**III. Simulation Assumptions**

The performance evaluation is based on a detailed multi-cell system level simulator which follows the guidelines in [10]. The simulation scenario is 3GPP Macro-cell case #1 with 7 sites and 3 sectors per site using the wrap-around technique. Spatial Channel Model (SCM) and 3D antenna pattern with default tilt of 15 degrees are used. Two contiguous CCs, each with 20 MHz bandwidth, are configured to form a wide band of 40 MHz. On each CC, up to two clusters are supported for uplink transmission. Imperfect Sounding Reference Signal (SRS) measurements and channel estimation
settings used in the system-level simulations. The average UE power back-off is set to be 6 dB (quite "aggressive" setting when considering the power back-off model in eqn.1). Table I summarizes the main parameter values used in our simulations.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propagation scenario</td>
<td>3GPP Macro case #1</td>
</tr>
<tr>
<td>Layout</td>
<td>7 sites - 3 sectors/site - wrap around</td>
</tr>
<tr>
<td>Channel profile</td>
<td>SCM channel model with 3D antenna</td>
</tr>
<tr>
<td>Component carriers</td>
<td>2 × 20 MHz contiguous @ 2GHz band</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>1 or 2 clusters</td>
</tr>
<tr>
<td>PRBs per sub-band</td>
<td>2 PRBs</td>
</tr>
<tr>
<td>Sounding resolution</td>
<td>2 PRBs</td>
</tr>
<tr>
<td>Sounding period</td>
<td>10 ms</td>
</tr>
<tr>
<td>Sounding method</td>
<td>Imperfect SRS with channel estimation error</td>
</tr>
<tr>
<td>eNode-B receiver</td>
<td>4-Rx MMSE</td>
</tr>
<tr>
<td>UE Tx bandwidth</td>
<td>ATB for multi-cluster</td>
</tr>
<tr>
<td>Packet scheduling</td>
<td>Throughput based Proportional Fair</td>
</tr>
<tr>
<td>Bursty traffic model</td>
<td>Finite buffer with Poisson arrival</td>
</tr>
<tr>
<td>Available MCSs</td>
<td>Fixed size of 2 Mbits per UE</td>
</tr>
<tr>
<td>Max UE power</td>
<td>200 mW [23 dBm]</td>
</tr>
<tr>
<td>HARQ</td>
<td>Synchronous and adaptive</td>
</tr>
<tr>
<td>BLER target</td>
<td>20%</td>
</tr>
<tr>
<td>Link adaptation</td>
<td>Fast AMC</td>
</tr>
<tr>
<td>α</td>
<td>0.6</td>
</tr>
<tr>
<td>Power spectral density</td>
<td>Load Adaptive Power Control</td>
</tr>
<tr>
<td>Average MPR $\overline{T_MPR}$</td>
<td>6 dB</td>
</tr>
</tbody>
</table>

TABLE I
SUMMARY OF MAIN SIMULATION PARAMETERS

error are assumed. The link to system level mapping is based on the actual value interface (AVI) method [11]. It is assumed that distance-dependent path loss and shadowing are maintained constant for each UE, but fast fading is updated every TTI independently on each CC based on the ITU-T Typical Urban power delay profile and UEs’ speed. Both full buffer and bursty traffic models are considered. In full-buffer traffic model, we assume each sector has a fixed number of UEs with full buffer. In bursty traffic model, each call in a sector follows a Poisson arrival process with a finite buffer of 2 Mbits payload. The offered load per cell can be obtained by multiplying the user arrival rate with the payload size. Proportional fair scheduling in frequency domain is used together with multi-cluster bandwidth allocation. Same open loop power control settings are applied on each CC, while independent Load Adaptive Power Control (LAPC) [12] is enabled on each CC to dynamically update the UE power spectral density $P_0$ based on the variable load conditions. The average UE power back-off $\overline{T_MPR}$ is set to be 6 dB (quite "aggressive" setting when considering the power back-off model in eqn.1). Table I summarizes the main parameter settings used in the system-level simulations.

IV. SIMULATION RESULTS

We start our analysis by first looking at the scenario with fixed number of UEs per sector and full-buffer traffic model. Only one CC is configured with 10 MHz bandwidth in this scenario. Fig. 3 shows the average throughput gain versus different number of UEs per sector. The reference scenario is single cluster scheduling without MU-MIMO. The gain of dual-cluster scheduling over single cluster scheduling increases as the number of UEs increases until reaching the maximum value at certain point, i.e., 6 UEs per sector in our case. Then the dual-cluster scheduling gain gradually decreases as the number of UEs increases. That is because when the number of UEs per sector is high, dual-cluster scheduling may even result in lower multi-user diversity gain. MU-MIMO gain increases monotonically as the number of UEs increases due to the reason that higher number of UEs will bring higher multi-user diversity. When MU-MIMO is combined with dual-cluster scheduling, the average throughput gain can be up to 56% with 10 UEs per sector compared with single cluster scheduling without MU-MIMO. It is much higher than the sum of gains brought by dual-cluster scheduling alone and MU-MIMO alone because MU-MIMO gains more chance to exploit frequency diversity than single cluster.

Fig. 3. Average user throughput gain versus different number of UEs per sector, 1 × 10 MHz with full-buffer traffic model

dual-cluster scheduling and can be applied in different scenarios. It is shown that the coverage of LTE-A UEs is almost the same as that of Rel’8 UEs. In other words, there is no gain in coverage by applying multi-cluster scheduling or CA in uplink. That is because at the cell edge, UEs usually experience high path loss and are limited by the maximum transmission power. In fact, with LAPC deployed, cell edge UEs are configured to transmit with maximum power. Even if those cell edge LTE-A UEs are assigned to multiple clusters and CCs, they do not have sufficient power to exploit the increased transmission bandwidth, and may even result in a performance loss due to the effect of MPR. Therefore, cell edge LTE-A UEs are assigned on only one cluster and one

Fig. 4. shows the cell edge user throughput versus the offered load in different scenarios. It is shown that the average throughput gain can be up to 56% with 10 UEs per sector compared with single cluster scheduling without MU-MIMO. It is much higher than the sum of gains brought by dual-cluster scheduling alone and MU-MIMO alone because MU-MIMO gains more chance to exploit frequency diversity than single cluster. Therefore it is recommended that dual-cluster scheduling is used in combination with MU-MIMO.

Next we evaluate the performance of multi-cluster scheduling with MU-MIMO and CA in bursty traffic model. Two CCs, each with 20 MHz bandwidth, are configured. Fig. 4 shows the average user throughput gain versus the offered load in different scenarios. It is shown that the coverage of LTE-A UEs is almost the same as that of Rel’8 UEs. In other words, there is no gain in coverage by applying multi-cluster scheduling or CA in uplink. That is because at the cell edge, UEs usually experience high path loss and are limited by the maximum transmission power. In fact, with LAPC deployed, cell edge UEs are configured to transmit with maximum power. Even if those cell edge LTE-A UEs are assigned to multiple clusters and CCs, they do not have sufficient power to exploit the increased transmission bandwidth, and may even result in a performance loss due to the effect of MPR. Therefore, cell edge LTE-A UEs are assigned on only one cluster and one
CC, which results in almost the same coverage performance compared with Rel’8 UEs.

Fig. 5 shows the average user throughput versus the offered load in different scenarios. With CA and multi-cluster scheduling, the average user throughput of LTE-A UEs is significantly higher than that of Rel’8 UEs. In our proposed CC selection algorithm, we distinguish between power-limited and non-power-limited LTE-A UEs based on users’ pathloss. Power-limited LTE-A UEs are assigned on only one CC and single cluster so that they will not experience any loss from being scheduled over multiple CCs and clusters, while LTE-A UEs not operating close to their maximum transmission power are assigned on multiple CCs and dual clusters so that they can benefit from the advantages of CA (i.e., increased transmission bandwidth) and multi-cluster scheduling (i.e., frequency domain diversity). MU-MIMO can further increase the user throughput, and the MU-MIMO gain increases as the offered load increases due to increased multi-user diversity.

V. CONCLUSIONS

In this paper, we have investigated the uplink multi-cluster scheduling with MU-MIMO for LTE-Advanced with carrier aggregation. The concept of multi-cluster scheduling, power back-off model for non-contiguous resource allocation, pathloss-threshold based CC selection, and MU-MIMO have been introduced. Two sets of simulation scenarios have been considered. With full-buffer traffic model, the simulation results show the gain of multi-cluster scheduling gets saturated at certain point, while MU-MIMO gain increases as the number of UEs increases in both single and dual-cluster transmissions. But MU-MIMO gain in dual-cluster scheduling is much higher than that in single-cluster scheduling because dual-cluster scheduling allows to fully exploit the gain of MU-MIMO. When multi-cluster scheduling is combined with MU-MIMO, the average user throughput gain can be up to 56% compared with single cluster scheduling without MU-MIMO. With finite-buffer bursty traffic model, cell edge user throughput of LTE-A UEs maintains almost the same as Rel’8 UEs, but substantial gains in average user throughput can be achieved with multi-CC multi-cluster scheduling and MU-MIMO.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to Nokia Siemens Networks colleagues Huiyu Yuan, Yuyu Yan, Jens Steiner, and Mads Brix, for their assistance in the development of system-level simulator.

REFERENCES