Efficient Coordination of Almost Blank Subframes with Coupling Macro Cells in Heterogeneous Networks

You-Chiun Wang and Bo-Jyun Huang

Abstract—Heterogeneous networking is an inevitable trend in mobile communication systems, where macro-cells and small-cells coexist to provide the service of wireless access. The technique of eICIC (enhanced inter-cell interference coordination) is popularly used to mitigate signal interference among cells. A number of subframes are picked as *almost blank subframes (ABSs)*, during which a macro-cell base station suspends data transmissions to improve signal quality of small-cells that it covers. Since the throughput of macro-cells drops to zero in an ABS, how to decide the number of ABSs (i.e., the ABS ratio) and their positions (i.e., the ABS placement) is critical. Many studies implicitly consider that every small-cell resides in only one macro-cell, and they manage ABSs for each individual macro-cell accordingly. In practice, some macro-cells may cover the same small-cells (called *coupling macro-cells*), which makes the management of their ABSs depend on each other. To address this issue, the paper proposes an *efficient coordination of ABS with coupling macro-cells (ECO)* algorithm. It divides the network into *component subnetworks* depending on whether macro-cells are coupling or not. Then, ECO computes the ABS ratios for the macro-cells in each component subnetwork based on multiple factors, including data rate, traffic demand, and QoS (quality of service) support. Moreover, ECO jointly places the ABSs of coupling macro-cells to shorten packet latency. Experimental results show that the ECO algorithm not only increases the network throughput but also reduces the loss rate of real-time data.

Index Terms—almost blank subframe (ABS), coupling macro-cells, heterogeneous network, interference management.

1 Introduction

To fulfill people's needs of high-speed wireless access, many operators deploy various types of base stations (also known as eNBs) in their mobile communication systems. These eNBs work together and organize a *heterogeneous network*. In particular, macro-cell eNBs take charge of providing seamless signal coverage of the service area. On the other hand, small-cell eNBs are placed within macro-cells to intensify signal quality in certain regions (e.g., stores or buildings) and also share their traffic loads to avoid network congestion [1]. Heterogeneous networking will be one of the most important features in 5G systems [2].

The transmitted power of a macro-cell eNB is typically 40 watts (i.e., 46 dBm), whereas that of a small-cell eNB is no more than 2 watts (i.e., 33 dBm) [3]. Inevitably, the *user equipments (UEs)* served in a small-cell will encounter large signal interference from nearby macro-cell eNBs. To deal with this problem, the Third Generation Partnership Project (3GPP) proposes *enhanced inter-cell interference coordination (eICIC)* that allows time-sharing of spectrum resources between macro-cells and small-cells. In each period, a macro-cell eNB inserts some muted slots, namely *almost blank subframes (ABSs)*, into its transmission schedule, where it sends nothing but merely control signals. As these signals occupy a small fraction of subcarriers, the interference to small-cell UEs will greatly decrease. Thus, a small-cell eNB can send data to its UEs with a much higher rate.

Intuitively, the idea behind eICIC is to increase the throughput of small-cells at expense of data transmissions of macro-cells in ABSs. Therefore, the number of ABSs in each period (i.e., the ABS ratio) plays a key role in deciding network performance. It is a challenge to find a suitable ABS ratio to balance between data transmissions in macro-cells and small-cells, so as to maximize the overall throughput. Most of existing studies implicitly assume that each small-cell is covered solely by one macro-cell, and strive to derive the optimal ABS ratio for a single macro-cell based on its status, for instance, traffic demand and the number of UEs. Nevertheless, how to decide the positions of ABSs in a period (i.e., the ABS placement), especially when some macro-cells are coupling, is rarely addressed.

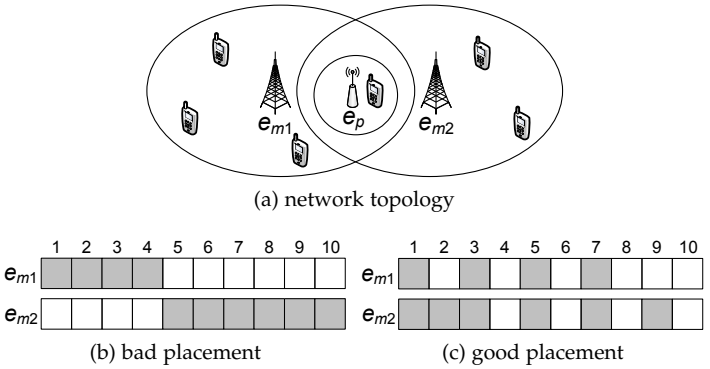


Fig. 1: Two examples of ABS placement, where gray slots indicate ABSs.

When two macro-cells cover the same small-cells (i.e., they reside in the overlap of these macro-cells), the two macro-cells are said to be *coupling* with each other. In this case, the management of their ABSs will depend on each other. Fig. 1 illustrates an example, where a small-cell eNB e_p locates inside both signal coverage of two macro-cell eNBs e_{m1} and e_{m2}. Suppose that the ABS ratios of e_{m1} and e_{m2} are set to 0.4 and

The authors are with the Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, 80424, Taiwan. E-mail: ycwang@cse.nsysu.edu.tw; m053040002@student.nsysu.edu.tw

0.6, respectively. If one takes the ABS placement in Fig. 1(b), the UEs served by e_p will be always interfered by either e_{m1} or e_{m2}, which degrades the throughput of that small-cell. By contrast, Fig. 1(c) gives better placement of ABSs, where e_p can transmit data to its UEs nearly without interference in subframes 1, 3, 5, and 7.

Based on the above motivation, this paper proposes an *efficient coordination of ABS with coupling macro-cells (ECO)* algorithm. The network is cut into *component subnetworks* based on the coupling relationship of macro-cells. To calculate the ABS ratio, ECO estimates the data rate supported by each eNB and checks whether using ABSs can improve the overall throughput in a component subnetwork. Then, it measures the amount of data required to be sent based on the QoS (quality of service) demand of each UE with real-time flows. On the premise that macro-cell UEs will not be starved, the ABS ratio is adaptively adjusted to raise the throughput of small-cells and satisfy QoS demands of their UEs. Moreover, when two macro-cells are coupling with each other, ABSs are arranged in their transmission schedules with the aim of minimizing the dropping ratio of real-time packets. Simulation results show that, by taking coupling macro-cells into consideration, the proposed ECO algorithm can significantly improve the network throughput and also reduce data loss of real-time flows, as compared with other methods.

This paper is outlined as follows: Section 2 surveys related work and Section 3 presents the network model. The ECO algorithm is proposed in Section 4, followed by performance evaluation in Section 5. Finally, Section 6 concludes this paper and gives the future work.

2 RELATED WORK

In the literature, it attracts attention to calculate an optimal ABS ratio based on the assumption that UEs always have downlink data to receive (called the *full-buffer model*). In particular, Pang et al. [4] use dynamic programming to estimate the ABS ratio, so as to achieve high throughput and keep transmission fairness. Trabelsi et al. [5] test all possible numbers of ABSs in each period to get the best answer, but such brute-force search is time-consuming. Singh and Andrews [6] model the locations of eNBs and UEs as two independent Poisson point processes, and derive the best ABS ratio through probabilistic analysis. Ayala-Romero et al. [7] adopt a response surface method to approximate the optimal ABS ratio. However, the full-buffer model can be applied to only few applications such as FTP (file transfer protocol) traffics.

A number of studies adjust the ABS ratio based on multiple factors in the network. Specifically, Vasudevan et al. [8] divide UEs into three categories: macro-cell, small-cell-center, and small-cell-edge UEs. Both throughput and fairness maximization problems are then formulated by sum-rate and product-rate utilities, which refer to data rates and queue lengths of the UEs in each category. Daeinabi et al. [9] compute the ABS ratio via a fuzzy logic controller, which takes the number of UEs, the throughput of macro-cell UEs, and the signal quality of small-cell UEs to be its inputs. Bartoli et al. [10] consider three scenarios for each eight-slot period: one ABS, two ABSs, and one ABS along with a reduced-power slot. According to traffic loads in different cells, a feasible scenario is picked to improve the total throughput. Lu et al. [11] estimate the throughput of UEs with and without the effect of ABS. Based on the difference between the above two throughput, the ABS ratio is iteratively fine-tuned (e.g., adding or subtracting by a fixed value), until there is no significant increase in the overall performance. However, the issue of packet latency is not addressed in these studies.

Some research efforts are dedicated to supporting QoS for real-time services. To provide better playback quality of videos, Daeinabi et al. [12] adopt a genetic algorithm to decide the ABS ratio, whose fitness function takes account of distortion, throughput, and delay of video flows. Argyriou et al. [13] allocate resources in ABSs and other subframes based on video quality, which is measured by the mean square error of video frames. Wang and Huang [14] jointly configure the parameters of ABS and DRX (discontinuous reception) to increase the energy efficiency of UEs while reducing their packet dropping ratios. Tang et al. [15] cope with both resource scheduling and ABS configuration by formulating a problem of stochastic optimization programming. They use the Lyapunov optimization technique to solve the problem, with the aim of raising the total throughput and reducing packet latency. Our previous work [16] proposes a *delay-aware ABS adjustment (DA3)* method to support QoS for real-time flows. DA3 measures the capacity of each eNB, and decides the ABS ratio to transmit as many urgent data as possible. However, none of the aforementioned studies consider the practical case where some macro-cells cover the same small-cells. This motivates us to develop the ECO algorithm to not only find the suitable ABS ratios but also place ABSs for coupling macro-cells, so as to increase the overall throughput and also reduce data loss of real-time flows.

3 NETWORK MODEL

Let us consider an LTE-A (long term evolution-advanced) heterogeneous network composed of macro-cells and small-cells, where all small-cells reside in macro-cells. Two macro-cells may cover the same small-cells (i.e., coupling), in the sense that these small-cells locate within their overlapping region, as shown in Fig. 2. The topology of cells is known in advance and eNBs can exchange their information via an X2 interface [17]. Besides, each UE is served by only one eNB, where the eNB maintains a queue to store its downlink data.

To specify the QoS requirement of a flow, 3GPP adopts *QoS class identifier (QCI)* which associates with two parameters [18]. The delay budget indicates the maximum latency that the flow can bear, and the loss rate gives an upper bound on the packet dropping ratio. Based on QCI, flows are classified into *guaranteed-bit-rate (GBR)* and *non-GBR* ones. In general, GBR flows support real-time applications with strict delay demands (e.g., VoIP and video), while non-GBR flows are applied to services with loose deadlines (e.g., web browsing). Therefore, GBR flows could have smaller QCIs and delay budgets than non-GBR ones. For ease of presentation, a UE with GBR flows is called a *GBR UE*.

The spectrum resource is divided into disjointed *physical resource blocks (PRBs)*, each with 0.5 ms duration and 180 kHz bandwidth. PRBs are considered as exclusive resources, which means that the same PRB cannot be given to multiple UEs [19]. The eNB presides over allotting PRBs to the UEs in its cell. When the bandwidth of a downlink channel is 1.4, 3, 5, 10, 15, and 20 MHz, there are respectively 6, 15, 25, 50, 75, and 100 PRBs available in a subframe, whose duration is 1 ms. The amount of data carried by a PRB depends on its *modulation and coding scheme (MCS)*. A more complex MCS lets the PRB carry more data, but it requires less noise interference, and vice

TABLE 1: Summary of notations.

notation	definition
$\hat{E}, \hat{E}_p$	set of all/small-cell eNBs in a component subnetwork
$\hat{U}_j, \hat{U}_j^{\mathbf{G}}$	set of all/GBR UEs served by eNB e_j
$S_{i,j}^{\mathbf{O}}, S_{i,j}^{\mathbf{A}}$	SINR of UE u_i served by eNB e_j in an OS/ABS
$Q_i^{\mathbf{O}}, Q_i^{\mathbf{A}}$	CQI of UE u_i in an OS/ABS
$D_j^{\mathbf{G},\mathbf{O}}, D_j^{\mathbf{G},\mathbf{A}}$	GBR data transmitted by eNB e_j in an OS/ABS
$D_j^{\mathbf{O}}$	the amount of data transmitted by eNB e_j in an OS
L_i	length of effective queued data of UE u_i
$L_i^{\mathrm{PU}}, L_i^{\mathrm{SU}}, L_i^{\mathrm{QD}}$	length of PU/SU/QD data of UE u_i
T	the number of subframes in a period
N_A	the number of ABSs in a period (ABS ratio: $\gamma = N_A/T$)
β	PU-relation indicator
δ	threshold on the average data loss rate of GBR UEs in a macro-cell

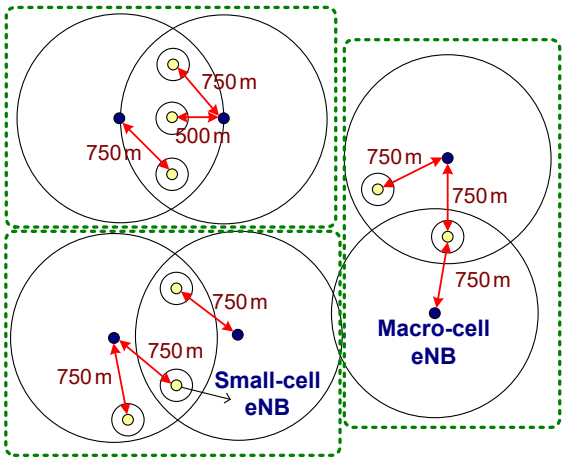


Fig. 2: An example of component subnetworks.

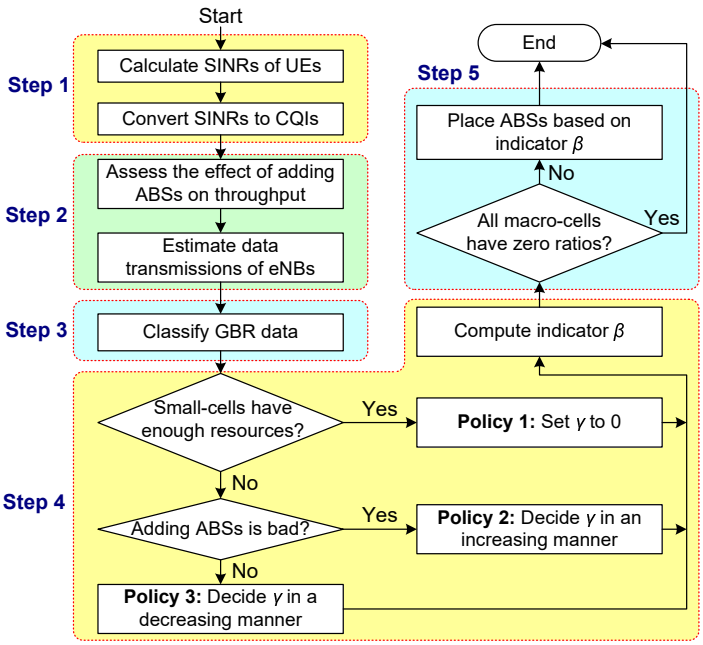


Fig. 3: Flowchart of the ECO algorithm.

versa. To help the eNB choose a suitable MCS, each UE has to report a *channel quality indicator (CQI)* which reveals and ranks its current channel condition.

The time axis is sliced into periods, where each period contains T subframes. Then, the *ABS management problem* asks how to decide the ABS ratio (denoted by γ) of each macro-cell and also its placement of ABSs in every period, such that the network throughput is maximized and the amount of data loss of GBR flows due to expiration is minimized. Furthermore, the problem puts a constraint that the average data loss rate of GBR UEs in macro-cells should keep below a threshold δ. Since a macro-cell has no throughput in ABSs, this constraint can prevent its GBR UEs from starvation. Table 1 summarizes the common notations.

4 The Proposed ECO Algorithm

From the viewpoint of macro-cells, we divide the network into non-overlapping *component subnetworks*. A component subnetwork is defined by the minimum set of macro-cells (together with their covered small-cells) such that 1) any macro-cell must be coupling with another macro-cell in the same component subnetwork and 2) any macro-cell in a component subnetwork will not be coupling with macro-cells in other component subnetworks. Fig. 2 gives an example, where the network is partitioned into three component subnetworks, each marked by a dotted rectangle.

Because macro-cells in any two component subnetworks are not coupling, we can separately decide ABS ratios and placement for the macro-cells in each component subnetwork. Therefore, our discussion aims at one component subnetwork, which contains a set $\hat{E}$ of eNBs. Fig. 3 illustrates the flowchart of our ECO algorithm, which comprises five steps to manage ABSs:

1) *Evaluate channel quality:* Each eNB collects CQIs of its UEs. To do so, a UE measures the values of SINR (signal-to-interference-plus-noise ratio) in an ABS and *ordinary subframe (OS)*.
2) *Assess network condition:* Given the channel quality of UEs, we check whether using more ABSs will decrease the overall throughput or not. In addition, each eNB calculates the amount of data transmissions in its cell.
3) *Classify GBR data:* For each GBR UE, we classify the downlink data in its queue according to packet deadlines and QoS demands.
4) *Set ABS ratios:* The ABS ratio of each macro-cell is decided by the network condition (assessed in step 2), the expected loss rate of GBR data in the macro-cell, and QoS demands of small-cell UEs. Moreover, we compute an indicator β based on the amount of urgent data in the macro-cell and its small-cells for the purpose of ABS placement.
5) *Decide ABS placement:* According to β and γ, macro-cell eNBs place their ABSs in the current period, with

TABLE 2: The MCS associated with each CQI and its required SINR (in dB).

CQI	MCS	SINR	CQI	MCS	SINR
1	QPSK (78)	-6.936	9	16QAM (616)	8.573
2	QPSK (120)	-5.147	10	64QAM (466)	10.366
3	QPSK (193)	-3.180	11	64QAM (567)	12.289
4	QPSK (308)	-1.253	12	64QAM (666)	14.173
5	QPSK (449)	0.761	13	64QAM (772)	15.888
6	QPSK (602)	2.699	14	64QAM (873)	17.814
7	16QAM (378)	4.694	15	64QAM (948)	19.829
8	16QAM (490)	6.525			

the objective of minimizing the number of dropped packets due to expiration.

4.1 Evaluate Channel Quality

Suppose that a UE u_i receives its downlink data from an eNB e_j through a subchannel c_k. Then, u_i's SINR can be calculated by

$$S_{i,j,k} = \frac{\tilde{\mathrm{P}}(u_i, e_j, c_k)}{\varepsilon N_0 B_k + \sum_{e_x \in \hat{E} - \{e_j\}} \tilde{\mathrm{P}}(u_i, e_x, c_k)}, \quad (1)$$

where $\tilde{\mathrm{P}}(u_i, e_j, c_k)$ is the power strength of e_j's signal received by u_i via c_k, which depends on the transmitted power of e_j and the distance between u_i and e_j. The signal from other eNBs (i.e., e_x) are considered as interference. The term $\varepsilon N_0 B_k$ gives the amount of interference caused by the environmental noise, where ε is the noise figure, N_0 is the power spectral density, and B_k is the bandwidth of c_k. Based on the 3GPP specification [20], N_0 is set to -174 dBm/Hz (to model the additive white Gaussian noise) and B_k is equal to 15 kHz.

Because the power strength in Eq. (1) also depends on whether the eNB is using an ABS or OS, two cases need to be further discussed.

Case 1: *e_j is a macro-cell eNB.* If e_j is using an OS, u_i's SINR (denoted by $S^{\mathbf{O}}_{i,j,k}$) is reckoned by Eq. (1). Otherwise, e_j will not send user data, so $\tilde{\mathrm{P}}(u_i, e_j, c_k)$ approximates to zero. By Eq. (1), u_i's SINR in an ABS (denoted by $S^{\mathbf{A}}_{i,j,k}$) is negligibly small.

Case 2: *e_j is a small-cell eNB.* When e_j resides in only one macro-cell with eNB e_m, u_i is in an OS or ABS if e_m is using an OS or ABS, respectively. However, when e_j is covered by two macro-cells whose eNBs are e_{m1} and e_{m2}, u_i is said to be in an OS if both e_{m1} and e_{m2} are using OSs. Otherwise, u_i is in an ABS. Here, we consider *partial ABS*, where either e_{m1} or e_{m2} is using an ABS, and compute two values of $S_{i,j,k}$ by setting either $\tilde{\mathrm{P}}(u_i, e_{m1}, c_k)$ or $\tilde{\mathrm{P}}(u_i, e_{m2}, c_k)$ to zero in Eq. (1). Then, we take their average to be the value of $S^{\mathbf{A}}_{i,j,k}$.

Let $\hat{C}$ be the set of subchannels used by e_j to send data to u_i. Then, the effective SINR of u_i in relation to e_j can be calculated as follows [21]:

$$S_{i,j} = -\tau \ln \left[\frac{1}{|\hat{C}|} \sum\nolimits_{k=1}^{|\hat{C}|} \exp(-S_{i,j,k}/\tau) \right], \quad (2)$$

where τ is a constant (e.g., $\tau = 1.57$) and $\exp(\cdot)$ is an exponential function. Again, we denote by $S^{\mathbf{O}}_{i,j}$ and $S^{\mathbf{A}}_{i,j}$ the effective SINRs of u_i in relation to e_j in an OS and ABS, respectively.

Table 2 gives the MCS associated with each CQI and its minimum required SINR [22], where QAM (quadrature amplitude modulation) is more complex than QPSK (quadrature phase-shift keying). In addition, each number in parentheses gives the corresponding code rate multiplied by 1024. According to Table 2, the CQI of u_i is the maximum one whose required SINR is smaller than $S_{i,j}$, and the eNB will encode u_i's data by using the associated MCS in its PRBs. For convenience, let us denote by $Q^{\mathbf{O}}_i$ and $Q^{\mathbf{A}}_i$ the CQIs of u_i in an OS and ABS, respectively. Lemma 1 analyzes the time complexity of step 1.

Lemma 1. Given n_U UEs, n_E eNBs, and n_C subchannels, the time complexity of step 1 is $O(n_U(n_E + n_C))$.

Proof: In step 1, we first compute the SINR of every UE by Eq. (1) for both ABS and OS cases, which spends time of $2n_U O(n_E)$. Then, calculating the effective SINR of each UE by Eq. (2) (also for both ABS and OS cases) spends time of $2n_U O(n_C)$. Finally, it takes a constant time to find the CQI of a UE by consulting Table 2. Since we need to decide both $Q^{\mathbf{O}}_i$ and $Q^{\mathbf{A}}_i$ for each UE, this operation consumes time of $2n_U O(1)$. To sum up, the time complexity of step 1 is $2n_U O(n_E) + 2n_U O(n_C) + 2n_U O(1) = O(n_U(n_E + n_C))$. □

4.2 Assess Network Condition

Afterwards, we estimate the theoretical data rate supported by each eNB. In particular, given the SINR of every UE served by a macro-cell eNB e_m, its maximum data rate in OSs can be calculated based on the Shannon-Hartley theorem [23] as follows:

$$R^{\mathbf{O}}_m = B \lg \left[1 + \sum_{u_i \in \hat{U}_m} \frac{S^{\mathbf{O}}_{i,m}}{|\hat{U}_m|} \right], \quad (3)$$

where $\hat{U}_m$ is the set of UEs served by e_m and $B = \sum_{c_k \in \hat{C}} B_k$ (i.e., the total bandwidth of all subchannels). According to the discussion in Section 4.1, since $S^{\mathbf{A}}_{i,m}$ is almost zero for each u_i in $\hat{U}_m$, e_m evidently has no data rate in ABSs. Similarly, the maximum data rate supported by all small-cell eNBs located in the macro-cell coordinated by e_m is calculated as below:

$$R^{\mathbf{O}}_p = \sum_{e_j \in \hat{E}_p} B \lg \left[1 + \sum_{u_i \in \hat{U}_j} \frac{S^{\mathbf{O}}_{i,j}}{|\hat{U}_j|} \right] \text{ for the OS case,} \quad (4)$$

$$R^{\mathbf{A}}_p = \sum_{e_j \in \hat{E}_p} B \lg \left[1 + \sum_{u_i \in \hat{U}_j} \frac{S^{\mathbf{A}}_{i,j}}{|\hat{U}_j|} \right] \text{ for the ABS case,} \quad (5)$$

where $\hat{E}_p$ denotes the set of small-cell eNBs covered by the macro-cell whose eNB is e_m. In case that a small-cell eNB is covered by two macro-cells whose eNBs are e_m and e_x, we consider that e_m is using an ABS but e_x is using an OS for the calculation of $S^{\mathbf{A}}_{i,j}$ in Eq. (5), as mentioned earlier in Section 4.1. Theorem 1 then analyzes the effect of using ABS on the overall throughput.

Theorem 1. If $R^{\mathbf{A}}_p - R^{\mathbf{O}}_p > R^{\mathbf{O}}_m$, adding more ABSs must increase the overall throughput. Otherwise, the overall throughput will be either flat or down with more ABSs when $R^{\mathbf{A}}_p - R^{\mathbf{O}}_p = R^{\mathbf{O}}_m$ or $R^{\mathbf{A}}_p - R^{\mathbf{O}}_p < R^{\mathbf{O}}_m$, respectively.

Proof: Let γ be the ABS ratio of a macro-cell eNB e_m. Then, the maximum data rate supported by e_m and its subordinate small-cell eNBs (i.e., those small-cell eNBs covered by the macro-cell whose eNB is e_m) can be calculated by

$$\begin{aligned} R &= (1-\gamma)R^{\mathbf{O}}_m + (1-\gamma)R^{\mathbf{O}}_p + \gamma R^{\mathbf{A}}_p \\ &= (R^{\mathbf{A}}_p - R^{\mathbf{O}}_m - R^{\mathbf{O}}_p) \times \gamma + (R^{\mathbf{O}}_m + R^{\mathbf{O}}_p). \end{aligned} \quad (6)$$

The coefficient $(R_p^{\mathbf{A}} - R_m^{\mathbf{O}} - R_p^{\mathbf{O}})$ decides the effect of the ABS ratio γ on the data rate R. Specifically, there are three cases to be discussed. Case 1 occurs when the coefficient is positive (i.e., $R_p^{\mathbf{A}} - R_p^{\mathbf{O}} > R_m^{\mathbf{O}}$). By Eq. (6), raising γ will also increase R. In other words, the overall throughput can be increased by using more ABSs. Case 2 is that the coefficient is zero (i.e., $R_p^{\mathbf{A}} - R_p^{\mathbf{O}} = R_m^{\mathbf{O}}$). In this case, R is always equal to $(R_m^{\mathbf{O}} + R_p^{\mathbf{O}})$, no matter how γ changes. Thus, the overall throughput keeps flat when $R_p^{\mathbf{A}} - R_p^{\mathbf{O}} = R_m^{\mathbf{O}}$. Finally, case 3 occurs when the coefficient is negative (i.e., $R_p^{\mathbf{A}} - R_p^{\mathbf{O}} < R_m^{\mathbf{O}}$). From Eq. (6), increasing γ will make R become smaller. Thus, adding more ABSs reduces the overall throughput instead. Consequently, the argument of these three cases verifies this theorem. □

As mentioned earlier in Section 3, the spectrum resource is materialized by PRBs for the purpose of easy allocation. Therefore, we estimate the amount of GBR and non-GBR data that an eNB can transmit in units of PRBs. In particular, the 3GPP standard [24] proposes three tables for estimating the number of data bits that can be received by a UE according to its CQI and the number of allocated PRBs. Specifically, the *CQI-to-MCS table* translates a CQI index into an MCS index. Then, the *MCS-to-TBS table* converts the MCS index into a TBS (transport block size) index, where TBS indicates the number of data bits carried by one PRB. Finally, the *TBS-to-bit table* takes the TBS index and the number of PRBs as inputs, and outputs the amount of data (in bits) that can be sent to the UE through these PRBs. Let us use a function $\tilde{\mathrm{F}}(Q_i, n_i)$ to represent the above table looking-up procedure, where Q_i is the CQI of a UE u_i and n_i is the number of PRBs allocated to u_i. To decide the value of n_i, one can refer to the result of a PRB scheduling method. For example, M-LWDF (modified largest weighted delay first) [25] is used as the PRB scheduling method in our simulations.

For each eNB e_j, the amount of GBR data that e_j can transmit in a subframe is calculated as follows:

$$D_j^{\mathbf{G},\mathbf{O}} = \sum\nolimits_{u_i \in \hat{U}_j^{\mathbf{G}}} \tilde{\mathrm{F}}(Q_i^{\mathbf{O}}, n_i) \text{ for the OS case,} \quad (7)$$

$$D_j^{\mathbf{G},\mathbf{A}} = \sum\nolimits_{u_i \in \hat{U}_j^{\mathbf{G}}} \tilde{\mathrm{F}}(Q_i^{\mathbf{A}}, n_i) \text{ for the ABS case,} \quad (8)$$

where $\hat{U}_j^{\mathbf{G}} \subseteq \hat{U}_j$ is the set of GBR UEs served by e_j. Moreover, we also estimate the amount of data sent by e_j in an OS:

$$D_j^{\mathbf{O}} = \sum\nolimits_{u_i \in \hat{U}_j} \tilde{\mathrm{F}}(Q_i^{\mathbf{O}}, n_i). \quad (9)$$

The above calculation in Eqs. (7), (8), and (9) together with the observation in Theorem 1 can give a course on how to decide the ABS ratio, which will be further discussed in Section 4.4. Lemma 2 then analyzes the time complexity of step 2.

Lemma 2. It takes $O(n_{Ep} n_{Up} + n_U)$ time to run step 2, where n_{Ep} is the number of small-cell eNBs, n_{Up} is the number of small-cell UEs, and n_U is the number of total UEs.

Proof: To get the observation in Theorem 1, we need to calculate $R_m^{\mathbf{O}}$, $R_p^{\mathbf{O}}$, and $R_p^{\mathbf{A}}$. Specifically, finding $R_m^{\mathbf{O}}$ by Eq. (3) takes $O(n_{Um})$ time, where n_{Um} is the number of macro-cell UEs. On the other hand, computing $R_p^{\mathbf{O}}$ and $R_p^{\mathbf{A}}$ by Eqs. (4) and (5) spends time of $2n_{Ep} \times O(n_{Up})$. After getting the values of $R_m^{\mathbf{O}}$, $R_p^{\mathbf{O}}$, and $R_p^{\mathbf{A}}$, checking the relationship between $(R_p^{\mathbf{A}} - R_p^{\mathbf{O}})$ and $R_m^{\mathbf{O}}$ consumes a constant time. Thus, the total time spent in the calculation of Theorem 1 is $O(n_{Um}) + 2n_{Ep} \times O(n_{Up}) + O(1) = O(n_{Ep} n_{Up})$. Then, finding $D_j^{\mathbf{G},\mathbf{O}}$, $D_j^{\mathbf{G},\mathbf{A}}$, and $D_j^{\mathbf{O}}$ by Eqs. (7), (8), and (9) takes $O(n_U)$ time, because we use the table looking-up procedure to calculate the amount of data that UEs can receive and the procedure spends a constant time to do the calculation for each UE. To sum up, the time complexity of step 2 is $O(n_{Ep} n_{Up} + n_U)$. □

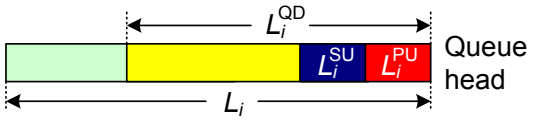


Fig. 4: Classification of the downlink data for a GBR UE.

4.3 Classify GBR Data

Since GBR data have strict delay demands, they should be given with a high priority for transmission. Therefore, it is necessary to take a deep look into the queue of each GBR UE u_i. In particular, we consider the *effective queued data*, where those packets that must expire (even though they are sent out immediately) will be discarded by the eNB to avoid wasting the bandwidth. Let L_i denote the length of u_i's effective queued data. Then, the data can be further categorized by their deadlines as follows:

- *Primarily urgent (PU) data:* A packet p_k belongs to the PU category if its delay time t_k satisfies the condition of

$$t_k \geq t_k^{\mathrm{DB}} - t_k^{\mathrm{TX}} - T\rho, \quad (10)$$

  where t_k^{TX} is the transmission time of p_k from the eNB to its UE (also known as the propagation delay), ρ is the length of a subframe (i.e., 1 ms), and t_k^{DB} is the delay budget of p_k's flow (measured in milliseconds). Thus, once p_k cannot be sent out in the current period, it will be dropped due to expiration, where a period contains T subframes. The length of u_i's PU data is denoted by L_i^{PU}.
- *Secondarily urgent (SU) data:* When the delay time t_k of packet p_k violates Eq. (10) but meets the condition of

$$t_k \geq t_k^{\mathrm{DB}} - t_k^{\mathrm{TX}} - 2T\rho, \quad (11)$$

  it is put in the SU category. Intuitively, p_k will become a PU packet in the next period. If the eNB can also send p_k out as soon as possible, it can reduce the amount of PU data in the following period (and decrease the packet dropping ratio accordingly). The length of u_i's SU data is denoted by L_i^{SU}.
- *QoS-demanded (QD) data:* When a communication system is said to be BIBO (bounded-input, bounded-output) stable, its output will be bounded in amplitude by giving a bounded input [26]. Obviously, an LTE-A system must be BIBO stable, because the demands of flows are finite (i.e., bounded input), so the eNB will not spend an infinite bandwidth for data transmissions (i.e., bounded output). Based on this observation, Piro et al. [27] show that the amount of data that must be sent in the x-th period in order to guarantee the QoS demand of a GBR UE can be calculated by the following equation:

$$L_i^{\mathrm{QD}}(x) = L_i(x) + \sum\nolimits_{y=0}^{M} \lambda(y) L_i(x-y) - \sum\nolimits_{y=1}^{M} \lambda(y) L_i(x-y+1), \quad (12)$$

where the function $\lambda(y)$ satisfies two conditions:

$$0 \leq \lambda(y) \leq 1, \text{ for } y = 0, 1, 2, \cdots,$$
$$\lambda(y) \leq \lambda(y-1), \text{ for } y \geq 2 \text{ with } \lambda(y) \in \mathbb{R}. \quad (13)$$

One example is to set $\lambda(0) = 0$ and $\lambda(y) = 1/2^{y-1}$ for $y \geq 1$. In other words, the amount of QD data depends on the effective queued data $L_i(x)$ in the current period and also those in previous periods (i.e., both $L_i(x-y)$ and $L_i(x-y+1)$). In Eq. (12), we suggest setting

$$M = \left\lfloor \frac{\min_{p_k \in u_i} t_k^{\mathrm{DB}}}{T\rho} \right\rfloor. \quad (14)$$

In this way, the eNB can send packets within their delay budgets. For the sake of consistency, the length of u_i's (current) QD data is denoted by L_i^{QD}.

Fig. 4 illustrates the classification of the downlink data for each GBR UE, where the following inequality must hold:

$$L_i^{\mathrm{PU}} + L_i^{\mathrm{SU}} \leq L_i^{\mathrm{QD}} \leq L_i. \quad (15)$$

In other words, both PU and SU data will be a subset of QD data. Then, Lemma 3 analyzes the time complexity of step 3.

Lemma 3. Given n_U UEs, it spends $O(MLn_U)$ time to carry out step 3, where $L = \max_{u_i \in \hat{U}} L_i$ and $\hat{U}$ is the set of all UEs.

Proof: Let us consider the case of one GBR UE u_i. In step 3, we first use Eqs. (10) and (11) to check whether each of u_i's packets belongs to PU and SU categories, respectively. The above check will take $O(L_i)$ time. Based on the observation in Eq. (15), PU and SU data must be parts of QD data. Thus, we can just check whether residual packets belong to the QD category. However, the worst case occurs when none of packets are PU or SU. In this case, we still have to check every packet by Eq. (12), which is proven to spend $O(M)$ time [27]. Thus, it takes time of $O(L_i) + L_i \times O(M) = O(ML_i)$ to do calculation for each UE. Since there are n_U UEs in $\hat{U}$, the time complexity of step 3 is thus $n_U \times O(M \times \max_{u_i \in \hat{U}} L_i) = O(MLn_U)$. □

4.4 Set ABS Ratios

Based on the knowledge of data rates supported by different eNBs and the downlink data owned by each GBR UE, in this step we compute the ABS ration γ with two objectives: 1) increasing the total throughput and 2) preventing the GBR UEs in macro-cells from starvation. From the flowchart in Fig. 3, there are three policies used to decide γ, which are discussed below. After that, we compute an indicator β according to these policies for the ABS placement in the next step.

4.4.1 Policy 1–Set γ to Zero

In this policy, we check whether each small-cell eNB can satisfy the traffic demands of its UEs in a period that contains merely OSs by

$$\sum\nolimits_{e_j \in \hat{E}_p} D_j^{\mathbf{O}} \times T \geq \sum\nolimits_{e_j \in \hat{E}_p} \sum\nolimits_{u_i \in \hat{U}_j} L_i. \quad (16)$$

Specifically, $D_j^{\mathbf{O}}$ indicates the number of data bits that a small-cell eNB can send in one OS, so the left term of Eq. (16) is the amount of data transmissions supported by all small-cell eNBs in a period which comprises only OSs. On the other hand, because L_i is the length of the effective queued data of a UE, the right term of Eq. (16) gives the amount of traffic demands of all small-cell UEs in the current period. Then, Theorem 2 proves that the optimal value of the ABS ratio γ must be zero when every small-cell eNB meets the condition of Eq. (16).

Theorem 2. If every small-cell eNB passes the check in policy 1, the optimal solution to maximizing the overall throughput is to set γ to zero.

Proof: For a small-cell, using ABS can increase the throughput because the signal quality of its UEs improves. Thus, when a small-cell eNB satisfies the condition of Eq. (16), it implies that the small-cell eNB has sufficient resources to send out the downlink data of all UEs even without ABS. On the other hand, using ABS will degrade the throughput of a macro-cell, because the macro-cell eNB cannot transmit user data in each ABS. Therefore, the optimal solution to maximizing the total throughput is to set $\gamma = 0$, when all small-cell eNBs pass the check in policy 1. □

4.4.2 Policy 2–Decide γ in an Increasing Manner

If some small-cell eNBs fail to pass the check in policy 1, we have to use a number of ABSs to increase their throughput. Policy 2 is adopted when $R_p^{\mathbf{A}} - R_p^{\mathbf{O}} < R_m^{\mathbf{O}}$. According to Theorem 1, adding more ABSs will decrease the overall throughput if $R_p^{\mathbf{A}} - R_p^{\mathbf{O}} < R_m^{\mathbf{O}}$. However, parts of the GBR UEs in small-cells may suffer from a high data loss rate due to their bad channel quality if no ABS is used. In this case, one better solution is to set a small ABS ratio γ, so as to keep the total throughput high while reducing data loss of the GBR UEs in small-cells.

Based on the above observation, starting from no ABS, we iteratively add one ABS until all PU data and a portion of SU data of the GBR UEs in small-cells can be sent out, on the condition that the average data loss rate of the GBR UEs in the macro-cell keeps below threshold δ. In this way, we can reduce the amount of data loss for the GBR UEs in each cell. Specifically, let N_A be the number of ABSs in the current period, which is initially set to zero. Then, we check whether Eq. (17) holds:

$$\sum\nolimits_{e_j \in \hat{E}_p} D_j^{\mathbf{G},\mathbf{A}} \times N_A + D_j^{\mathbf{G},\mathbf{O}} \times (T - N_A)$$
$$\geq \sum\nolimits_{e_j \in \hat{E}_p} \sum\nolimits_{u_i \in \hat{U}_j} L_i^{\mathrm{PU}} + \theta L_i^{\mathrm{SU}}, \quad (17)$$

where $\theta \in [0, 1]$ is a ratio of SU data to be sent in order to reduce the amount of PU data in the next period. In Eq. (17), since $D_j^{\mathbf{G},\mathbf{A}}$ and $D_j^{\mathbf{G},\mathbf{O}}$ indicate the amount of GBR data that a small-cell eNB can send in one ABS and OS, respectively, the left term gives the total amount of GBR data transmitted by the eNB when there are N_A ABSs in a period. On the other hand, the right term means that each GBR UE in small-cells can receive its PU data and a fraction θ of SU data to avoid packet expiration. The violation of Eq. (17) implies that the number of ABSs (i.e., N_A) is not enough to ward off data loss of the GBR UEs in small-cells. In this case, we iteratively increase N_A by one, until any of the following three conditions holds: 1) the condition of Eq. (17) is satisfied, 2) $N_A = T$ (i.e., all subframes in the current period are set as ABSs), or 3) the following condition is violated:

$$D_m^{\mathbf{G},\mathbf{O}} \times (T - N_A) \geq \sum\nolimits_{u_i \in \hat{U}_m^{\mathbf{G}}} (1 - \delta) L_i^{\mathrm{PU}}. \quad (18)$$

Specifically, since the macro-cell eNB is able to transmit user data only in $(T - N_A)$ OSs, Eq. (18) checks whether at least a

ratio $(1-\delta)$ of PU data owned by the GBR UEs in the macro-cell can be sent out. Then, the ABS ratio γ will be set to N_A/T. Theorem 3 shows that we can minimize the amount of data loss for the GBR UEs in small-cells by policy 2, on the premise that the GBR UEs in the macro-cell are not starved (i.e., their average data loss rate will not exceed threshold δ).

Theorem 3. By using policy 2, we can find the optimal γ value to minimize the amount of data loss for the GBR UEs in small-cells, under the condition that the average data loss rate of the GBR UEs in the macro-cell is below δ.

Proof: Starting from zero ABS, in policy 2 we iteratively add one ABS and check whether Eq. (17) holds. If so, the GBR UEs in small-cells will never incur packet loss, because all of their PU data and a ratio θ of SU data can be sent out in the current period. The only two cases that ABSs cannot be added any more in policy 2 are 1) N_A reaches T and 2) Eq. (18) cannot hold. For the first case, since all subframes are set as ABSs, the throughput of small-cells has been maximized, thereby minimizing the amount of data loss for the GBR UEs in small-cells. For the second case, Eq. (18) is used to guarantee that the average data loss rate of the GBR UEs in the macro-cell is below threshold δ. In this case, the value of N_A is optimal in the sense that it can minimize the amount of data loss for the GBR UEs in small-cells, under the condition that the average data loss rate of the GBR UEs in the macro-cell keeps below δ. □

4.4.3 Policy 3–Decide γ in a Decreasing Manner

This policy is complementary to policy 2, which is adopted when $R_p^{\mathbf{A}} - R_p^{\mathbf{O}} \geq R_m^{\mathbf{O}}$. In this case, adding more ABSs could help increase the overall throughput based on the observation in Theorem 1. Therefore, we calculate the minimum value of N_A to support QoS for the GBR UEs in small-cells. If necessary, N_A is decreased to keep the data loss rate of the GBR UEs in the macro-cell below threshold δ.

To do so, we estimate the amount of increase in the throughput of the GBR UEs in small-cells with one ABS as follows:

$$D_{\text{IN}} = \sum\nolimits_{e_j \in \hat{E}_p} D_j^{\mathbf{G},\mathbf{A}} - D_j^{\mathbf{G},\mathbf{O}}. \tag{19}$$

Then, the number of ABSs used to satisfy QoS demands of the GBR UEs in small-cells is calculated by

$$\xi = \frac{\sum_{e_j \in \hat{E}_p} \sum_{u_i \in \hat{U}_j^{\mathbf{G}}} L_i^{\text{QD}} - \sum_{e_j \in \hat{E}_p} D_j^{\mathbf{G},\mathbf{O}} \times T}{D_{\text{IN}}},$$
$$N_A = \min\{\xi, T\}. \tag{20}$$

Here, the term $(\sum_{e_j \in \hat{E}_p} \sum_{u_i \in \hat{U}_j^{\mathbf{G}}} L_i^{\text{QD}})$ gives the amount of QD data owned by all GBR UEs in small-cells, and the term $(\sum_{e_j \in \hat{E}_p} D_j^{\mathbf{G},\mathbf{O}} \times T)$ indicates the amount of GBR data sent by small-cell eNBs without ABS. Therefore, their difference divided by D_{IN} will be the number of required ABSs. Since a period contains T subframes, the constraint of $N_A \leq T$ must hold in Eq. (20). However, if the condition of Eq. (18) is violated, we iteratively decrease N_A by one until Eq. (18) is satisfied, so as to avoid starving the GBR UEs in the macro-cell. Theorem 4 shows that we can support QoS for the GBR UEs in small-cells by policy 3.

Theorem 4. Small-cell eNBs can send out as more QD data of their GBR UEs as possible by policy 3, on the premise that the average data loss rate of the GBR UEs in the macro-cell is below δ.

Proof: In policy 3, we adopt Eq. (20) to decide the initial value of N_A (i.e., the number of ABSs), and then iteratively decrease N_A by one if Eq. (18) cannot be satisfied. In particular, there are three cases to be discussed. The first case is that N_A is equal to $\left(\sum_{e_j \in \hat{E}_p} \sum_{u_i \in \hat{U}_j^{\mathbf{G}}} L_i^{\text{QD}} - \sum_{e_j \in \hat{E}_p} D_j^{\mathbf{G},\mathbf{O}} \times T\right) / D_{\text{IN}}$. In this case, we actually find out the enough number of ABSs for small-cell eNBs to send out all QD data of their GBR UEs (in other words, their QoS demands are satisfied). The second case is that N_A is equal to T. Since all subframes are set as ABSs, the throughput of small-cells has been maximized. Thus, small-cell eNBs can transmit the maximum amount of QD data. The last case occurs when N_A is below the value calculated by Eq. (20). In this case, the number of ABSs is reduced to prevent the GBR UEs in the macro-cell from starvation. Therefore, we still find out the best γ value to let small-cell eNBs send out as more QD data of their GBR UEs as possible, while keeping the average data loss rate of the GBR UEs in the macro-cell below δ. □

4.4.4 Compute indicator β

Both policies 2 and 3 result in a positive ABS ratio γ, which makes a macro-cell eNB insert ABSs in the current period. To facilitate the ABS placement procedure in Section 4.5, a *PU-relation indicator* β is computed based on the amount of PU data in the macro-cell and its subordinate small-cells:

- $\beta = 0$ if $\sum_{u_i \in \hat{U}_m^{\mathbf{G}}} L_i^{\text{PU}} = 0$ (i.e., there is no PU data in the macro-cell).
- $\beta = 1$ if $\sum_{e_j \in \hat{E}_p} \sum_{u_i \in \hat{U}_j^{\mathbf{G}}} L_i^{\text{PU}} = 0$ (i.e., there is no PU data in small-cells).
- $\beta = 2$ if $\sum_{u_i \in \hat{U}_m^{\mathbf{G}}} L_i^{\text{PU}} \leq \sum_{e_j \in \hat{E}_p} \sum_{u_i \in \hat{U}_j^{\mathbf{G}}} L_i^{\text{PU}}$ (i.e., the macro-cell has no more PU data than its small-cells).
- $\beta = 3$ otherwise.

Then, Lemma 4 analyzes the time complexity of step 4.

Lemma 4. The time complexity of step 4 is $O(T(n_{Ep} + n_U))$, where n_{Ep} is the number of small-cell eNBs and n_U is the number of UEs.

Proof: In step 4, three policies are used to decide the ABS ratio. The core of policy 1 is Eq. (16), which spends time of $O(n_{Ep} + n_{Up})$. For policy 2, it uses both Eqs. (17) and (18) to do the check in every iteration. Eq. (17) takes time of $O(n_{Ep} + n_{Up})$ and Eq. (18) spends time of $O(n_{Um})$, where n_{Um} denotes the number of macro-cell UEs. The worst case in policy 2 occurs when N_A is iteratively increased from zero to T. Thus, it consumes time of $T \times (O(n_{Ep} + n_{Up}) + O(n_{Um})) = O(T(n_{Ep} + n_U))$ to do the calculation in policy 2. Similarly, we adopt both Eqs. (18) and (20) in policy 3 to do the check in every iteration, where Eq. (20) takes time of $O(n_{Ep} + n_{Up})$. The worst case in policy 3 occurs when N_A is initially set to T and then iteratively decreased by one, until N_A becomes zero. Thus, the calculation in policy 3 takes time of $T \times (O(n_{Um}) + O(n_{Ep} + n_{Up})) = O(T(n_{Ep} + n_U))$. Since the calculation of indicator β can be embedded into the above operations and the execution of policies 2 and 3 is mutual exclusive (i.e., we only pick one of these two policies to execute), the overall time complexity of step 4 is thus $O(n_{Ep} + n_{Up}) + \max\{O(T(n_{Ep} + n_U)), O(T(n_{Ep} + n_U))\} = O(T(n_{Ep} + n_U))$. □

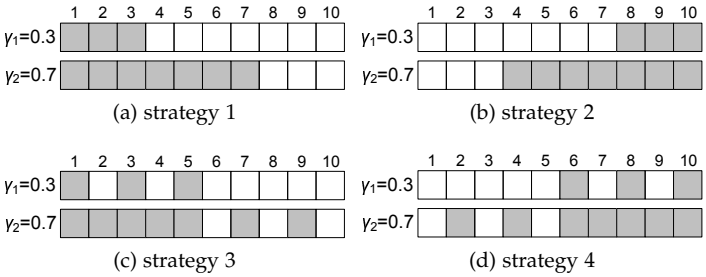


Fig. 5: Four strategies to place ABSs.

4.5 Decide ABS Placement

Finally, we decide how to place ABSs for each macro-cell. Evidently this step will be skipped if all macro-cells have no ABS, as shown by the flowchart in Fig. 3. Let us consider a general situation where two macro-cells cover the same small-cells and both of them have positive ABS ratios. Suppose that their PU-relation indicators are β_1 and β_2. Then, there are four strategies proposed to efficiently place ABSs.

Strategy 1: $\max\{\beta_1, \beta_2\} = 0$. Since both macro-cells have no PU data, ABSs are placed from the beginning of the period in a consecutive manner. In this way, small-cell eNBs can send PU data as soon as possible. Fig. 5(a) gives an example, where two macro-cell eNBs e_{m1} and e_{m2} have ABS ratios of 0.3 and 0.7, respectively. Then, e_{m1} places ABSs in the first three subframes, and e_{m2} places ABSs in the first seven subframes.

Strategy 2: $\max\{\beta_1, \beta_2\} = 1$. This strategy is opposite to the previous strategy, where none of small-cells has PU data but at least one macro-cell does. Thus, ABSs are placed from the end of the period in a consecutive manner to let macro-cell eNBs send PU data as soon as possible. Fig. 5(b) shows an example, where e_{m1} places ABSs in the last three subframes and e_{m2} places ABSs in the last seven subframes.

Strategy 3: $\max\{\beta_1, \beta_2\} = 2$. Each macro-cell has no more PU data than its small-cells. Thus, we place ABSs from the beginning of the period in an interlaced manner. Fig. 5(c) presents an example, where e_{m1} places ABSs in subframes 1, 3, and 5, and e_{m2} places ABSs in subframes 1, 3, 5, 7, 9, 2, and 4. Therefore, small-cell UEs will not be interfered in subframes 1, 3, and 5. In addition, a macro-cell eNB can send out PU data in the early stage of the period to reduce data loss for its served GBR UEs.

Strategy 4: $\max\{\beta_1, \beta_2\} = 3$. At least one macro-cell has more PU data than its small-cells. To prevent this macro-cell eNB from discarding many PU data, ABSs are placed from the end of the period in an interlaced manner, as illustrated in Fig. 5(d). Specifically, e_{m1} places ABSs in subframes 10, 8, and 6, and e_{m2} then places ABSs in subframes 10, 8, 6, 4, 2, 9, and 7.

For other situations (e.g., a macro-cell has no shared small-cells with others, or either of two coupling macro-cells has a zero ABS ratio), the macro-cell eNB can follow the above four strategies to place ABSs. Lemma 5 then gives an analysis on the time complexity of step 5.

Lemma 5. Running step 5 takes $O(Tn_{Em})$ time, where n_{Em} is the number of macro-cell eNBs.

Proof: In step 5, a macro-cell eNB will decide whether each subframe is ABS based on one of the four strategies. Since there are n_{Em} macro-cell eNBs and each period contains T subframes, the time complexity of step 5 is thus $O(Tn_{Em})$. □

4.6 Discussion

The idea behind our ECO algorithm is to estimate the amount of data transmissions supported by macro-cell and small-cell eNBs in each ABS and OS. According to the relationship between $(R_p^{\mathbf{A}} - R_p^{\mathbf{O}})$ and $R_m^{\mathbf{O}}$, ECO decides the ABS ratio γ with the aim of increasing the total throughput and avoiding starving the GBR UEs in a macro-cell. Then, ABSs are arranged in coupling macro-cells to send out urgent data as soon as possible, so as to reduce data loss of GBR flows.

At first glance, the proposed ECO algorithm has a similar flowchart with our previous DA3 method [16]. However, there are four differences between ECO and DA3 in essence. First of all, the design of ECO considers the case of coupling macro-cells, whereas DA3 assumes one single macro-cell. Thus, the evaluation of not only channel quality of UEs but also data transmissions supported by eNBs (i.e., capacity) in ECO and DA3 is different. In fact, ECO can calculate more accurate signal quality for UEs and also capacities of eNBs than DA3. Second, ECO addresses SU data of GBR UEs, which is ignored by DA3. Therefore, ECO allows some eNBs to send out SU data when they have enough resources. In this way, ECO can alleviate traffic loads of eNBs and also reduce packet loss of urgent data, as compared with DA3. Third, ECO takes different policies with DA3 in deciding the ABS ratio due to the consideration of SU data and coupling macro-cells. Besides, a PU-relation indicator β is calculated based on the amount of PU data in different cells to facilitate the ABS placement. Finally, ECO deals with the ABS placement problem (but DA3 does not), which adopts four placement strategies according to the PU-relation indicators of coupling macro-cells to reduce packet latency of GBR flows. The above designs distinguishes the ECO algorithm from the previous DA3 method. Moreover, the experimental results in Section 5 also show that ECO outperforms DA3 in terms of both throughput and data loss of GBR flows. Theorem 5 then analyzes the time complexity of the ECO algorithm.

Theorem 5. Given n_U UEs, n_E eNBs, and n_C subchannels, the ECO algorithm takes time of $O(n_U(n_E + n_C + ML + T))$, where L is the maximum queue length of UEs.

Proof: According to Lemmas 1, 2, 3, 4, and 5, the overall time complexity of ECO is $O(n_U(n_E + n_C)) + O(n_{Ep}n_{Up} + n_U) + O(MLn_U) + O(T(n_{Ep} + n_U)) + O(Tn_{Em})$. Because $n_{Ep} < n_E$, $n_{Em} < n_E$, $n_{Up} \leq n_U$, and $n_E < n_U$ (i.e., the number of UEs is usually larger than the number of eNBs), we can simplify the above time complexity to $O(n_U(n_E + n_C + ML + T))$. □

Generally speaking, the overlapping area of two macro-cells occupy merely a small portion of each of these macro-cells [28]. It can be expected that the most common situation for coupling macro-cells is that two macro-cells cover the same small-cells. Another case, though may rarely happen, is that three or more macro-cells are coupling with each other (i.e., covering the same small-cells). Consequently, we make some extensions to the ECO algorithm to deal with this special case. Suppose that a set $\hat{E}_C \subseteq \hat{E}$ of macro-cells are coupling with each other in a component subnetwork, where $|\hat{E}_C| \geq 3$. Then, each step in ECO is modified as follows:

- *Step 1 (evaluate channel quality):* For a UE u_i served in a small-cell (with eNB e_j) that is covered by the macro-cells whose eNBs are in $\hat{E}_C$, u_i is considered to be in an OS if every eNB in $\hat{E}_C$ is using an OS. Otherwise,

TABLE 3: Simulation parameters.

parameter	value
transmission-related parameters:	
channel bandwidth	5 MHz (i.e., 25 PRBs per subframe)
path loss	macro-cell: $128.1 + 37.6 \log \mathrm{dist}(\mathrm{eNB}, \mathrm{UE})$
	small-cell: $140.7 + 36.7 \log \mathrm{dist}(\mathrm{eNB}, \mathrm{UE})$
	$\mathrm{dist}(\mathrm{eNB}, \mathrm{UE})$ is measured in kilometers.
penetration loss	10 dB
shadowing fading	log-normal distribution with 0 dB mean and 8 dB standard deviation
fast fading	Jakes model (for Rayleigh fading)
eNB-related parameters:	
frame structure	frequency division duplexing (FDD)
MCS	referring to Table 2
transmitted power	macro-cell: 46 dBm, small-cell: 30 dBm
cell range	macro-cell: 1 km, small-cell: 150 m
PRB scheduling	M-LWDF
UE-related parameters:	
the number of UEs	135, 180, 225, 270, 315, 360
traffic flow	GBR flow: H.264 video (242 kbps)
	non-GBR flow: constant-bit-rate (12 kbps)
distribution of UEs	uniform distribution: 50% of UEs reside in small-cells
	hotspot distribution: 75% of UEs reside in small-cells

u_i is in an ABS. For the latter case, we iteratively pick one eNB e_y in $\hat{E}_C$ and set $\tilde{\mathrm{P}}(u_i, e_y, c_k)$ to zero in Eq. (1) to compute $S_{i,j,k}$. Then, the SINR of u_i in an ABS (i.e., $S^{\mathbf{A}}_{i,j,k}$) will be the average of these $S_{i,j,k}$ values.

- *Step 2 (assess network condition):* The calculation of the maximum data rate $R^{\mathbf{A}}_p$ supported by small-cell eNBs for the ABS case in Eq. (5) needs to be changed. In particular, suppose that a small-cell eNB e_j is covered by the macro-cells whose eNBs are in $\hat{E}_C$. To decide the SINR of a UE u_i served by e_j in an ABS, we iteratively select one eNB e_y in $\hat{E}_C$ to use an ABS and let other eNBs in $\hat{E}_C$ use an OS. Then, the final value of $S^{\mathbf{A}}_{i,j}$ in Eq. (5) will be the average of these SINRs for u_i.
- *Step 3 (classify GBR data):* Since we only check the packets of each GBR UE in step 3, which is irrelevant to whether macro-cells are coupling or not, there is no need to modify this step.
- *Step 4 (set ABS ratio):* Each macro-cell eNB solely decides its ABS ratio by the three policies (without consulting other coupling macro-cells). Moreover, the indicator β is calculated based on the amount of PU data in the macro-cell and its subordinate small-cells. Thus, this step can keep the way it is now.
- *Step 5 (decide ABS placement):* Let β_y denote the PU-relation indicator of a macro-cell eNB e_y. Then, if the value of $\max_{e_y \in \hat{E}_C}\{\beta_y\}$ is 0, 1, 2, and 3, strategies 1, 2, 3, and 4 will be used to place ABSs (referring to Section 4.5), respectively.

In our simulations, we consider the general case where two macro-cells are coupling with each other.

5 Performance Evaluation

We adopt LTE-Sim to evaluate system performance, which is an open-source simulator written in C++ for modeling transmission behavior of LTE-A networks [29]. Table 3 lists simulation parameters and Fig. 2 gives the topology of cells. A number of UEs are placed within cells based on *uniform* or *hotspot* distributions. About 80% of UEs are GBR UEs, each with a video flow. Every other UE has a non-GBR constant-bit-rate flow. The number of subframes in a period (i.e., T) is set to 10.

TABLE 4: Improvement ratios for the total throughput by ECO.

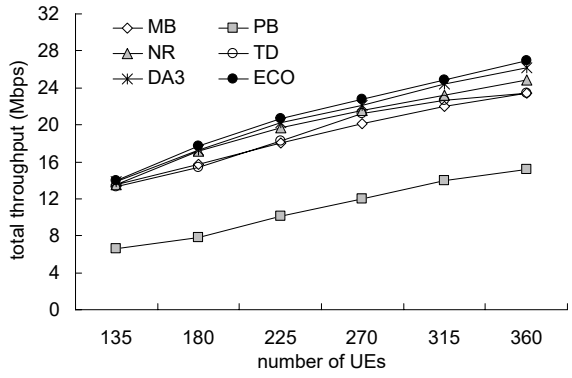
distribution of UEs	MB	PB	NR	TD	DA3
uniform distribution	12.54%	93.62%	11.20%	5.89%	2.44%
hotspot distribution	32.62%	10.77%	12.56%	6.27%	5.23%

Five methods are considered for comparison. Specifically, the *macro-cell benefited (MB) method* sets γ to 0.1 for macro-cells to observe the effect of using just few ABSs. On the contrary, the *small-cell benefited (PB) method* keeps γ to 0.9 to observe the effect of using many ABSs. In the *neutral ratio (NR) method*, γ is set to 0.5, so a half of subframes will be ABSs. Besides, we select two methods discussed in Section 2 that dynamically adjust γ based on network conditions. In particular, the *throughput difference (TD) method* [11] computes γ by referring to the difference of UEs' throughput in an ABS and OS. The *DA3 method* [16] estimates the amount of data transmitted by each eNB and computes γ to send out as many urgent data as possible. In the ECO algorithm, we set parameter δ to 0.3.

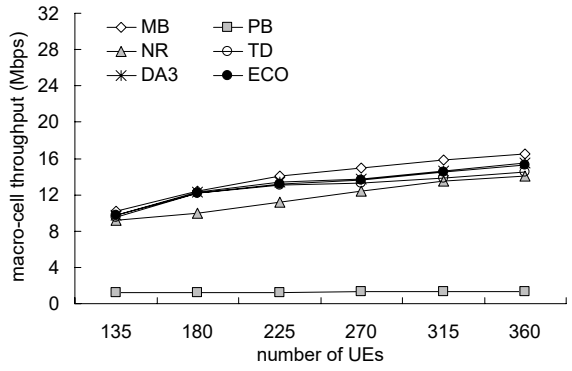
5.1 Comparison on Throughput

Fig. 6(a) presents the total throughput under the uniform distribution of UEs. Generally speaking, the total throughput increases as the number of UEs increases, because the PRB utilization also increases. Since a half of UEs locate in macro-cells, it is not a good idea to use too many ABSs (or around 50% of UEs will be starved). That is why the PB method always results in the lowest total throughput. On the other hand, the MB, NR, and TD methods perform similarly, because UEs are evenly distributed over macro-cells and small-cells. Both DA3 and ECO methods raise the throughput of small-cells on the prerequisite that the throughput of macro-cells can be still kept, so they have higher total throughput than others. Since the ECO algorithm seeks to satisfy QoS demands of the GBR UEs in small-cells when the network condition allows, it can further improve the total throughput as compared with the DA3 method.

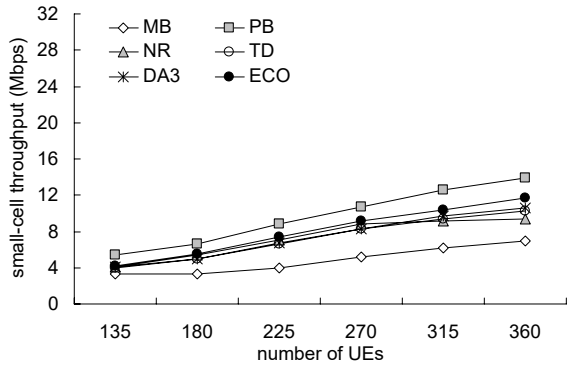
Let us take a look at the throughput of macro-cells and small-cells, as shown in Fig. 6(b) and (c), respectively. In particular, the MB method sets γ to 0.1, so it keeps the highest throughput of macro-cells but results in the lowest throughput of small-cells (due to high signal interference from macro-cell eNBs). On the contrary, the PB method sets γ to 0.9,



(a) total throughput

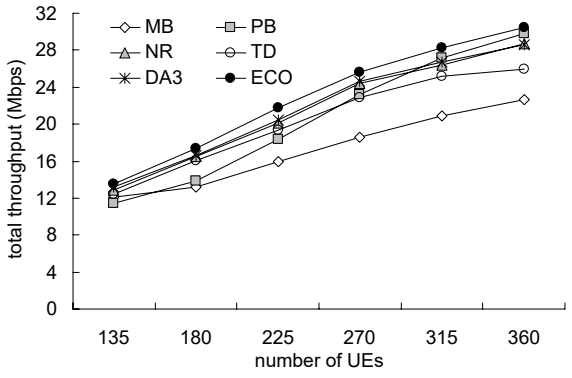


(b) macro-cell throughput

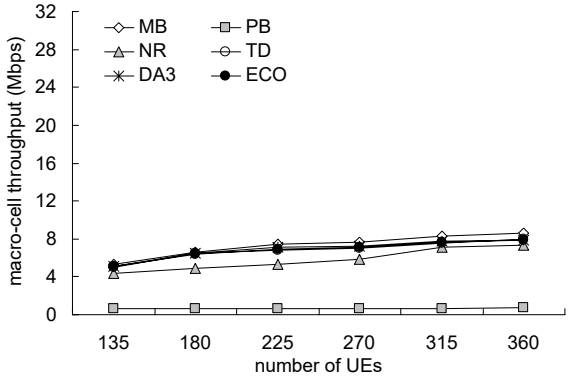


(c) small-cell throughput

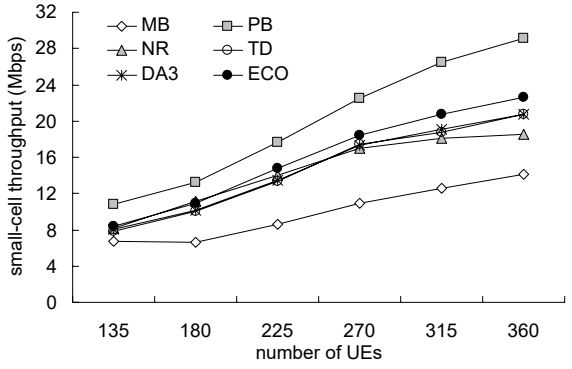
Fig. 6: Comparison on throughput under the uniform distribution of UEs.



(a) total throughput



(b) macro-cell throughput



(c) small-cell throughput

Fig. 7: Comparison on throughput under the hotspot distribution of UEs.

so it significantly improves the throughput of small-cells but almost has no throughput in macro-cells (i.e., starvation). Since UEs are uniformly distributed in cells, the NR method can balance data transmissions in macro-cells and small-cells by setting γ to 0.5, as compared with the MB and PB methods. On the other hand, all dynamic methods (i.e., the TD, DA3, and ECO methods) have higher throughput in different cells than the NR method, which verifies the effectiveness of adaptively adjusting the ABS ratio. Although the throughput of macro-cells by these three methods is pretty close, the ECO algorithm can flexibly increase the throughput of small-cells in contrast with both TD and DA3 methods, thereby improving the total throughput.

Fig. 7(a) illustrates the total throughput under the hotspot distribution of UEs, where three quarters of UEs reside in small-cells. Comparing with Fig. 6(a), the total throughput of each method increases because the channel quality of more UEs is improved by using ABS. Since the PB method sets γ to 0.9, its increase in the total throughput is the most significant. Thanks to the three policies for deciding the ABS ratio (discussed in Section 4.4), our ECO algorithm can always have the highest total throughput among all methods.

Fig. 7(b) and (c) then give the throughput of macro-cells and small-cells under the hotspot distribution of UEs, respectively. In contrast to Fig. 6(b) and (c), the throughput of macro-cells decreases while the throughput of small-cells increases. The reason is that most of UEs locate in small-cells. In this case, using more ABSs (e.g., the PB method) can greatly increase the throughput of small-cells. However, macro-cell UEs are still starved in the PB method because of the large ABS ratio. That is why our ECO algorithm adopts Eq. (18) to keep transmission fairness for macro-cell UEs. Table 4 summarizes the improvement ratio for the total throughput by the ECO algorithm as compared with each of other methods.

5.2 Comparison on Data Loss of GBR Flows

Fig. 8(a) shows the total loss rate of GBR data under the uniform distribution of UEs. Specifically, the total loss rate

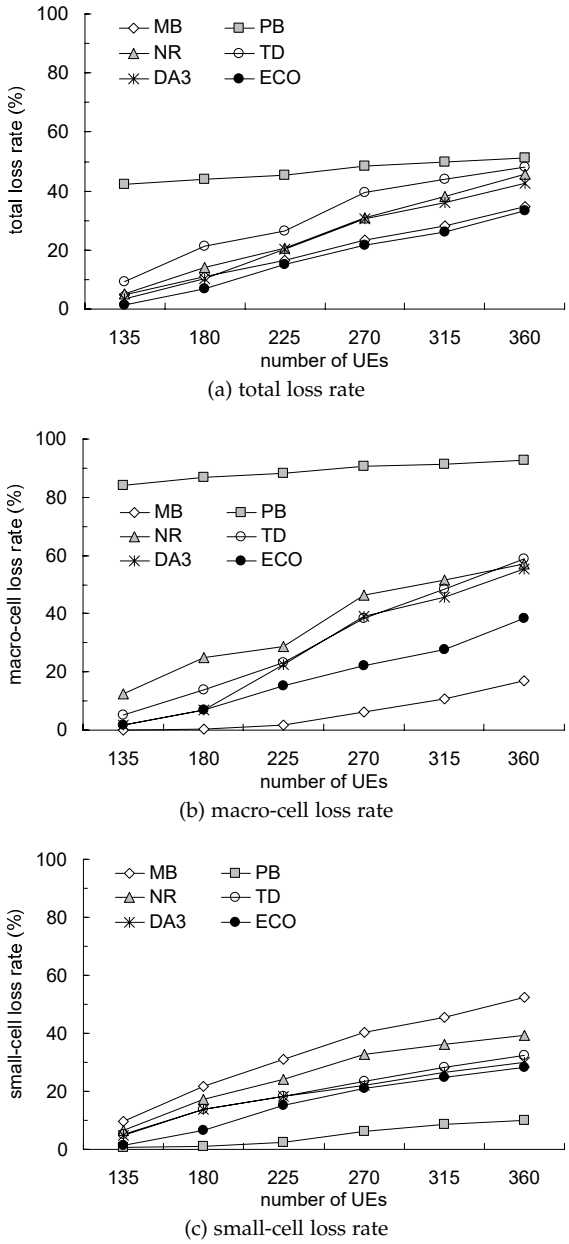
TABLE 5: Reduction ratios for data loss of GBR flows by ECO.

distribution of UEs	MB	PB	NR	TD	DA3
uniform distribution	11.64%	62.70%	44.48%	32.23%	26.97%
hotspot distribution	42.66%	35.76%	42.17%	24.68%	24.02%

increases as the number of UEs increases, because more UEs compete for the fixed number of PRBs. Since 50% of UEs locate in macro-cells, using more ABSs will cause more of their GBR packets to be dropped due to expiration. That is why the PB and MB methods have larger and smaller total loss rates than most of other methods, respectively. On the other hand, our ECO algorithm not only increases the total throughput (referring to Fig. 6(a)) but also deals with the ABS placement of coupling macro-cells to send PU data as soon as possible. Consequently, the ECO algorithm can keep the total loss rate lower, as compared with all other methods.

Then, Fig. 8(b) and (c) analyze the loss rates of GBR data in macro-cells and small-cells, respectively. For the PB method, its loss rate of macro-cells always keeps higher than 80%, though it has a pretty smaller loss rate of small-cells than others. The MB method is opposite to the PB method, where it has the smallest and largest loss rates of macro-cells and small-cells, respectively. By using the four strategies discussed in Section 4.5 to place ABSs based on the β indicator, the ECO algorithm reduces both loss rates of macro-cells and small-cells, as compared with the NR, TD, and DA3 methods. This experimental result demonstrates that our ECO algorithm can better support QoS for GBR flows than these methods.

In Fig. 9, we present the data loss rates of GBR flows under the hotspot distribution of UEs. Since the number of UEs in small-cells is three times greater than that in macro-cells, the total loss rate in the PB method decreases but that in the MB method increases. On the other hand, both total and small-cell loss rates of each other method slightly decrease, because small-cell eNBs can reduce the number of dropped GBR packets due to the increase in throughput (referring to Fig. 7(c)). By carefully placing ABSs with the consideration of PU data and coupling macro-cells, our ECO algorithm always has the smallest total loss rate of GBR data under the hotspot distribution of UEs. Table 5 summarizes the reduction ratio for data loss of GBR flows by the ECO algorithm as compared with each of other methods.



(a) total loss rate

(b) macro-cell loss rate

(c) small-cell loss rate

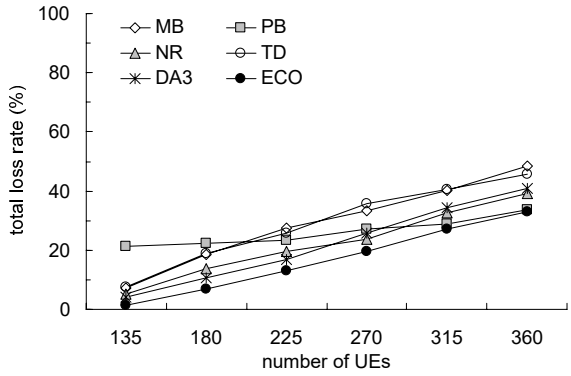
Fig. 8: Comparison on data loss of GBR flows under the uniform distribution of UEs.

5.3 Computation Time of Each Step in ECO

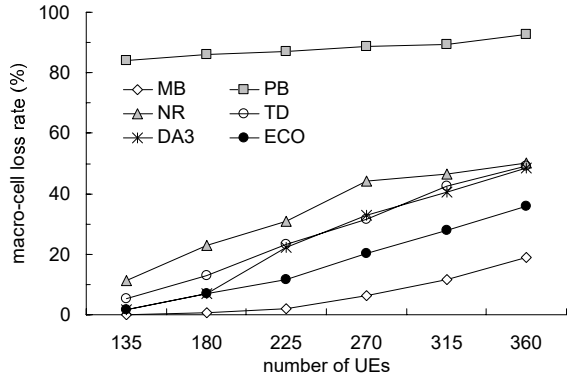
The proposed ECO algorithm comprises five steps to attain improvement in both throughput and data loss of GBR flows. In this section, we compare the computation time of each step in ECO. However, since the LTE-Sim program for our simulations contains many components other than ECO such as traffic generation in upper layers, packet scheduling in the MAC (medium access control) layer, and channel modeling in the physical layer, it is quite difficult to measure the computation time of each individual step in the ECO algorithm. Instead, we compare the (theoretical) computation time of each step based on the analysis of computational complexity in Lemmas 1–5.

In this experiment, we set the simulation parameters as below. According to the cell topology in Fig. 2, there are six macro-cell eNBs and eight pico-cell eNBs, so we have $n_{Em} = 6$, $n_{Ep} = 8$, and $n_E = 14$. Since the downlink channel has bandwidth of 5 MHz, there are 25 subchannels available (i.e., $n_C = 25$) [30]. The delay budget of video flows defined in the 3GPP specification [18] is 150 ms, so we can calculate that $M = \lfloor 150/10 \rfloor = 15$ by Eq. (14). The queue length (i.e., L) is set to 20 packets.

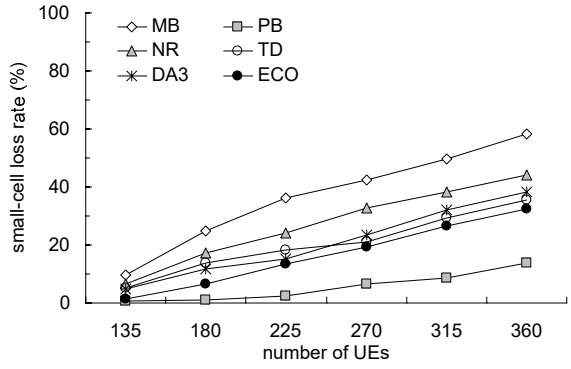
Fig. 10 compares the computation time of each step in the ECO algorithm, where it spends one unit time to conduct each operation. As can be seen, the total computation time of ECO grows linearly as the number of UEs increases. On the average, steps 1, 2, 3, 4, and 5 account for 22.09%, 3.96%, 67.96%, 5.85%, and 0.14% of ECO's computation time, respectively. Evidently both steps 1 and 3 dominate the computational complexity of ECO. However, step 1 is not particular to ECO, since UEs have to regularly report their CQIs to eNBs for reference in LTE-A [24]. Therefore, the computational complexity of ECO can be substantially reduced if we can efficiently decrease the computation time of step 3. One possible solution is to associate each packet in the queue with its current and next-round categories (e.g., an SU packet must become a PU packet in the next round). In this way, every time when step 3 is executed, we only need to check the newly incoming packets,



(a) total loss rate



(b) macro-cell loss rate



(c) small-cell loss rate

Fig. 9: Comparison on data loss of GBR flows under the hotspot distribution of UEs.

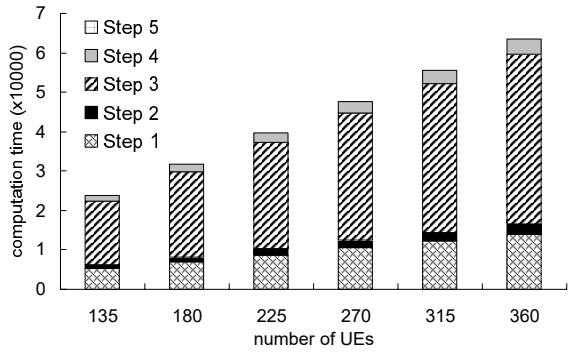


Fig. 10: Comparison on the computation time of each step in the ECO algorithm.

which avoids repeated calculation for those old packets that already have associations. Thus, the computation time of step 3 can be greatly decreased.

6 Conclusion and Future Work

To mitigate signal interference in heterogeneous networks, 3GPP proposes eICIC to make macro-cell eNBs stop sending user data in ABSs to better channel quality in small-cells. The management of ABSs is critical in deciding network performance. However, most studies neither consider a practical case where two macro-cells cover the same small-cells nor solve the ABS placement problem for the case of coupling macro-cells. Therefore, this paper develops the ECO algorithm for efficient ABS management by addressing these issues. It evaluates channel quality, assesses network condition, and classifies GBR data, so as to estimate the best ABS ratio to raise the total throughput on the premise that the GBR UEs in macro-cells are not starved. Based on their ratios and PU data, ECO jointly places ABSs for two coupling macro-cells with the aim of sending out urgent data as soon as possible. Through simulations by LTE-Sim, we show that the proposed ECO algorithm significantly increases the total throughput and also reduces data loss of GBR flows, as compared with the MB, PB, NR, TD, and DA3 methods, under both uniform and hotspot distributions of UEs.

In this paper, we aim to improve the performance of LTE-A eICIC with the consideration of coupling macro-cells. For 5G systems, mmWave (millimeter wave) massive MIMO (multiple-input multiple-output) communication is accomplished by using hybrid transceivers, and various beam-forming techniques are also developed [31]. A number of research efforts thus make use of beam-forming to alleviate the co-channel interference. For example, Mosleh et al. [32] propose an interference alignment scheme with limited feedback to mitigate the inter-cell interference and also a leakage-based coordinated beam-forming strategy to reduce the intra-cell interference. Deng et al. [33] consider using two-hop D2D (device-to-device) relaying to extend the coverage of mmWave signals. Both the issues of opportunistic relay selection and mmWave beam-forming are addressed with the aim of limiting the signaling overhead. Liu et al. [34] discuss how to exploit beam-forming to provide user-centric coordinated transmissions in 5G networks, especially for the case when UEs move across different cells. To support dense frequency reuse, Abohamra et al. [35] find the directions of the desired user and the interferers in a cell, and then use the beam-former to produce a beam towards the desired user and nulls in the direction of the interferers. Since beam-forming (along with multiple antenna signal processing) provides more flexibility for data transmissions, in the future work we will consider how to take advantage of its property together with some designs in our ECO algorithm to further increase the overall throughput, keep transmission fairness, and support QoS for real-time services in heterogeneous networks.

References

[1] Y.C. Wang and K.C. Chien, "EPS: energy-efficient pricing and resource scheduling in LTE-A heterogeneous networks," *IEEE Trans. Vehicular Technology*, vol. 67, no. 9, pp. 8832–8845, 2018.

[2] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: a comprehensive survey," *IEEE Comm. Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[3] Y.C. Wang and C.A. Chuang, "Efficient eNB deployment strategy for heterogeneous cells in 4G LTE systems," *Computer Networks*, vol. 79, no. 14, pp. 297–312, 2015.

[4] J. Pang, J. Wang, D. Wang, G. Shen, Q. Jiang, and J. Liu, "Optimized time-domain resource partitioning for enhanced inter-cell interference coordination in heterogeneous networks," *Proc. IEEE Wireless Comm. and Networking Conf.*, 2012, pp. 1613–1617.

[5] N. Trabelsi, L. Roullet, and A. Feki, "A generic framework for dynamic eICIC optimization in LTE heterogeneous networks," *Proc. IEEE Vehicular Technology Conf.*, 2014, pp. 1–6.
[6] S. Singh and J.G. Andrews, "Joint resource partitioning and offloading in heterogeneous cellular networks," *IEEE Trans. Wireless Comm.*, vol. 13, no. 2, pp. 888–901, 2014.
[7] J.A. Ayala-Romero, J.J. Alcaraz, J. Vales-Alonso, and E. Egea-Lopez, "Online optimization of interference coordination parameters in small cell networks," *IEEE Trans. Wireless Comm.*, vol. 16, no. 10, pp. 6635–6647, 2017.
[8] S. Vasudevan, R.N. Pupala, and K. Sivanesan, "Dynamic eICIC: a proactive strategy for improving spectral efficiencies of heterogeneous LTE cellular networks by leveraging user mobility and traffic dynamics," *IEEE Trans. Wireless Comm.*, vol. 12, no. 10, pp. 4956–4969, 2013.
[9] A. Daeinabi, K. Sandrasegaran, and P. Ghosal, "An enhanced intercell interference coordination scheme using fuzzy logic controller in LTE-advanced heterogeneous networks," *Proc. IEEE Int'l Symp. Wireless Personal Multimedia Comm.*, 2014, pp. 520–525.
[10] G. Bartoli, R. Fantacci, D. Marabissi, and M. Pucci, "Adaptive muting ratio in enhanced inter-cell interference coordination for LTE-A systems," *Proc. IEEE Int'l Wireless Comm. and Mobile Computing Conf.*, 2014, pp. 990–995.
[11] S.H. Lu, W.P. Lai, and L.C. Wang, "Time domain coordination for inter-cell interference reduction in LTE hierarchical cellular systems," *Proc. IEEE Int'l Conf. Heterogeneous Networking for Quality, Reliability, Security and Robustness*, 2014, pp. 51–55.
[12] A. Daeinabi, K. Sandrasegaran, and S. Barua, "A dynamic almost blank subframe scheme for video streaming traffic model in heterogeneous networks," *Proc. Int'l Conf. Electrical Engineering/Electronics, Computer, Telecomm. and Information Technology*, 2015, pp. 1–6.
[13] A. Argyriou, D. Kosmanos, and L. Tassiulas, "Joint time-domain resource partitioning, rate allocation, and video quality adaptation in heterogeneous cellular networks," *IEEE Trans. Multimedia*, vol. 17, no. 5, pp. 736–745, 2015.
[14] Y.C. Wang and C.C. Huang, "Efficient management of interference and power by jointly configuring ABS and DRX in LTE-A HetNets," *Computer Networks*, vol. 150, pp. 15–27, 2019.
[15] L. Tang, Y. Wei, W. Chen, and Q. Chen, "Delay-aware dynamic resource allocation and ABS configuration algorithm in HetNets based on Lyapunov optimization," *IEEE Access*, vol. 5, pp. 23764–23775, 2017.
[16] Y.C. Wang and S.T. Chen, "Delay-aware ABS adjustment to support QoS for real-time traffic in LTE-A HetNet," *IEEE Wireless Comm. Letters*, vol. 6, no. 5, pp. 590–593, 2017.
[17] 3GPP, "LTE; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 Application Protocol (X2AP)," ETSI TS 136 423 V15.2.0, 2018.
[18] ETSI, "Policy and charging control architecture (release 15)," 3GPP TS 23.203 V15.0.0, 2017.
[19] Y.C. Wang and D.R. Jhong, "Efficient allocation of LTE downlink spectral resource to improve fairness and throughput," *Int'l J. Comm. Systems*, vol. 30, no. 14, pp. 1–13, 2017.
[20] Y.C. Wang and S. Lee, "Small-cell planning in LTE HetNet to improve energy efficiency," *Int'l J. Comm. Systems*, vol. 31, no. 5, pp. 1–18, 2018.
[21] R. Giuliano and F. Mazzenga, "Exponential effective SINR approximations for OFDM/OFDMA-based cellular system planning," *IEEE Trans. Wireless Comm.*, vol. 8, no. 9, pp. 4434–4439, 2009.
[22] C. Mehlfuhrer, M. Wrulich, J.C. Ikuno, D. Bosanska, and M. Rupp, "Simulating the long term evolution physical layer," *Proc. European Signal Processing Conf.*, 2009, pp. 1471–1478.
[23] L.M. Surhone, M.T. Timpledon, and S.F. Marseken, *Shannon-Hartley Theorem*, VDM Publishing, 2010.
[24] ETSI, "Evolved universal terrestrial radio access (E-UTRA); physical layer procedures (release 14)," 3GPP TS 36.213 V14.1.0, 2016.
[25] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Comm. Magazine*, vol. 39, no. 2, pp. 150–154, 2001.
[26] K.J. Astrom and B. Wittenmark, *Computer Controlled Systems: Theory and Design*, Dover Publications, 2012.
[27] G. Piro, L.A. Grieco, G. Boggia, R. Fortuna, and P. Camarda, "Two-level downlink scheduling for real-time multimedia services in LTE networks," *IEEE Trans. Multimedia*, vol. 13, no. 5, pp. 1052–1065, 2011.
[28] M. Tayyab, X. Gelabert, and R. Jantti, "A survey on handover management: from LTE to NR," *IEEE Access*, vol. 7, pp. 118907–118930, 2019.
[29] G. Piro, L.A. Grieco, G. Boggia, F. Capozzi, and P. Camarda, "Simulating LTE cellular systems: an open-source framework," *IEEE Trans. Vehicular Technology*, vol. 60, no. 2, pp. 498–513, 2011.
[30] F. Capozzi, G. Piro, L. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: key design issues and a survey," *IEEE Comm. Surveys & Tutorials*, vol. 15, no. 2, pp. 678–700, 2013.
[31] I. Ahmed, H. Khammari, A. Shahid, A. Musa, K.S. Kim, E.D. Poorter, and I. Moerman, "A survey on hybrid beamforming techniques in 5G: architecture and system model perspectives," *IEEE Comm. Surveys & Tutorials*, vol. 20, no. 4, pp. 3060–3097, 2018.
[32] S. Mosleh, J.D. Ashdown, J.D. Matyjas, M.J. Medley, J. Zhang, and L. Liu, "Interference alignment for downlink multi-cell LTE-advanced systems with limited feedback," *IEEE Trans. Wireless Comm.*, vol. 15, no. 12, pp. 8107–8121, 2016.
[33] J. Deng, O. Tirkkonen, R. Freij-Hollanti, T. Chen, and N. Nikaein, "Resource allocation and interference management for opportunistic relaying in integrated mmWave/sub-6 GHz 5G networks," *IEEE Comm. Magazine*, vol. 55, no. 6, pp. 94–101, 2017.
[34] J. Liu, K. Au, A. Maaref, J. Luo, H. Baligh, H. Tong, A. Chassaigne, and J. Lorca, "Initial access, mobility, and user-centric multi-beam operation in 5G new radio," *IEEE Comm. Magazine*, vol. 56, no. 3, pp. 35–41, 2018.
[35] Y.A. Abohamra, M.R. Soleymani, and Y.R. Shayan, "Using beamforming for dense frequency reuse in 5G," *IEEE Access*, vol. 7, pp. 9181–9190, 2019.