Delay-aware ABS Adjustment to Support QoS for Real-time Traffic in LTE-A HetNet

You-Chiun Wang and Shih-Ting Chen

Abstract—To solve interference in a heterogeneous network, LTE-A uses *almost blank subframe (ABS)* to improve channel quality of picocells at the expense of macrocell's transmission. Many studies find the ABS ratio based on user number or traffic load. However, the issue of real-time (RT) packet delay is not well studied. We thus develop a *delay-aware ABS adjustment (DA3)* method to improve performance and lower packet dropping. DA3 estimates cells' capacity to decide the ratio, and analyzes queue content to alleviate packet discard. Simulation results show that DA3 increases cells' goodput while reducing RT packet dropping.

Index Terms—ABS, eICIC, HetNet, LTE-A, real-time traffic.

◆

1 INTRODUCTION

LTE-A offers broadband wireless service, and *heterogeneous network (HetNet)* is one feature, where various eNBs can cooperate in the same service area. Macrocell (MC) eNBs give large-scale coverage, and picocell (PC) eNBs enhance signals in hot spots. HetNet provides flexible eNB deployment [1] but worsens interference. Therefore, LTE-A uses *enhanced inter-cell interference coordination (eICIC)* to solve the problem. It allows cells to share the same band, and uses *almost blank subframe (ABS)* to reduce interference. In a period, some slots are selected as *ABS slots* where the MC eNB sends only low-power control signals. Thus, PC *user equipments (UEs)* can have better channel quality. *Further enhanced ICIC (FeICIC)* extends this method by using *reduced power ABS (RP-ABS)* to allow an MC eNB to send data at a lower power level [2].

The ABS ratio decides HetNet performance, as MC UEs receive no data in ABS slots. It attracts attention to find the ratio. Both [3], [4] assume that UEs have full buffers, and find the optimal ratio by brute-force search. The study [5] derives the HetNet's rate distribution by various ratios and backhaul capacity to find the best solution. However, the full-buffer assumption is impractical. The study [6] raises the ratio with a probability as the number of PC UEs grows, but it cannot handle the case where there are few MC UEs with large demand. Given an ABS ratio, the study [7] uses sum- and product-rate utilities to model throughput and fairness maximization, and finds the optimal value. However, it assumes that packets have no latency constraint. The study [8] uses 3 cases in each 8-slot period: 1 ABS, 2 ABSs, and 1 ABS + 1 low-power slot. It selects a suitable case based on network load. Obviously, there are too few choices. The study [9] uses fuzzy logic to get the ratio, whose inputs include the number of UEs, SINRs of PC UEs, and throughput of MC UEs. The study [10] finds UEs' throughput with and without interference. It adjusts the ratio iteratively by 12.5% to improve performance. However, [9], [10] do not consider packet delay.

Only few studies address RT traffic. The study [11] uses a genetic algorithm to allocate ABS slots, where the fitness function considers throughput, interference, and loss/delay of video flows. But, it has to repeat many iterations to get a stable result (e.g., more than 10 iterations), which incurs high computation overhead. Thus, we propose a *delay-aware ABS adjustment (DA3) method* to support QoS for RT traffic. To decide the ABS ratio, DA3 derives a network-status indicator τ to check the suitability of using ABS slots and estimates each cell's capacity. It also analyzes queued data with different urgent degrees to alleviate packet discard. Through simulations, we show that DA3 not only improves network goodput, but also reduces RT packet dropping, which verifies its effectiveness.

2 PROBLEM DEFINITION

We consider a HetNet with one MC containing some PCs. It is easy to extend the solution to a multi-MC HetNet, as each MC decides the ABS ratio γ on its own. Let $\mathcal{E}$, $\mathcal{E}_m$, and $\mathcal{E}_p$ be the sets of all, MC, and PC eNBs, respectively. We divide time into periods, each with T_S slots and T_L length. We use single-user MIMO for transmission, so *resource blocks (RBs)* are non-sharable. Let $\mathcal{U}_m$ and $\mathcal{U}_p$ be the sets of UEs linked to MC and PC eNBs, respectively. Each UE has either RT or non-RT traffic. Our problem asks how to find γ to increase goodput while reducing RT packet dropping. Note that the optimal solution to the first objective may be $\gamma \approx 1$, but it starves MC UEs. Hence, we define a lower bound $\varphi \in (0, 1)$ of *RT loss rate* for the MC to avoid such unfair situation.

3 THE PROPOSED DA3 METHOD

Fig. 1 gives DA3's flowchart. Phase 1 computes SINR of each UE. Phase 2 derives each cell's capacity and τ. Phase 3 analyzes queue content to find urgent data. Phase 4 finally decides γ. Below, we detail each phase and discuss DA3.

3.1 Phase 1–SINR Calculation

Since MC transmission is paused in an ABS slot, PC UEs are interfered only by other PC eNBs. Thus, the SINR value $\Gamma_i^{\mathbf{A}}$ of a UE u_i is

$$\Gamma_i^{\mathbf{A}} = \begin{cases} \frac{P_{i,j}}{\eta B + \sum_{e_a \in \mathcal{E}_p, e_a \neq e_j} I_{i,a}}, e_j \in \mathcal{E}_p & \text{if } u_i \in \mathcal{U}_p \\ 0 & \text{if } u_i \in \mathcal{U}_m, \end{cases} \quad (1)$$

Y. C. Wang and S. T. Chen are with the Department of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, 80424, Taiwan.
E-mail: ycwang@cse.nsysu.edu.tw; ago30425@gmail.com.

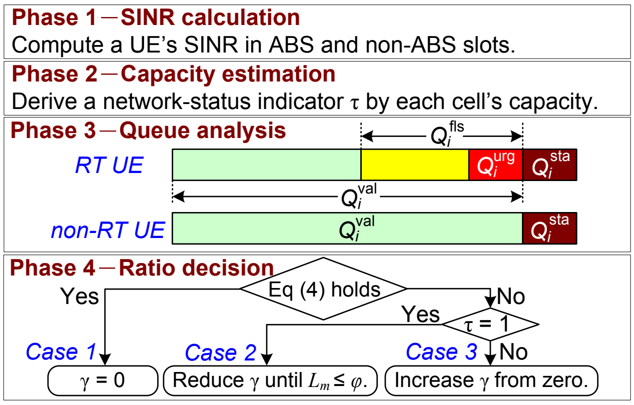


Fig. 1: The flowchart of our DA3 method.

where $P_{i,j}$ is u_i's received power from its eNB e_j, η is noise effect, which is the product of noise figure and noise spectral density, B is channel bandwidth, and $I_{i,a}$ is power of signal interference from another eNB e_a. In a non-ABS slot, u_i may be interfered by nearby eNBs, so we compute its SINR by

$$\Gamma_i^{\mathbf{O}} = \frac{P_{i,j}}{\eta B + \sum_{e_a \in \mathcal{E}, e_a \neq e_j} I_{i,a}}, \tag{2}$$

where $e_j \in \mathcal{E}_p$ if $u_i \in \mathcal{U}_p$ and $e_j \in \mathcal{E}_m$ if $u_i \in \mathcal{U}_m$. In Eqs. 1 and 2, $P_{i,j}$ depends on e_j's transmitted power and the distance from e_j to u_i. Typical transmitted power of PC and MC eNBs is respectively 30dBm and 46dBm [1], so the MC eNB can provide higher received power to its UEs than a PC eNB.

3.2 Phase 2–Capacity Estimation

We use the Shannon's theorem to estimate the capacity of each cell. Since all UEs in a cell share the same channel, we take their average SINR to be the cell's SINR. Let $\bar{\Gamma}_j^{\mathbf{A}}$ and $\bar{\Gamma}_j^{\mathbf{O}}$ be average SINR of UEs in a PC with eNB e_j in ABS and non-ABS slots, respectively. The capacity of that cell is $\xi_j^{\mathbf{A}} = B \lg(1 + \bar{\Gamma}_j^{\mathbf{A}})$ and $\xi_j^{\mathbf{O}} = B \lg(1 + \bar{\Gamma}_j^{\mathbf{O}})$ in ABS and non-ABS slots, respectively. Thus, the overall capacity of all PCs is $\xi_p^{\mathbf{A}} = \sum_{e_j \in \mathcal{E}_p} \xi_j^{\mathbf{A}}$ and $\xi_p^{\mathbf{O}} = \sum_{e_j \in \mathcal{E}_p} \xi_j^{\mathbf{O}}$ in ABS and non-ABS slots, respectively. The MC has no capacity in an ABS slot (i.e., $\xi_m^{\mathbf{A}} = 0$), and its capacity in a non-ABS slot is $\xi_m^{\mathbf{O}} = B \lg(1 + \bar{\Gamma}_m^{\mathbf{O}})$, where $\bar{\Gamma}_m^{\mathbf{O}}$ is average SINR of MC UEs in a non-ABS slot.

Then, we find the HetNet's capacity with an ABS ratio γ:

$$\begin{aligned}\xi &= \frac{(1-\gamma)T_L \cdot (\xi_p^{\mathbf{O}} + \xi_m^{\mathbf{O}}) + \gamma T_L \cdot (\xi_p^{\mathbf{A}} + \xi_m^{\mathbf{A}})}{T_L} \\ &= (\xi_p^{\mathbf{A}} - \xi_m^{\mathbf{O}} - \xi_p^{\mathbf{O}}) \cdot \gamma + \xi_p^{\mathbf{O}} + \xi_m^{\mathbf{O}}. \end{aligned} \tag{3}$$

Eq. (3) is a linear equation with one unknown γ, whose slope is $\xi_p^{\mathbf{A}} - \xi_m^{\mathbf{O}} - \xi_p^{\mathbf{O}}$. The slope decides the effect of ratio γ on ξ. When the slope is positive, ξ grows as we increase γ. Besides, ξ reduces as we enlarge γ when the slope is negative. By the slope, we define a *network-status indicator* $\tau = 1$ when $\xi_p^{\mathbf{A}} - \xi_m^{\mathbf{O}} - \xi_p^{\mathbf{O}} > 0$, or $\tau = 0$ otherwise. Here, we use Eq. (3) to observe how γ affects the theoretical capacity of the HetNet. When $\tau = 1$, it means that we can increase γ to improve capacity, and vice versa. Note that we need not consider the actual amount of traffic in Eq. (3), as it is reflected by the queue content of each UE, which we will analyze in phase 3.

We also estimate the amount of RT and non-RT data sent by an eNB. In LTE-A, each UE reports the *channel quality indicator (CQI)* to its eNB. Then, the standard [12] gives three tables to compute the amount of data that the UE can receive. The first table maps between the *modulation and coding scheme (MCS)* index and the CQI index for the UE. The second table maps between the *transport block size (TBS)* index and the MCS index. Then, given the TBS index and the number of RBs of that UE, we can find the amount of data carried by these RBs from the third table. Let us use a term $f(\alpha_i, n_i)$ to denote the number of bits received by a UE u_i via the above three-table mapping, where α_i is the CQI index of u_i and n_i is the number of RBs assigned to u_i. Then, each eNB e_j can transmit the amount of RT data in an ABS slot: $\zeta_j^{\mathbf{A},\mathrm{rt}} = \sum\{f(\alpha_i^{\mathbf{A}}, n_i) \,|\, \forall u_i \text{ is an RT UE and links to } e_j\}$, where $\alpha_i^{\mathbf{A}}$ is u_i's CQI in an ABS slot. Similarly, we can define $\zeta_j^{\mathbf{A},\mathrm{nrt}}$, $\zeta_j^{\mathbf{O},\mathrm{rt}}$, and $\zeta_j^{\mathbf{O},\mathrm{nrt}}$, which are the amount of non-RT data in an ABS slot, RT data in a non-ABS slot, and non-RT data in a non-ABS slot that e_j can send, respectively. We remark that n_i depends on the RB scheduler. One can refer to the RB scheduling result in the previous period to estimate n_i.

3.3 Phase 3–Queue Analysis

Fig. 1 shows our analysis on a UE's queue. For each RT UE u_i, we consider four cases:

Stale data (Q_i^{sta}): A packet q_x is in Q_i^{sta} if its delay d_x satisfies that $d_x + \sigma(q_x) > d_i^{\max}$, where $\sigma(q_x)$ is propagation time to send q_x and $d_i^{\max}$ is u_i's delay tolerant time. Obviously, these data should be removed to save bandwidth.

Valid data (Q_i^{val}): Except for Q_i^{sta}, all other packets can catch up their deadlines in theory.

Frame level scheduler (FLS) data (Q_i^{fls}): It is proven in [13] that FLS can estimate the necessary amount of RT data transmitted in each period to meet a UE's QoS demand. In the lth period, we compute the amount of FLS data by $Q_i^{\mathrm{fls}}(l) = Q_i^{\mathrm{val}}(l) + y_i(t) \sum_{t=2}^{\beta} (Q_i^{\mathrm{val}}(l-t+1) - Q_i^{\mathrm{val}}(l-t+2) - Q_i^{\mathrm{fls}}(l-t+1))$, where $Q_i^{\mathrm{val}}(l)$ is the amount of valid data in the lth period, $y_i(t)$ is a coefficient in $[0,1]$ where $y_i(t) \geq y_i(t+1)$, $\forall t \geq 2$, and β is the number of observing slots (e.g., $\beta = T_S$).

Urgent data (Q_i^{urg}): A packet $q_x \in Q_i^{\mathrm{fls}}$ is urgent if its delay satisfies that $d_x + \sigma(q_x) + T_L > d_i^{\max}$. In other words, if q_x will not be sent in the next period, it must expire.

For a non-RT UE, since it has loose delay requirement, we consider only Q_i^{sta} and Q_i^{val}, as shown in Fig. 1.

3.4 Phase 4–Ratio Decision

With the capacity and queue analysis, we compute the ABS ratio γ by three cases below.

Case 1: PC capacity $\geq$ valid queue length, that is,

$$T_S \sum\nolimits_{e_j \in \mathcal{E}_p} \zeta_j^{\mathbf{O},\mathrm{rt}} + \zeta_j^{\mathbf{O},\mathrm{nrt}} \geq \sum\nolimits_{u_i \in \mathcal{U}_p} Q_i^{\mathrm{val}}. \tag{4}$$

As PC eNBs have enough capacity to send out valid data of their UEs, there is no need to use ABS slots. So, we set $\gamma = 0$.

Case 2: $\tau = 1$. It is better to use ABS slots to improve goodput, but we should avoid starving MC UEs. Thus, we estimate the RT loss rate of MC with a given ratio γ by

$$L_m = \frac{\sum_{u_i \in \mathcal{U}_m^{\mathrm{rt}}} Q_i^{\mathrm{urg}} - (1-\gamma) T_S \cdot \zeta_j^{\mathbf{O},\mathrm{rt}}}{\sum_{u_i \in \mathcal{U}_m^{\mathrm{rt}}} Q_i^{\mathrm{urg}}}, \; e_j \in \mathcal{E}_m, \tag{5}$$

where $\mathcal{U}_m^{\mathrm{rt}}$ is the set of RT UEs in the MC. Here, $\sum_{u_i \in \mathcal{U}_m^{\mathrm{rt}}} Q_i^{\mathrm{urg}}$ is the amount of urgent data required by MC UEs, and $(1-\gamma) T_S \cdot \zeta_j^{\mathbf{O},\mathrm{rt}}$ is the amount of RT data that can be transmitted

TABLE 1: Effect of CQI and queue length on the ABS ratio γ.

UE	CQI	queue length
MC	CQI $\uparrow \Rightarrow \gamma \downarrow$	$Q_i^{\mathrm{urg}} \uparrow \Rightarrow \gamma \downarrow$
PC	CQI $\uparrow \Rightarrow \gamma \downarrow$	$Q_i^{\mathrm{val}} \uparrow, Q_i^{\mathrm{fls}} \uparrow \Rightarrow \gamma \uparrow$

by the MC eNB. By taking their difference, the numerator in Eq. (5) is the amount of urgent data that cannot be sent out before expiration (i.e., RT loss data). Then, we compute goodput improvement by a PC with one ABS slot:

$$\mu_p = \sum\nolimits_{e_j \in \mathcal{E}_p} \zeta_j^{\mathbf{A},\mathrm{rt}} + \zeta_j^{\mathbf{A},\mathrm{nrt}} - \zeta_j^{\mathbf{O},\mathrm{rt}} - \zeta_j^{\mathbf{O},\mathrm{nrt}}. \tag{6}$$

From Eq. (6), we estimate the number of ABS slots required to send out valid data of PC UEs:

$$t_p = \frac{\sum_{u_i \in \mathcal{U}_p} Q_i^{\mathrm{val}} - T_S \sum_{e_j \in \mathcal{E}_p} \zeta_j^{\mathbf{O},\mathrm{rt}} + \zeta_j^{\mathbf{O},\mathrm{nrt}}}{\mu_p}, \tag{7}$$

where the numerator is the difference between the amount of valid data of PC UEs and the capacity of PCs without ABS. By dividing it to μ_p, we can predict how many ABS slots are used to meet the demand of all PC UEs. Then, starting from t_p, we iteratively decrease it by one, until $L_m \leq \varphi$ (i.e., the MC's loss rate is met). Thus, the new ABS ratio is $\gamma = \lfloor t_p/T_S \rfloor$.

Case 3: $\tau = 0$. Using ABS may not improve goodput. However, as PC UEs queue lots of RT data, we should allocate ABS slots to send them. Starting from $t_p = 0$, we iteratively add it by one and set $\gamma = \lfloor t_p/T_S \rfloor$, until either $L_m \geq \varphi$ or

$$T_S((1-\gamma) \sum_{e_j \in \mathcal{E}_p} \zeta_j^{\mathbf{O},\mathrm{rt}} + \gamma \sum_{e_j \in \mathcal{E}_p} \zeta_j^{\mathbf{A},\mathrm{rt}}) \geq \delta \sum_{u_i \in \mathcal{U}_p^{\mathrm{rt}}} Q_i^{\mathrm{fls}}, \tag{8}$$

where $\mathcal{U}_p^{\mathrm{rt}}$ denotes RT UEs in PCs. Eq. (8) checks if the ABS ratio γ allows PC eNBs to have enough capacity (i.e., the left term) to send a δ ratio of RT data (i.e., the right term). Here, δ is a system-defined parameter, where $0 < \delta \leq 1$. Since Q_i^{fls} is the amount of data used to meet an RT UE u_i's QoS demand, Eq. (8) allows the PC eNB to send out at least $\delta \cdot Q_i^{\mathrm{fls}}$ amount of RT data to reduce u_i's packet dropping.

We remark on these cases. In case 1, HetNet already has enough capacity to meet all UEs' demand, so we set $\gamma = 0$ to get the maximum goodput. Case 2 implies that the maximum goodput occurs when we maximize PC goodput by enlarging γ. However, it may starve MC UEs. Thus, we find the number of ABS slots t_p for PC UEs to get the most data and iteratively decrease t_p until the MC's loss rate L_m is below φ. Case 3 is opposite to case 2, where the maximum goodput occurs as we maximize MC goodput by reducing γ. So, we gradually increase t_p to reduce RT packet dropping in PCs, under the premise of $L_m \leq \varphi$. This design considers QoS for RT traffic and distinguishes DA3 from existing work.

3.5 Discussion

We study the effect of CQI and queue length on γ, as given in Table 1. When MC UEs have larger CQIs, $\zeta_j^{\mathbf{O},\mathrm{rt}}$ grows. By Eq. (5), we reduce γ to lower the MC's loss rate L_m. Besides, increasing Q_i^{urg} enlarges L_m, so we prefer reducing γ. On the other hand, when PC UEs have larger CQIs, $\zeta_j^{\mathbf{O},\mathrm{rt}}$ and $\zeta_j^{\mathbf{O},\mathrm{nrt}}$ increase. By Eqs. 7 and 8, we should reduce γ. Intuitively, when PC UEs have good channel quality, using ABS may not much improve PC goodput but hurt MC goodput. Thus, it is better to lower γ to increase total goodput. By Eq. (7), enlarging Q_i^{val} increases the number of ABS slots t_p (i.e., γ increases). Also, increasing Q_i^{fls} makes it slower to meet the condition of Eq. (8).

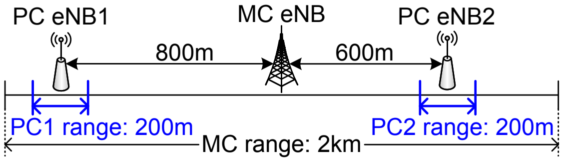


Fig. 2: The network topology used in the simulations.

Since we gradually increase t_p from zero in case 3 of phase 4, it results in a larger γ value. Theorem 1 analyzes the time complexity of DA3.

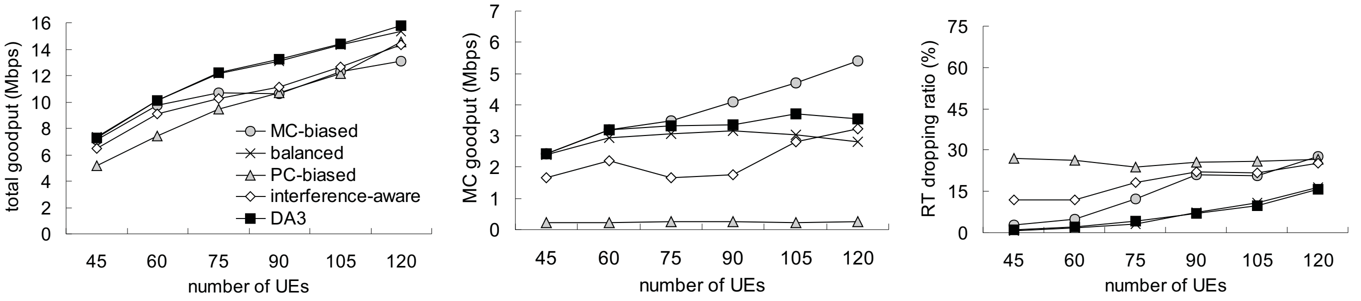
Theorem 1. Given h UEs and k eNBs, the computation complexity of the DA3 method is $O(hqT_S + k(h + T_S))$, where q is the maximum queue length.

Proof: Let h_m and h_p denote the number of MC and PC UEs, respectively. Phase 1 takes $O(hk)$ time to find SINR of each UE based on signal power of eNBs. Phase 2 decides τ by $\xi_p^{\mathbf{A}} - \xi_m^{\mathbf{O}} - \xi_p^{\mathbf{O}}$, where we find average SINR to get each variable. It spends $O(3h)$ time. Then, we sum up $f(\alpha_i, n_i)$ of UEs to get $\zeta_j^{\mathbf{A},\mathrm{rt}}$, $\zeta_j^{\mathbf{A},\mathrm{nrt}}$, $\zeta_j^{\mathbf{O},\mathrm{rt}}$, and $\zeta_j^{\mathbf{O},\mathrm{nrt}}$. To find $f(\alpha_i, n_i)$, we refer to the three tables in [12], where a table-mapping operation takes constant time. So, phase 2 requires time of $O(3h + 3h) = O(h)$. In phase 3, the worst case occurs as Q_i^{fls} has all packets. By [13], it takes $O(T_S)$ time to find $Q_i^{\mathrm{fls}}(l)$. As we check all packets of each UE, phase 3 takes $O(hqT_S)$ time. Phase 4 has 3 cases. Case 1 checks Eq. (4) with $O(k + h_p)$ time. Case 2 uses Eq. (7) to find t_p, which takes $O(h_p + k)$ time. The worst case is when we reduce t_p from T_S to 0. Each iteration finds L_m by Eq. (5). Thus, case 2 spends $O(h_p + k + T_S h_m)$ time. Case 3 has at most T_S iterations, each using Eqs. 5 and 8. It takes time of $O(T_S(h_m + h_p + k))$. So, phase 4 spends $O(T_S(h + k))$ time. To sum up, DA3 requires time of $O(hqT_S + k(h + T_S))$. □

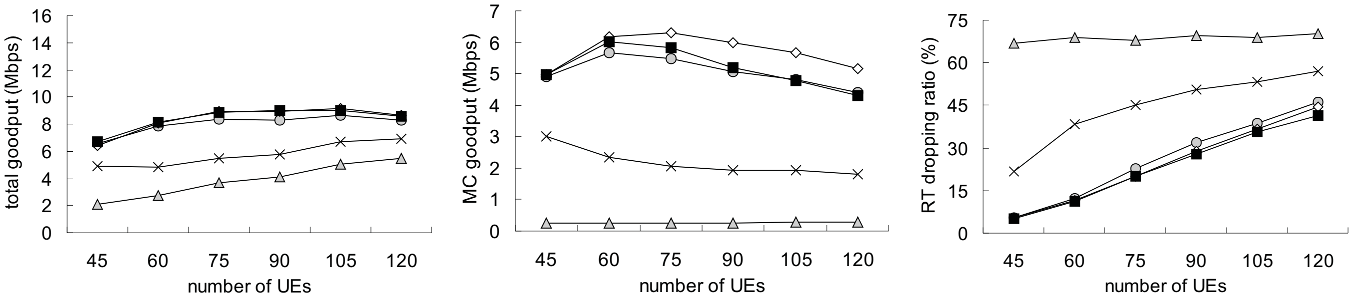
4 SIMULATION STUDY

We use LTE-Sim [14] to evaluate performance, which considers an MC with two PCs sharing a 5MHz channel. Fig. 2 gives network topology, where MC and PC radiuses are 1km and 100m, respectively. We use this layout for two reasons. First, it helps us observe the effect of ABS more obviously by eliminating the effect of other factors (e.g., eNB deployment). Second, it roughly keeps UE density in cells even if UEs have mobility (i.e., 3km/h speed). Thus, we can measure the effect of UE number more precisely.

The transmitted power of the MC eNB is set to 46dBm. Because LTE-Sim applies the small-cell scenario of only femtocells, we set the transmitted power of femtocell eNBs to 30dBm to simulate PC ones. An MC UE u_i has path loss of $128.1 + 37.6 \log D_i$, where D_i is the distance from u_i to its eNB in kilometers. The path loss of PC UEs is $140.7 + 36.7 \log D_i$. We use the log-normal distribution with 0dB mean and 8dB standard deviation for shadowing effect. Since LTE-Sim does not support ABS, we divide time into 1ms slots. In ABS slots, MC eNB sends a 30byte packet using 1/3 transmitted power to imitate control signals. Such packets are not counted in MC goodput. Besides, 2/3 UEs are RT UEs, each with a 242kbps video flow whose packet deadline is 100ms. Each non-RT UE has a 12kbps flow. Two scenarios are used for UE distribution. About 2/3 and 1/6 UEs locate in PCs in scenarios I and II, respectively. We use three methods, *MC-biased*, *balanced*, and *PC-biased*, for comparison, which set γ to 0.1, 0.5, and 0.9,



(a) Scenario I: total goodput (b) Scenario I: MC goodput (c) Scenario I: RT dropping ratio



(d) Scenario II: total goodput (e) Scenario II: MC goodput (f) Scenario II: RT dropping ratio

Fig. 3: Experimental results.

respectively. We also compare DA3 with the *interference-aware* method [10]. In DA3, we set $\varphi = 0.3$ and $\delta = 0.5$.

Fig. 3(a) gives total goodput in scenario I. As each cell has similar number of UEs, MC-biased and PC-biased methods respectively degrade PC and MC performance, so they have lower goodput than the balanced method. The interference-aware method uses throughput difference to adjust γ but does not perform well. Our DA3 method considers not only SINR but also queue content, so it has total goodput similar to the balanced method. Fig. 3(b) shows MC goodput in scenario I. The MC-biased method achieves the highest MC goodput, as it uses the fewest ABS slots. The PC-biased method uses the most ABS slots, so its MC goodput is close to zero. DA3 ensures MC transmission by using φ, so it has higher MC goodput than balanced and interference-aware methods. Fig. 3(c) gives the RT dropping ratio in scenario I, which is inversely proportional to total goodput. Thus, both DA3 and balanced methods have lower dropping ratios.

Fig. 3(d)–(f) give the results in scenario II, where MC UEs are the majority. The balanced method is no longer the winner (in fact, it performs worse than most methods), which shows the necessity to change γ. Since most UEs locate in the MC, the interference-aware method lowers γ to improve performance. Thus, it has the highest MC goodput in Fig. 3(e). DA3 also allocates fewer ABS slots to increase overall goodput (based on case 3 in phase 4), but it seeks to meet QoS demand of PC UEs by Eq. (8). Thus, DA3 uses more ABS slots than the interference-aware method, which decreases MC goodput, as shown in Fig. 3(e). In this way, DA3 can significantly reduce RT packet dropping (especially for PC UEs), as shown in Fig. 3(f).

5 CONCLUSION AND FUTURE WORK

Many studies derive an optimal ABS ratio by fixed network condition or change it based on UEs' demand, but few of them address RT traffic. Considering cell capacity, network status, and queue content, we propose the DA3 method to improve performance while reducing RT delay. By LTE-Sim experiments, we show that DA3 increases overall goodput, ensures MC transmission, and reduces RT packet dropping. We have three issues for future work. First, we assume that each MC has its own PCs. When some MCs cover the same PC, DA3 uses a divide-and-conquer method to let each MC decide its ABS ratio. We can improve DA3 by allowing MC eNBs to cooperatively compute their ratios. Second, we expect to integrate ABS with carrier aggregation, which allows an eNB to combine multiple channels to get large bandwidth. Since MC and PC eNBs can use different channels, it deserves investigation how to select channels with ABS to send data. Finally, as LTE-A proposes RP-ABS in FeICIC, we will consider using RP-ABS to help MC eNBs send RT traffic.

REFERENCES

[1] 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.
[2] F. Alfarhan, R. Lerbour, and Y.L. Helloco, "An optimization framework for LTE eICIC and reduced power eICIC," *Proc. IEEE Global Comm. Conf.*, 2015, pp. 1–6.
[3] M. Al-Rawi, M. Simsek, and R. Jantti, "Utility-based resource allocation in LTE-advanced heterogeneous networks," *Proc. IEEE Int'l Wireless Comm. and Mobile Computing Conf.*, 2013, pp. 826–830.
[4] 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.
[5] 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.
[6] M. Al-Rawi, J. Huschke, and M. Sedra, "Dynamic protected-subframe density configuration in LTE heterogeneous networks," *Proc. IEEE Int'l Conf. Computer Comm. and Networks*, 2012, pp. 1–6.
[7] 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.
[8] 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.

[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] 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.

[11] A. Daeinabi, K. Sandrasegaran, and S. Barua, "A dynamic almost blank subframe scheme for video streaming traffic model in heterogeneous networks," *Proc. IEEE Int'l Conf. Electrical Engineering/Electronics, Computer, Telecomm. and Information Technology*, 2015, pp. 1–6.

[12] ETSI, "Evolved universal terrestrial radio access (E-UTRA); physical layer procedures (release 14)," 3GPP TS 36.213 V14.1.0, 2016.

[13] 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.

[14] 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.