Hong and Park *EURASIP Journal on Wireless Communications and Networking* (2015) 2015:101
DOI 10.1186/s13638-015-0335-0

EURASIP Journal on Wireless Communications and Networking
a SpringerOpen Journal

RESEARCH Open Access

Recursive load balance scheme for two-tier cellular networks using enhanced inter-cell interference coordination

Myunghoon Hong and Seungyoung Park*

Abstract

Two-tier cellular networks in which small cells are overlaid with macro-cells have attracted considerable attention because these networks are expected to accommodate the rapidly increasing mobile traffic. Enhanced inter-cell interference coordination (eICIC) has also been proposed to further enhance the overall system performance by offloading more macro-cell users to small cells. Even with this approach, it is likely that macro-cell users are effectively offloaded to the small cells in the network because of the following reasons: i) some macro-cell users are offloaded to small cells by simply comparing their received signal strengths, and ii) the almost blank subframe (ABS) in the small cell is strictly assigned to users who experience strong interference from its nearby macro-cells. In this paper, we propose a load balance scheme for improving the downlink network-wide proportional fairness (PF) for two-tier cellular networks using eICIC. In particular, employing a cumulative distribution function-based scheduling scheme that supports sharing ABS across all users in each small cell, the proposed scheme recursively combines i) the user association between the cells and ii) ABS rate adaptation to improve the network PF. Through system-level simulations, we demonstrate that the proposed scheme noticeably improves the network PF compared to the eICIC scheme regardless of the number of macro-cell users offloaded to small cells.

Keywords: Load balance; Enhanced inter-cell interference coordination; eICIC; Two-tier cellular network; Almost blank subframe; ABS; CDF-based scheduling

1 Introduction

The wireless network has been evolved with the latest technology such as Long-Term Evolution (LTE) to meet the exponential traffic growth in mobile broadband services. However, LTE is likely to be still insufficient to meet this traffic growth because mobile data traffic has been predicted to increase 13-fold between 2012 and 2017 [1]. To meet the increasing traffic demands, the concept of spatial reuse has attracted considerable attention.

To enable improved spatial reuse, two-tier cellular networks in which small cells are overlaid by macro-cells have been introduced [2]. A large number of cells deployed in a fixed area can reduce the distances between transmitter-receiver pairs; hence, their radio link qualities can be enhanced to possibly support larger data rates. Typically, the coverage of a macro-cell is much larger than that of a small cell because the transmit power of the macro-cell is much higher. Thus, a severe imbalance results with respect to the numbers of users associated with macro-cells and small cells, if a user is served by a cell for which the received signal strength is strongest.

To offload macro-cell users to small cells, cell range expansion (CRE) has been employed in which a user can be served by a small cell if the received signal strengths from its original serving macro-cell and the small cell is less than a prespecified value [3]. Specifically, the coverage of small cell i is extended by CRE bias, i.e., the cell associated with an arbitrary user k is determined as

$$\text{User } k\text{'s associated cell} = \max_i \left(\Gamma_i \, \text{RSRP}_{i,k}\right)$$

where $\text{RSRP}_{i,k}$ and Γ_i denote the average received signal power of user k from cell i and CRE bias of cell i,

*Correspondence: s.young.park@kangwon.ac.kr
College of Information Technology, Kangwon National University, 192-1, Hyoja-dong, 200-701 Chuncheon, South Korea

respectively. Also, the CRE bias satisfies that $\Gamma_i > 1$ if cell i is a small cell and $\Gamma_i = 1$ if cell i is a macro-cell. As a result, this user will experience severe cross-tier interference from the macro-cell. To alleviate the cross-tier interference problem, enhanced inter-cell interference coordination (eICIC) has been introduced [3]. In particular, the frames are divided in the time domain into two subframes whose rates are the same for all cells, and every macro-cell is simultaneously muted during a prespecified subframe in every frame, which is called the almost blank subframe (ABS). Under this frame structure, the small-cell user who becomes a victim because of CRE can avoid the cross-tier interference if the ABS is assigned. Further, eICIC has been employed in carrier aggregation (CA) that utilizes multiple component carriers to expand the effective bandwidth supporting a high peak data rate [2,4]. In particular, some component carriers are used for only macro-cells, and the others are shared by macro-cells and small cells. Also, each macro-cell avoids transmissions or reduces its transmission power of data during some subframes on the shared component carriers. In this case, the macro-cell edge user can receive data from its associated macro-cell on the macro-cell component carriers and the closest small-cell on the shared component carriers that are used by macro-cells and small cells, simultaneously [4,5]. Studies have shown that the overall system throughput performance is significantly improved if the joint interference management and cell association in time and frequency domains is employed [6]. In particular, the CRE bias and the transmit power allocation of each cell are adjusted using reinforcement learning. As the joint interference management and cell association has been proposed for two-tier networks with CA, it is beyond the scope of this paper. We restrict ourselves to considering two-tier networks without CA.

The performance of two-tier networks using eICIC is known to be sensitive to the setting of CRE bias and the ABS rate [7]. Thus, several schemes have been proposed to maximize the sum of users' utilities in the network (i.e., network-wide utility) through adapting the ABS rate and determining users that should receive the resource within the ABS or non-ABS [8-10]. However, it is unlikely that these schemes effectively balance the loads among the macro-cells and the small cells. The reason is that some macro-cell users are offloaded to the small cells by simply comparing the received signal strengths due to CRE.

In a single-tier macro-cell network, a load balance (LB) problem was formulated as maximizing the network-wide utility such as proportional fairness (PF). As the problem was shown to be non-deterministic polynomial-time (NP)-hard, a heuristic algorithm was proposed such that each user is offloaded to another cell only when the network utility increases in a distributed manner [11,12]. Therefore, we expect that the network load of a two-tier network using eICIC becomes better balanced if a heuristic algorithm similar to those in [11,12] is employed instead of CRE.

Typically, the numbers of victim and non-victim users are quite different among the cells. For a given ABS rate, when the number of the non-victim users is much larger than that of the victim users in some small cells, non-victim users in these small cells could not be assigned a sufficient resource while victim users would be assigned too much resource. This unbalanced resource allocation cannot be mitigated because the ABS rate which is identical to all cells cannot be adjusted only for these small cells. However, as ABS can be assigned to the non-victim users without any interference problem, the unbalanced allocation problem is expected to be mitigated if ABS is shared across the non-victim and the victim users. Computer simulations have demonstrated that the throughput performance can be improved if a similar assignment policy sharing both subframes among the victim and the non-victim users is employed using a proportional fair scheduling [13]. In this case, each small-cell user separately estimates its transmission rates corresponding to ABS and non-ABS, respectively, because the interference power levels at a small cell may be quite different from ABS and non-ABS [14].

To achieve this, the small cell transmits a pilot signal and each small-cell user estimates its channels on ABS and non-ABS using the pilot signal. Note that the pilot signal is not transmitted on the entire frame. Particularly, in LTE, the pilot signal occupies prespecified locations of time and frequency resources within the frame [15]. After the channel estimation, the received pilot signals are regenerated. Then, the signals of interference plus noise on ABS and non-ABS are estimated by subtracting the corresponding regenerated pilot signals from the corresponding received signals [16]. After the interference-plus-noise powers are estimated using these signals, the signal-to-interference-plus-noise power ratios (SINRs) on ABS and non-ABS are estimated. From the estimated SINRs, the transmission rates are estimated using rate estimation techniques such as the exponential effective SINR [17].

After the estimated rates are sent to its serving small cell, each resource in ABS and non-ABS is assigned to the user with the largest proportional fair metric irrespective of whether the user is a victim or a non-victim. However, if non-ABS is assigned to the victim user, the transmission rate would be almost near zero due to strong cross-tier interference from nearby macro-cells. For this reason, in this paper, non-ABS is allowed to be assigned strictly to the non-victim users while ABS is

allowed to be assigned to the non-victim and the victim users.

Unfortunately, it is unlikely that the sharing assignment using proportional fair scheduling can be easily employed in a LB problem for two-tier networks using eICIC. This is because a particular user's throughput needs to be estimated in advance to determine whether it is offloaded [11]. However, the theoretical evaluation of user throughput has not yet been known. This difficulty can be avoided if a cumulative distribution function (CDF)-based scheduling is employed, because the throughput of a particular user can be easily evaluated, while the performance is similar to that of proportional fair scheduling [18,19].

Therefore, in this paper, we propose a LB scheme to improve the downlink network-wide utility using the CDF-based scheduling that supports sharing ABS across victim and non-victim users. In particular, the proposed scheme recursively combines i) the user association between cells and ii) the adaptation of the ABS rate such that the network utility is improved. System-level simulations demonstrate that the proposed scheme improves the network utility compared to the eICIC scheme regardless of the number of macro-cell users offloaded to small cells.

The remainder of this paper is organized as follows. Section 2 describes the frame structure supporting eICIC in a two-tier cellular network, the received system model, and the user association. The scheduling process supporting the assignment policy that shares ABS across victim as well as non-victim users is described in Section 3. The proposed recursive LB scheme is described in Section 4. System-level simulation results are presented in Section 5. Finally, we draw a conclusion in Section 6.

The notations adopted in this paper are as follows. We use $|\mathcal{U}|$ to denote the cardinality of set $\mathcal{U}$, abs(a) to denote the absolute value of complex number a, $(\cdot)^*$ to denote complex conjugate, $\delta(\cdot)$ to denote the Kronecker-delta function, and $F_X(\cdot)$ to denote a CDF of random variable X, respectively. Further, a complex random variable Z is called a zero-mean circularly symmetric complex Gaussian random variable, if and only if the real and imaginary parts of Z are statistically independent and identically distributed (i.i.d.) Gaussian with a zero mean and a variance of $\frac{\sigma^2}{2}$ and it is expressed as $Z \sim \mathcal{CN}\left(0, \sigma^2\right)$.

2 System description

2.1 Frame structure

Consider a two-tier cellular network in which each cell site is divided into three sectored macro-cells under the hexagonal cellular structure shown in Figure 1. Each macro-cell underlays L small cells. Assume that orthogonal frequency division multiple access (OFDMA) is employed with the frequency reuse factor of one. In

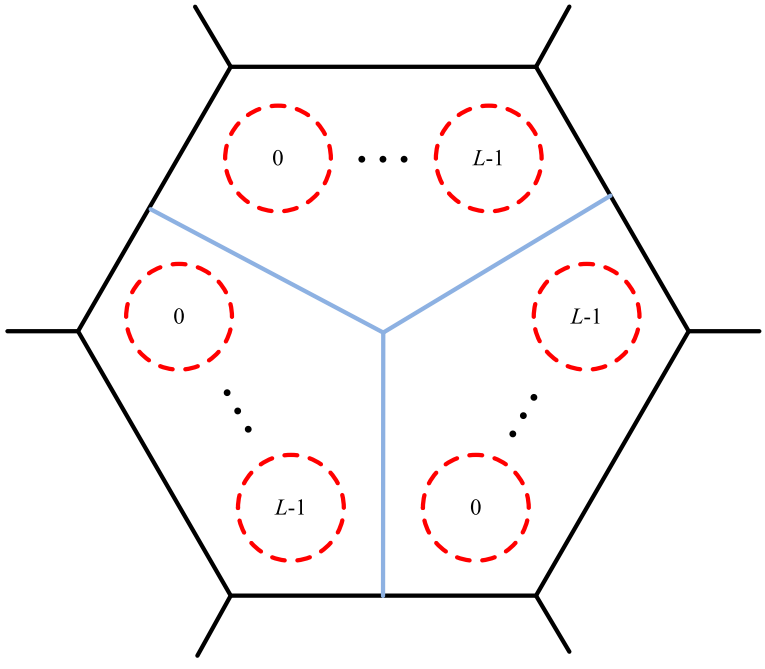

Figure 1 Cell structure for two-tier cellular network.

particular, assume that each frame has B physical resource blocks (PRBs) and each PRB occupies Δt in time and Δf in frequency [15].

Figure 2 shows a frame structure supporting eICIC. Assuming that accurate time synchronization between all cells is achieved, each frame is divided in the time domain into non-ABS and ABS. Note that the ratio of the ABS length and the frame length is defined as ABS rate α (for $0 < \alpha < 1$), which is the same for all cells. In practice, however, the ABS rate can be chosen from a set of prespecified values instead of any value for $0 < \alpha < 1$ because each frame consists of multiple subframes and each subframe can be either ABS or non-ABS. For example, when each frame consists of ten subframes, all possible values of α are 0.1, 0.2, ..., and 0.9. However, in this paper, we assume that the ABS rate can be chosen as any value for $0 < \alpha < 1$.

In practice, each macro-cell does not transmit data, but it still transmits the most essential information such as the pilot signal during ABS [20]. To fully benefit from ABS, each victim user who receives data from a small cell during ABS may employ an interference cancellation receiver such that data is detected after the pilot signals from nearby macro-cells are estimated and subtracted from the received signal [21]. Assuming that the pilot signals from macro-cells are perfectly cancelled on ABS duration, the victim users would not experience strong cross-tier interference on ABS duration. Therefore, we assume that the small-cell users do not experience the interference from the macro-cells during ABS.

2.2 Received signal model

Consider that each cell and each user are equipped with a single antenna. Assuming that each subcarrier within a PRB experiences the same flat fading, the downlink discrete-time complex baseband signal received by user k at PRB b in frame q is

$$y_{k,b}^{q} = \sum_{i \in \mathcal{T}} \sqrt{P_i G_{i,k}} h_{i,k,b}^{q} s_i + n_{k,b}, \tag{1}$$

where P_i, s_i, $G_{i,k}$, $h_{i,k,b}^{q}$, $n_{k,b}$, and $\mathcal{T}$ denote the transmit power of cell i, the transmit signal of cell i with $E\{s_i\} = 0$ and $E\{s_i^* s_{i'}\} = \delta(i - i')$, the path gain between cell i and user k including the effects of path loss and shadowing, a Rayleigh fading channel between cell i and user k at PRB b in frame q, an additive white Gaussian noise distributed according to $\mathcal{CN}(0, N_0)$, and

$$\mathcal{T} = \begin{cases} \mathcal{M} \cup \mathcal{S} & \text{if } b \in \mathcal{A}_\alpha \\ \mathcal{S} & \text{if } b \in \mathcal{B}_\alpha, \end{cases} \tag{2}$$

respectively. In (2), $\mathcal{M}$, $\mathcal{S}$, $\mathcal{A}_\alpha$, and $\mathcal{B}_\alpha$ denote a set of macro-cells, a set of small cells, a set of PRBs within non-ABS, and a set of PRBs within ABS for a given value of α, respectively. Assume that $h_{i,k,b}^{q}$ is static during each frame, independent from PRB to PRB, known perfectly to the receiver, distributed according to $\mathcal{CN}(0,1)$, and $E\left\{\left(h_{i,k,b}^{q}\right)^* h_{i',k',b'}^{q'}\right\} = \delta(i - i')\delta(k - k')\delta(b - b')\delta(q - q')$.

Under this signal model, the downlink instantaneous maximum achievable rates for user k associated with cell i at b in $\mathcal{A}_\alpha$ and $\mathcal{B}_\alpha$ of frame q are

$$R_{i,k,b}^{q} = \Delta f \log_2 \left(1 + \frac{P_i G_{i,k} \left(\text{abs}\left(h_{i,k,b}^{q}\right)\right)^2}{\sum_{i' \in (\mathcal{M} \cup \mathcal{S}) \setminus \{i\}} P_{i'} G_{i',k} \left(\text{abs}\left(h_{i',k,b}^{q}\right)\right)^2 + N_0}\right) \tag{3}$$

and

$$\tilde{R}_{i,k,b}^{q} = \begin{cases} 0 & \text{for } i \in \mathcal{M} \quad (4) \\ \Delta f \log_2 \left(1 + \dfrac{P_i G_{i,k} \left(\text{abs}\left(h_{i,k,b}^{q}\right)\right)^2}{\sum_{i' \in \mathcal{S} \setminus \{i\}} P_{i'} G_{i',k} \left(\text{abs}\left(h_{i',k,b}^{q}\right)\right)^2 + N_0}\right) & \text{for } i \in \mathcal{S}, \quad (5) \end{cases}$$

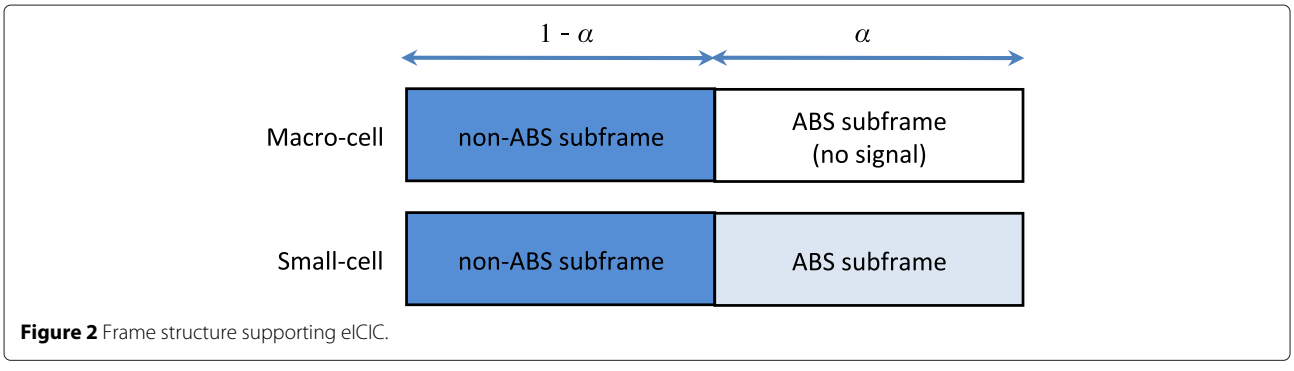

Figure 2 Frame structure supporting eICIC.

respectively. Note that (5) is derived using the assumption that small-cell users do not experience the interference from the macro-cells during ABS as discussed in Section 2.1. In addition, assume that the downlink data queues of different users are infinitely backlogged and any queueing dynamics are not considered in this paper.

In the remainder of this paper, for simplicity, we drop the superscript q indicating the frame index.

2.3 CRE and eICIC

We describe how a user is associated with a cell if the eICIC scheme is employed. User k determines its associated cell as follows. First, user k finds cell i such that

$$i = \arg\max_{i' \in \mathcal{M} \cup \mathcal{S}} P_{i'} G_{i',k}. \tag{6}$$

If $i \in \mathcal{S}$, user k is associated with cell i. Otherwise, user k re-finds cell $\hat{i}$ such that

$$\hat{i} = \arg\max_{i' \in \mathcal{M} \cup \mathcal{S}} \Gamma_{i'} P_{i'} G_{i',k} \tag{7}$$

and it is associated with cell $\hat{i}$, where Γ_i denotes the CRE bias of cell i such that $\Gamma_i > 1$ for $i \in \mathcal{S}$, and $\Gamma_i = 1$ for $i \in \mathcal{M}$. Note that the CRE mechanism does not change the network operation considerably, because it uses the framework already available to carry out the handover procedure.

After the user association, a user associated with a macro-cell is assigned non-ABS (i.e., macro-cell user). Further, a user that is associated with a small cell and receives the strongest signal from it is assigned non-ABS (i.e., non-victim user). Besides, a user forcibly associated with a small cell satisfying (7) is assigned only ABS because it experiences strong cross-tier interference (i.e., victim user).

We have described the user association and resource assignment policy in the eICIC scheme where non-ABS and ABS are strictly assigned to the non-victim and the victim users, respectively. In the following, we describes the user association and resource assignment to support sharing ABS across the non-victim and victim users for the proposed scheme.

2.4 User association

Each user can be categorized into one of the three different types of users depending on which type of cell is associated and which type of resource is assigned. Unlike those in the eICIC scheme, a macro-cell user, a non-victim user, and a victim user are assigned PRBs in i) non-ABS in a macro-cell, ii) non-ABS and ABS in a small cell, or iii) ABS in a small cell, respectively. From this, assume that user k satisfies

$$\sum_{i \in \mathcal{M}} X_{i,k} + \sum_{i \in \mathcal{S}} (Y_{i,k} + Z_{i,k}) = 1, \tag{8}$$

where $X_{i,k}$, $Y_{i,k}$, and $Z_{i,k}$ denote variables to indicate which resource in cell i is assigned to user k. In particular, $X_{i,k} = 1$ if user k is assigned PRBs in non-ABS of cell i for $i \in \mathcal{M}$, and 0 otherwise. Besides, $Y_{i,k} = 1$ if user k is assigned PRBs in non-ABS and ABS of cell i for $i \in \mathcal{S}$, and 0 otherwise. $Z_{i,k} = 1$ if user k is assigned PRBs in ABS of cell i for $i \in \mathcal{S}$, and 0 otherwise. Note that a given non-victim user k receiving the strongest signal from its associated small cell is allowed to be assigned non-ABS and ABS (i.e., $Y_{i,k} = 1$) because non-ABS can be assigned without a severe cross-tier interference problem. However, the victim user receiving the strongest interference signal from its nearby macro-cells is allowed to be assigned ABS only (i.e., $Z_{i,k} = 1$).

In addition, define $\mathcal{U}_i$ as a set of users associated with cell i as

$$\mathcal{U}_i = \begin{cases} \{k \mid X_{i,k} = 1\} & \text{for } i \in \mathcal{M} \\ \mathcal{V}_i \cup \mathcal{W}_i & \text{for } i \in \mathcal{S}, \end{cases} \tag{9}$$

where $\mathcal{V}_i = \{k \mid Y_{i,k} = 1\}$ and $\mathcal{W}_i = \{k \mid Z_{i,k} = 1\}$. From (8), it is satisfied that $\mathcal{U}_i \cap \mathcal{U}_j = \emptyset$ for $i \neq j$ and $\mathcal{V}_i \cap \mathcal{W}_i = \emptyset$ for all i.

3 Description of scheduling process

In this section, we describe a CDF-based scheduling scheme supporting the assignment policy in which ABS is shared among the non-victim and the victim users in a small cell.

3.1 Description of CDF-based scheduling

First, we describe CDF-based scheduling under a single-tier cellular network as shown in the following lemma.

Lemma 1. [18] *Assume that cell i transmits a pilot signal for PRB b and user k on a given frame measures its downlink maximum achievable rate $R_{i,k,b}$ using this pilot signal. When the CDF of $R_{i,k,b}$ (i.e.,$F_{R_{i,k,b}}(\cdot)$) is perfectly known to the cell, user k sends $R_{i,k,b}$ to the cell and it is converted to $V_{i,k,b}$ ($\equiv F_{R_{i,k,b}}(R_{i,k,b})$). Using $\{V_{i,k,b}\}$ from all users associated with cell i, PRB b is assigned to user $k^\dagger$ such that*

$$k^\dagger = \arg\max_k \left(V_{i,k,b}\right)^{1/w_{i,k}}, \tag{10}$$

where $w_{i,k}$ is a non-negative weighting factor and identical to all PRBs for given cell i and user k, and $\sum_k w_{i,k} = 1$. Then, cell i sends the downlink signal with a rate of $R_{i,k^\dagger,b}$ to user $k^\dagger$ on PRB b.

Under this setup, the average resource allocation rate and the average throughput of user k are

$$r_{i,k} = w_{i,k} \tag{11}$$

and

$$T_{i,k} = \sum_{b} \int_0^1 u^{(1/w_{i,k}-1)} F^{-1}_{R_{i,k,b}}(u)du, \tag{12}$$

respectively.

The proof of this lemma can be found in [18]. In this scheduling, each user feels as if virtually $1/w_{i,k}$ users with the rate distributions identical to its own distribution were competing for a given PRB [18]. It implies that CDF-based scheduling exploits multiuser diversity. Also, when the average resource allocation rate is identical for all users in the cell (i.e., $w_{i,k}$ is the same for all users), each user's throughput performance approaches that of a proportional fair scheduling scheme under a Rayleigh fading channel as its moving window size to estimate the average throughput increases.

Further, we can extend to the case when the spatial diversity from multiple transmit antennas is exploited along with the CDF-based scheduling if employing the unitary beamforming technique in which multiple signal streams are simultaneously transmitted through the orthonormal beamforming vectors at the transmitter as suggested in [22]. In particular, the transmitter sends the multiple orthogonal pilot signals through the orthonormal beamforming vectors. Using these pilot signals, each user estimates the rates of the signal streams and sends them to the scheduler. Then, each stream is assigned to the user with the maximum scheduling metric. As shown in [22], it is expected that the throughput performance is increased roughly proportional to the number of the streams when the number of users in the scheduler is sufficiently large.

In addition, from (11), we see that the average resource allocation rates among the users can be easily controlled by adjusting $\{w_{i,k}\}$ and their corresponding throughput performances. Due to these properties, we apply CDF-based scheduling in the proposed scheme to support sharing ABS among the non-victim and victim users in the following.

3.2 Scheduling process of proposed scheme

Macro-cell case: Assume that macro-cell i transmits a pilot signal for a given frame. Recalling that a macro-cell assigns PRBs in only non-ABS, each user k in macro-cell i estimates its downlink maximum achievable rates $\{R_{i,k,b}\}$ for PRBs in $\mathcal{A}_\alpha$. Assuming that the CDF of $R_{i,k,b}$ is perfectly known to cell i, user k sends $R_{i,k,b}$ to cell i and it is converted to $V_{i,k,b}(\equiv F_{R_{i,k,b}}(R_{i,k,b}))$. Using $\{V_{i,k,b}\}$, PRB b in $\mathcal{A}_\alpha$ is assigned to user k^* satisfying

$$k^* = \arg\max_{k\in\mathcal{U}_i} \left(V_{i,k,b}\right)^{1/w_{i,k}} \tag{13}$$

with $\sum_{k\in\mathcal{U}_i} w_{i,k} = 1$.

Under this scheduler, assume that every user's average resource allocation rate in macro-cell i is the same. To satisfy this, we employ the following weight factor that is identical to all PRBs as

$$w_{i,k} = \frac{1}{|\mathcal{U}_i|} \text{ for any } k \in \mathcal{U}_i. \tag{14}$$

As only non-ABS is available and it occupies $(1-\alpha)$ of all PRBs, the average resource allocation rate and the average throughput of user k are evaluated using Lemma 1 as

$$r_{i,k} = \frac{(1-\alpha)}{|\mathcal{U}_i|} \tag{15}$$

and

$$T_{i,k}(\mathcal{U}_i) = \sum_{b\in\mathcal{A}_\alpha} \int_0^1 u^{(1/w_{i,k}-1)} F^{-1}_{R_{i,k,b}}(u)du, \tag{16}$$

respectively.

Small-cell case: First, consider a resource assignment policy where ABS and non-ABS are strictly allocated to the victim and the non-victim users in a given small cell i, respectively. In this case, assuming that each user in a given user set ($\mathcal{V}_i$ or $\mathcal{W}_i$) has the same average resource allocation rate, the corresponding average resource allocation rate is given as

$$r_{i,k} = \begin{cases} \dfrac{(1-\alpha)}{|\mathcal{V}_i|} & \text{if } k \in \mathcal{V}_i \\ \dfrac{\alpha}{|\mathcal{W}_i|} & \text{if } k \in \mathcal{W}_i. \end{cases} \tag{17}$$

With this in mind, consider the condition when larger resources are allocated to victim users on an average, i.e.,

$$\frac{\alpha}{|\mathcal{W}_i|} > \frac{(1-\alpha)}{|\mathcal{V}_i|}. \tag{18}$$

In this case, recalling that ABS can be also assigned to non-victim users, we can fairly assign the resource among the users at the small cell if ABS is allowed to be assigned to both victim and non-victim users. To do this, we employ the scheduling scheme as follows.

First, assume that small cell i transmits a pilot signal on both non-ABS and ABS. Recalling that a small cell assigns PRBs in both non-ABS and ABS to non-victim user k in $\mathcal{V}_i$, this user estimates its downlink maximum achievable rates $\{R_{i,k,b}\}$ and $\{\tilde{R}_{i,k,b}\}$ on non-ABS and ABS, respectively. On the other hand, because victim user k in $\mathcal{W}_i$ can be assigned PRBs only on ABS to avoid the cross-tier interference, this user estimates its downlink maximum achievable rates $\{\tilde{R}_{i,k,b}\}$.

Assuming that the CDFs of $\{R_{i,k,b}\}$ (and $\{\tilde{R}_{i,k,b}\}$) are perfectly known to cell i, user k sends $R_{i,k,b}$ (and $\tilde{R}_{i,k,b}$) to cell i and it is converted to $V_{i,k,b}\left(\equiv F_{R_{i,k,b}}\left(R_{i,k,b}\right)\right)\left(\text{and } \tilde{V}_{i,k,b}\left(\equiv F_{\tilde{R}_{i,k,b}}\left(\tilde{R}_{i,k,b}\right)\right)\right)$. Then, PRB b on non-ABS at a given frame of small cell i is assigned to non-victim user $k^{\dagger} \in \mathcal{V}_i$ satisfying

$$k^{\dagger} = \arg\max_{k\in\mathcal{V}_i} \left(V_{i,k,b}\right)^{1/w_{i,k}} \tag{19}$$

with $\sum_{k\in\mathcal{V}_i} w_{i,k} = 1$ and the value of $w_{i,k}$ is assumed to be identical to all PRBs. As ABS can be assigned to non-victim users as well as the victim users, PRB b on ABS at a given frame is assigned to user $k^{\ddagger} \in \mathcal{U}_i\ (= \mathcal{V}_i \cup \mathcal{W}_i)$ satisfying

$$k^{\ddagger} = \arg\max_{k\in\mathcal{U}_i} \left(\tilde{V}_{i,k,b}\right)^{1/\tilde{w}_{i,k}} \tag{20}$$

with $\sum_{k\in\mathcal{U}_i} \tilde{w}_{i,k} = 1$ and $\tilde{w}_{i,k}$ is assumed to be identical to all PRBs.

Under this scheduler, assume that every user's average resource allocation rate in small cell i is the same as

$$r_{i,k} = \frac{1}{|\mathcal{U}_i|}. \tag{21}$$

To satisfy this, the weight factors $\{w_{i,k}\}$ in (19) and $\{\tilde{w}_{i,k}\}$ in (20) are employed as

$$w_{i,k} = \frac{1}{|\mathcal{V}_i|} \quad \text{for} k \in \mathcal{V}_i \tag{22}$$

and

$$\tilde{w}_{i,k} = \begin{cases} \dfrac{|\mathcal{U}_i| - |\mathcal{W}_i|/\alpha}{|\mathcal{V}_i||\mathcal{U}_i|} & \text{for } k \in \mathcal{V}_i \\ \dfrac{1}{\alpha|\mathcal{U}_i|} & \text{for } k \in \mathcal{W}_i. \end{cases} \tag{23}$$

As a result, the average throughputs of user k in $\mathcal{V}_i$ and $\mathcal{W}_i$ are

$$S_{i,k}(\mathcal{V}_i, \mathcal{W}_i) = \sum_{b\in\mathcal{A}_\alpha} H_{i,k,b} + \sum_{b\in\mathcal{B}_\alpha} \tilde{H}_{i,k,b} \tag{24}$$

and

$$Q_{i,k}(\mathcal{V}_i, \mathcal{W}_i) = \sum_{b\in\mathcal{B}_\alpha} \tilde{H}_{i,k,b}, \tag{25}$$

respectively, where

$$H_{i,k,b} = \int_0^1 u^{(1/w_{i,k}-1)} F^{-1}_{R_{i,k,b}}(u)du \tag{26}$$

and

$$\tilde{H}_{i,k,b} = \int_0^1 u^{(1/\tilde{w}_{i,k}-1)} F^{-1}_{\tilde{R}_{i,k,b}}(u)du. \tag{27}$$

However, if (18) is not satisfied, i.e., $\frac{\alpha}{|\mathcal{W}_i|} > \frac{(1-\alpha)}{|\mathcal{V}_i|}$, we cannot achieve the average allocation rate of (21). The reason is that the victim user needs to be assigned non-ABS to achieve (18). However, if it is, the victim user would receive strong interference from its nearby macro-cells. Therefore, in this case, non-ABS and ABS are strictly assigned to only users in $\mathcal{V}_i$ and $\mathcal{W}_i$, respectively. Consequently, the average resource allocation rate of user k is equivalent to (17), and the average throughput for $\mathcal{V}_i$ and $\mathcal{W}_i$ are

$$S_{i,k}(\mathcal{V}_i, \mathcal{W}_i) = \sum_{b\in\mathcal{A}_\alpha} H_{i,k,b} \tag{28}$$

and

$$Q_{i,k}(\mathcal{V}_i, \mathcal{W}_i) = \sum_{b\in\mathcal{B}_\alpha} \tilde{H}_{i,k,b}, \tag{29}$$

respectively.

3.3 Implementation issues

Note that CDF-based scheduling requires the condition that all users' rate distributions $\left\{F_{R_{i,k,b}}(\cdot)\right\}$ and $\left\{F_{\tilde{R}_{i,k,b}}(\cdot)\right\}$ are perfectly known to the scheduler. However, because this condition is not feasible practically, their empirical CDFs are employed [23,24]. To achieve this, user k estimates the transmission rate on PRB b using the pilot signal from its associated cell i and sends it to cell i. Then, cell i generates the empirical CDF of the rate on PRB b by accumulating user k's rate feedback information on PRB b.

The throughput performance loss due to the empirical CDF of any arbitrary channel is shown to be negligible as the number of each user's rate feedbacks to estimate its empirical CDF grows roughly $O(K^2)$, where K is the number of users at the scheduler [24]. In particular, when the rate of a given PRB is statistically independent from frame to frame, at most $\frac{K}{2\epsilon}$ rate feedbacks per user are needed to achieve the performance loss of ϵ [24]. For example, when the number of users is 50 in a cell, at most 1,000 rate feedbacks per user are needed to achieve the loss of 2.5%.

Assuming that $F_{R_{i,k,b}}(\cdot)$ is i.i.d. for any PRB b of user k, cell i generates the empirical CDF of $R_{i,k,b}$, which is identical to any PRB of user k, by accumulating the rate feedback information of all PRBs from user k. Subsequently, its empirical CDF can be easily estimated with negligible throughput performance loss within a few frame durations because each frame typically consists of a large number of PRBs. However, in the OFDMA system, the channel fading correlations between adjacent PRBs for a given user are typically large. This means that $F_{R_{i,k,b}}(\cdot)$ is not i.i.d. for each PRB b of user k. To make the actual channel fading experienced on each PRB to be approximately i.i.d., we can employ random beamforming vectors that are identical within one PRB but independently generated between PRBs, because the random beamforming approach reduces the correlation between PRBs [25].

Now, we discuss how many rate feedbacks are required to acquire the empirical CDF for a given user. Recall that the number of rate feedbacks to acquire the empirical

CDF is at most 1,000 for the number of users in a cell of 50 and performance loss of 2.5% [24]. In LTE with a 10-MHz bandwidth and 10-ms frame duration, the number of PRBs per frame is 1,000 [15]. It means that the empirical CDF can be acquired if each user estimates its rates of all PRBs within one frame.

Furthermore, we can reduce the feedback load to acquire the empirical CDF because the number of modulation and coding scheme (MCS) levels supported in a real system is limited. This means that the rate is one of the transmission rates supported by the MCS levels. Therefore, the probability mass function (PMF) of the MCS levels is required to perform the scheduling instead of the CDF [18]. Assuming that the rate of each PRB is i.i.d., the PMF can be estimated as follows. Each user estimates the probability that a given MCS level is estimated on each PRB by accumulating the number of PRBs supporting the corresponding MCS level among all the PRBs [18]. Then, the estimated PMF is sent to the scheduler. As the number of the probabilities in PMF sent by each user is the same as that of MCS levels, each user's feedback loads can be significantly reduced compared to the empirical CDF case. In addition, the PMF needs to be updated at an appropriate interval when users move. However, further work is required to determine how often the PMF should be updated.

4 Description of the proposed LB scheme

4.1 Problem formulations

Recalling that from (16), (24), and (25), a user's average throughput is affected by the number of users associated with the same cells and the ABS rate α, the optimal LB algorithm finds a set of indicator variables

$$\{X_{i,k}\}_{i\in\mathcal{M}}, \{Y_{i,k}, Z_{i,k}\}_{i\in\mathcal{S}}$$

and the ABS rate α such that

$$\sum_{k\in\mathcal{U}} U_k(\gamma_k) \tag{30}$$

is maximized subject to the constraints in (8), where $\mathcal{U}$, $U_k(\cdot)$, and γ_k denote a set of all users in the network, user k's utility function that is continuously differentiable, monotonically increasing, and strict concave [26], and the average throughput of user k as

$$\gamma_k = \sum_{i\in\mathcal{M}} X_{i,k} T_{i,k} + \sum_{i\in\mathcal{S}} \left(Y_{i,k} S_{i,k} + Z_{i,k} Q_{i,k}\right) \tag{31}$$

for given $\{X_{i,k}\}_{i\in\mathcal{M}}$, $\{Y_{i,k}, Z_{i,k}\}_{i\in\mathcal{S}}$, and α, respectively. Note that if the utility function is linear, e.g., $U_k(\gamma_k) = \gamma_k$ for all k, (30) is maximized if the maximal-rate scheduling that selects the user with the largest transmission rate is employed [27]. Although the maximal-rate scheduling is throughput-optimal, it is not a satisfactory solution in terms of fairness among users. Instead of the linear function, in this paper, we employ a logarithmic function, i.e., $U_k(\cdot) = \log(\cdot)$ for all k because it has been known that the throughput and the fairness are well balanced among users [28]. Note that the sum of log throughputs of all users in the network, $\sum_{k\in\mathcal{U}} \log \gamma_k$, is defined as a network-wide PF. In addition, owing to the property of the logarithmic function, the network PF can be noticeably increased if the throughput of the users with low rates is improved rather than that of the users with high rates.

To find an optimal solution, the value of network-wide PF should be evaluated for all possible combinations of $\{X_{i,k}\}_{i\in\mathcal{M}}$, $\{Y_{i,k}, Z_{i,k}\}_{i\in\mathcal{S}}$, and α. In particular, because the number of all possible combinations of indicator variables for a given α is $(|\mathcal{M}| + 2|\mathcal{S}|)^{|\mathcal{U}|}$, this problem has been known to be NP-hard [11]. Thus, we propose a heuristic algorithm to find a suboptimal solution in this work. In particular, the proposed scheme combines i) the user association among the cells and ii) the ABS rate update in a recursive manner while improving the network-wide PF.

Now, we describe the main steps of the proposed scheme qualitatively. The user association consists of i) the intra-cell user association in which each user within a given small cell is associated between non-victim and victim user groups and ii) the inter-cell user association in which each user within a given cell is associated with different cells. In particular, in the intra-cell user association, the small cell evaluates the sum of log of throughputs of the users within its cell if one of the users in the non-victim user group moves to the victim user group and vice versa. After the evaluation is performed for all users within a given small cell, one user maximizing the sum of log user throughputs is selected and moved to a different user group.

In the inter-cell user association, on the other hand, a given cell evaluates the sum of log of throughputs of the users within it and one of its neighboring cells, if one of the users within the cell is offloaded to a corresponding neighboring cell. To achieve this, the rate CDFs of a given user measured at the neighboring cells are required. Recalling that each user measures and reports its neighboring cells' channel state information to its associated cell to facilitate the handover procedure, it is expected that the required rate CDFs can be obtained without changing the network operation considerably. After evaluating the sum of log user throughputs, we select one of the combinations of the users and the neighboring cells that maximizes the sum of log user throughputs. Then, the chosen user is offloaded to the corresponding neighboring cell.

In addition, assume that each cell performs the intra-cell association and the inter-cell association at every M_{intra} frames and every M_{inter} frames, respectively. In particular, cell i performs the inter-cell association at frame $(lM_{\text{inter}} + \varpi_i)$ where ϖ_i is a randomly chosen value for cell i among

$\{0, 1, \ldots, M_{\text{inter}}-1\}$. Similarly, cell i performs the intra-cell association at frame $(lM_{\text{intra}} + \tilde{\varpi}_i)$ where $\tilde{\varpi}_i$ is a randomly chosen value for cell i among $\{0, 1, \ldots, M_{\text{intra}} - 1\}$.

In the following sections, we describe in detail the processes of the proposed scheme.

4.2 User association update

Intra-cell user association: At every M_{intra} frames, small cell i chooses user $k^\star$ in $\mathcal{V}_i$ ($\mathcal{W}_i$) that achieves the largest improvement in the sum of log throughputs of all users in $\mathcal{U}_i$ if it moves to $\mathcal{W}_i$ ($\mathcal{V}_i$). In particular, user $k^\star$ is chosen such that

$$k^\star = \arg\max_{k \in \mathcal{U}_i} \underbrace{\left(\psi_{i,k} - \sum_{k' \in \mathcal{U}_i} \log \gamma_{k'} \right)}_{\equiv \Lambda_{i,k}} \tag{32}$$

where $\psi_{i,k}$ is

$$\psi_{i,k} = \sum_{k' \in \tilde{\mathcal{V}}_i} \log S_{i,k'} \left(\tilde{\mathcal{V}}_i, \tilde{\mathcal{W}}_i \right) + \sum_{k' \in \tilde{\mathcal{W}}_i} \log Q_{i,k'} \left(\tilde{\mathcal{V}}_i, \tilde{\mathcal{W}}_i \right). \tag{33}$$

In (33), $\tilde{\mathcal{V}}_i$ and $\tilde{\mathcal{W}}_i$ are denoted as

$$\tilde{\mathcal{V}}_i = \begin{cases} \mathcal{V}_i \backslash \{k\} & \text{if } k \in \mathcal{V}_i \\ \mathcal{V}_i \cup \{k\} & \text{if } k \in \mathcal{W}_i \end{cases}$$

and

$$\tilde{\mathcal{W}}_i = \begin{cases} \mathcal{W}_i \cup \{k\}, & \text{if } k \in \mathcal{V}_i \\ \mathcal{W}_i \backslash \{k\}, & \text{if } k \in \mathcal{W}_i, \end{cases}$$

respectively.

If $\Lambda_{i,k^\star} > 0$, user $k^\star$ moves to another user set within cell i. Equivalently, the association variable of user $k^\star$ is updated as

$$Y_{i,k^\star} = 0 \text{ and } Z_{i,k^\star} = 1, \text{ if } k^\star \in \mathcal{V}_i$$

and

$$Y_{i,k^\star} = 1 \text{ and } Z_{i,k^\star} = 0, \text{ if } k^\star \in \mathcal{W}_i.$$

Inter-cell user association: Assume that user k can be possibly associated with any cell in a set of candidate cells as

$$\begin{aligned} \mathcal{C}_k = &\left\{ i \mid G_{i,k} \geq G_{\pi_{(K-1)},k} \text{ for } i \in \mathcal{M} \right\} \\ &\cup \left\{ i \mid G_{i,k} \geq G_{\phi_{(\tilde{K}-1)},k} \text{ for } i \in \mathcal{S} \right\} \end{aligned} \tag{34}$$

where π and ϕ denote permutations such that $G_{\pi_0,k} \geq G_{\pi_1,k} \geq \cdots \geq G_{\pi_{(|\mathcal{M}|-1)},k}$ for $\pi_j \in \mathcal{M}$ and $G_{\phi_0,k} \geq G_{\phi_1,k} \geq \cdots \geq G_{\phi_{(|\mathcal{S}|-1)},k}$ for $\phi_j \in \mathcal{S}$, respectively; K and $\tilde{K}$ denote the numbers of the candidate macro-cells and small cells, respectively. With this, at every M_{inter} frames, user k in cell i finds cell i_k^* that achieves the largest improvement in the sum of log throughputs of the users in $\mathcal{U}_i$ and $\mathcal{U}_{i_k^*}$ if user k is offloaded to cell i_k^*. In particular, cell i_k^* is chosen such that

$$i_k^* = \arg\max_{i' \in \mathcal{C}_k \backslash \{i\}} \underbrace{\left(\Psi_{i,k} + \Phi_{i',k} - \sum_{k' \in \mathcal{U}_i \cup \mathcal{U}_{i'}} \log \gamma_{k'} \right)}_{\equiv \Delta_{i,i'}^k} \tag{35}$$

where $\Psi_{i,k}$ denotes the sum of log throughputs of users in $\mathcal{U}_i$ if user k is offloaded to the other cell as

$$\Psi_{i,k} = \begin{cases} \sum_{k' \in \mathcal{U}_i \backslash \{k\}} \log T_{i,k'}(\mathcal{U}_i \backslash \{k\}) & \text{for } i \in \mathcal{M} \\ \sum_{k' \in (\mathcal{V}_i \backslash \{k\}) \cup \mathcal{W}_i} \log S_{i,k'}(\mathcal{V}_i \backslash \{k\}, \mathcal{W}_i) & \text{for } i \in \mathcal{S} \text{ and } k \in \mathcal{V}_i \\ \sum_{k' \in \mathcal{V}_i \cup (\mathcal{W}_i \backslash \{k\})} \log Q_{i,k'}(\mathcal{V}_i, \mathcal{W}_i \backslash \{k\}) & \text{for } i \in \mathcal{S} \text{ and } k \in \mathcal{W}_i \end{cases} \tag{36}$$

and $\Phi_{i',k}$ denotes the sum of log throughputs of users in $\mathcal{U}_{i'}$ if user k is offloaded to cell i' as

$$\Phi_{i',k} = \begin{cases} \sum_{k' \in \mathcal{U}_{i'} \cup \{k\}} \log T_{i',k'}(\mathcal{U}_{i'} \cup \{k\}) & \text{for } i' \in \mathcal{M} \\ \max \left[J_{i',k}, \tilde{J}_{i',k} \right] & \text{for } i' \in \mathcal{S}. \end{cases} \tag{37}$$

In (37),

$$\begin{aligned} J_{i',k} = &\sum_{k' \in \mathcal{V}_{i'} \cup \{k\}} \log S_{i',k'}(\mathcal{V}_{i'} \cup \{k\}, \mathcal{W}_{i'}) \\ &+ \sum_{k' \in \mathcal{W}_{i'}} \log Q_{i',k'}(\mathcal{V}_{i'} \cup \{k\}, \mathcal{W}_{i'}), \end{aligned}$$

and

$$\begin{aligned} \tilde{J}_{i',k} = &\sum_{k' \in \mathcal{V}_{i'}} \log S_{i',k'}(\mathcal{V}_{i'}, \mathcal{W}_{i'} \cup \{k\}) \\ &+ \sum_{k' \in \mathcal{W}_{i'} \cup \{k\}} \log Q_{i',k'}(\mathcal{V}_{i'}, \mathcal{W}_{i'} \cup \{k\}). \end{aligned}$$

After all the users in cell i find $\{i_k^*\}$, cell i chooses the user that achieves the largest improvement in $\Delta_{i,i_k^*}^k$. In particular, user $k^\S$ is chosen such that

$$k^\S = \arg\max_{k \in \mathcal{U}_i} \Delta_{i,i_k^*}^k. \tag{38}$$

If $\Delta_{i,i_{k^\S}^*}^{k^\S} > 0$, user $k^\S$ in cell i is offloaded to cell $i_{k^\S}^*$. Equivalently, user $k^\S$'s association variable corresponding to cell i is updated as follows. If $i \in \mathcal{M}$,

$$X_{i,k^\S} = 0.$$

Otherwise,

$$Y_{i,k^{\S}} = 0 \text{ for } k^{\S} \in \mathcal{V}_i \text{ and } Z_{i,k^{\S}} = 0 \text{ for } k^{\S} \in \mathcal{W}_i.$$

In addition, its association variable corresponding to cell $i^*_{k^{\S}}$ is updated as follows. If $i^*_{k^{\S}} \in \mathcal{M}$

$$X_{i^*_{k^{\S}},k^{\S}} = 1.$$

If $i^*_{k^{\S}} \in \mathcal{S}$,

$$Y_{i^*_{k^{\S}},k^{\S}} = 1 \text{ for } J_{i^*_{k^{\S}},k^{\S}} > \tilde{J}_{i^*_{k^{\S}},k^{\S}} \text{ and } Z_{i^*_{k^{\S}},k^{\S}} = 1 \text{ otherwise.}$$

Note that the user association can be performed in a distributed manner because user k can be offloaded to one of the candidate cells in $\mathcal{C}_k$, which exchange the association information irrespective of whether they are macro-cells or small cells. Specifically, the association information to be exchanged is as follows. For a given user k in cell i, $\Psi_{i,k}$, $\left\{\sum_{k' \in \mathcal{U}_{i'}} \gamma_{k'}\right\}_{i' \in \mathcal{C}_k}$, and $\{\Phi_{i',k}\}_{i' \in \mathcal{C}_k \setminus \{i\}}$ are required to determine the most feasible candidate cell i^*_k as shown in (35). From (36), we see that $\Psi_{i,k}$ can be easily calculated because the throughputs of the users in $\mathcal{U}_i$ are already available. On the other hand, to calculate $\Phi_{i',k}$ in (37), first user k measures the downlink rates from cell i' and sends them to cell i. Using them, cell i acquires the rate CDF of user k for cell i' and sends it to cell i' using inter-cell interface such as the X2 interface in LTE [15]. Using the received rate CDF, cell i' calculates $\Phi_{i',k}$ and sends it to cell i as well as $\sum_{k' \in \mathcal{U}_{i'}} \gamma_{k'}$. This process for user k is performed for all cells in $\mathcal{C}_k \setminus \{i\}$. After receiving $\{\Phi_{i',k}\}_{i' \in \mathcal{C}_k \setminus \{i\}}$, cell i determines the most feasible candidate cell i^*_k for user k. Compared to the handover procedure, it is expected that the additional overhead to exchange the association information is not a heavy burden.

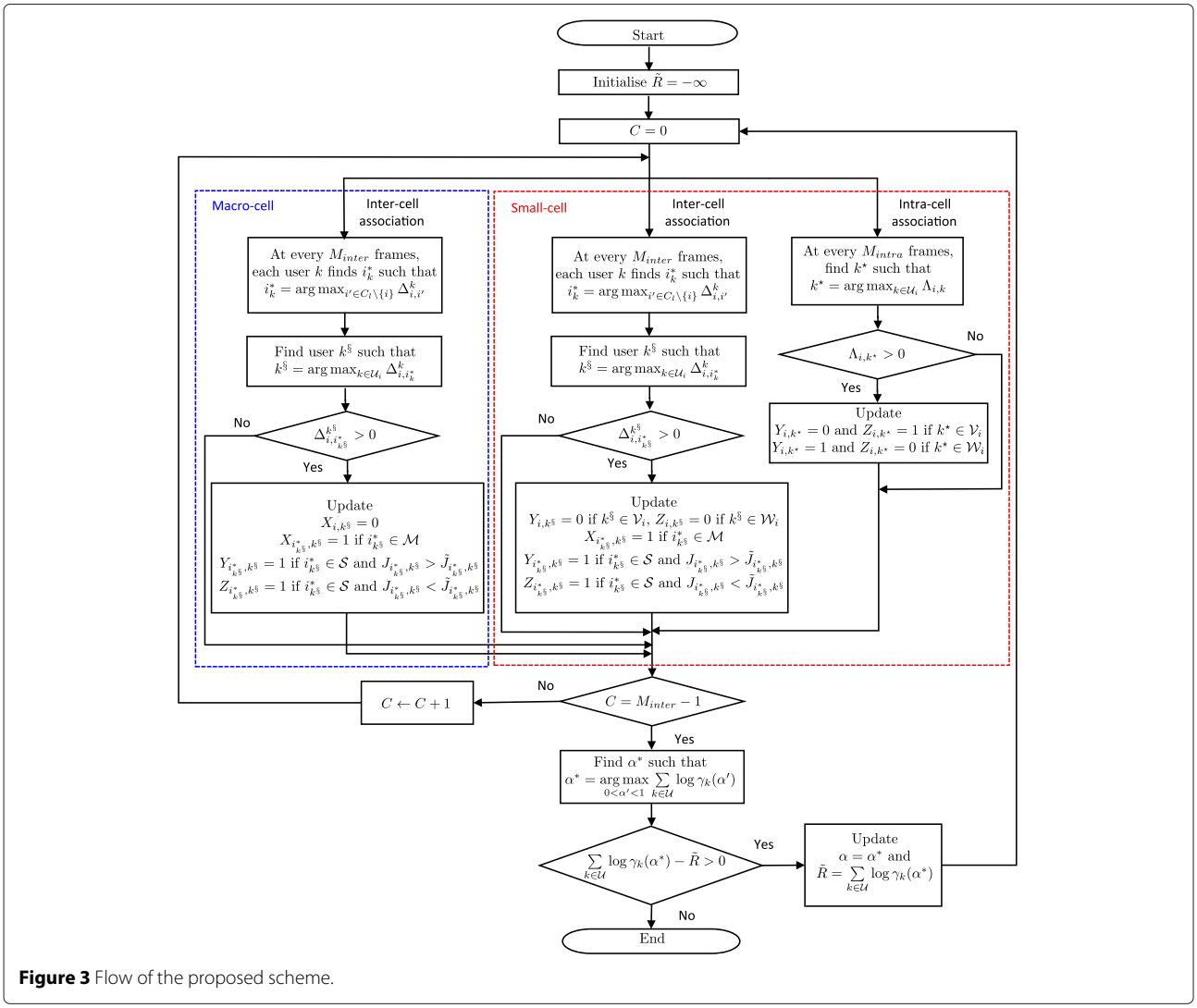

Figure 3 Flow of the proposed scheme.

4.3 ABS rate update

After every cell completes its user association process, the ABS rate is updated such that

$$\alpha^* = \arg\max_{0<\alpha'<1} \sum_{k\in\mathcal{U}} \log \gamma_k\left(\alpha'\right). \tag{39}$$

Then, the user association processes as described in Section 4.2 are performed again with the value of α^*. The recursive process that combines user associations and the ABS rate determination iterates as long as the value of network-wide PF, i.e., $\sum_{k\in\mathcal{U}} \log \gamma_k(\alpha^*)$, at the current iteration is larger than that of the previous iteration.

To perform the ABS rate update, it is required that the information required to estimate the throughput of all users in (39) should be provided to the central controller. Thus, to reduce this complexity, we might i) increase the period of ABS update, which increases the convergence time, or ii) perform the ABS update using only some of cells, which results in performance loss.

Finally, Figure 3 shows the overall flow of the proposed scheme.

4.4 Extension to N-tier cellular network

Consider the case when the inter-cell user association is employed in a N-tier cellular network which consists of N spatially and spectrally coexisting tiers, and each tier is distinguished by its transmit power, cell density, and data rate [29]. In this case, the information required to perform the inter-cell user association might exchange among the cells in only some tiers. For example, it is not allowed that a user is offloaded to a cell for a closed subscriber group (CSG) if it is not a member of the corresponding CSG [30]. It means that a given user in a cell might not be offloaded to some neighboring cells. This limits the performance improvement compared to the case when the information exchanges among all cells in all tiers.

Also, when only macro-cells employ ABS, the proposed ABS rate update can be employed without any modification. However, when the ABS rate is different for each tier in a N-tier cellular network, it is not easy to determine the ABS rates for all tiers maximizing the network PF at once. The simplest way to update the ABS rates is that the ABS rate for an arbitrary chosen tier is updated every after all cells complete their user association processes. Therefore, a further work is required to develop the ABS rate update algorithm supporting different ABS rates in a N-tier cellular network.

4.5 Power control aspects

Although we have considered that the transmit power levels of the cells are fixed, it is expected that the performance can be further improved if each cell's transmission power is adaptively controlled. However, it is very hard to determine the optimal power levels of the cells to improve network PF. The reason is that increasing the power in one cell increases interference in the other and vice versa, making most of the cases a non-convex or non-concave problem. To solve this problem, we may combine the power control process with the association process and the ABS rate update process in a recursive manner. The power control process is described as follows.

Assume that each cell can select its power level from a set of multiple power levels $\tilde{\mathcal{P}}$ that is common to all cells. Under this setup, at every given frame, a given cell estimates the sum of log throughputs of the users within itself for all possible combinations of transmit power levels of itself and its neighboring cells. Then, the information on the sum of log user throughputs is exchanged among them. Using the received information received from its neighboring cells, the cell evaluates the sum of log throughputs of the users in itself and its neighboring cells. Finally, the cell selects a combination of power levels that maximizes the sum of log throughputs of the users in itself and its neighboring cells and applies the power levels on itself and its neighboring cells. Similar to the association

Table 1 Summary of simulation assumptions

Parameter	Value
Number of macro-cell sites	19
Inter-site distance of macro-cell site	500 m
Number of sectors (macro-cell) per macro-cell site	3
Number of small-cells per macro-cell	4
Number of users per macro-cell	60
Transmit antenna pattern of macro-cell	70° 3-dB beamwidth, 18 dBi
Transmit antenna pattern of small-cell	Omnidirectional, 5 dBi
Receive antenna pattern of user	Omnidirectional, 0 dBi
System bandwidth	9 MHz
PRB bandwidth	180 kHz
Thermal noise density	−174 dBm/Hz
Noise figure at user	7 dB
Transmit power of macro-cell	46 dBm
Transmit power of small-cell	37 dBm
Standard deviation of shadowing	10 dB
Shadowing correlation	0.5
Penetration loss	20 dB
Minimum distance between macro- and small-cells	75 m
Minimum distance between small-cells	40 m
Minimum distance between macro-cell and user	35 m
Minimum distance between small-cell and user	10 m

process and the ABS rate update process, the power control process for each cell is iterated every M_p frames. The detailed process is described as follows.

First, define $\mathcal{P}_{i,j}$ as power level set j for cell i and each power level in this set is employed at cell i and cells in $\mathcal{N}_i$, where $\mathcal{N}_i$ denotes a set of neighboring cells of cell i. In particular, $\mathcal{P}_{i,j}$ consists of $(|\mathcal{N}_i|+1)$ power levels, each of which is chosen from power level set $\tilde{\mathcal{P}}$. Therefore, the number of all possible sets of $\{\mathcal{P}_{i,j}\}$ is $(|\tilde{\mathcal{P}}|)^{(|\mathcal{N}_i|+1)}$. With this, at every M_p frame, cell i evaluates

$$\Theta_i(\mathcal{P}_{i,j}) = \sum_{k\in\mathcal{U}_i} \log\left(\gamma_k(\mathcal{P}_{i,j})\right) \tag{40}$$

and cell i' in $\mathcal{N}_i$ evaluates

$$\Theta_{i'}(\mathcal{P}_{i,j}) = \sum_{k\in\mathcal{U}_{i'}} \log\left(\gamma_k(\mathcal{P}_{i,j})\right) \tag{41}$$

where $\gamma_k(\mathcal{P}_{i,j})$ denotes user k's throughput provided that cells of $\mathcal{N}_i \cup \{i\}$ employ the power levels of $\mathcal{P}_{i,j}$. After exchanging $\{\Theta_i(\mathcal{P}_{i,j})\}$ with cells in $\mathcal{N}_i$, cell i finds j^* such that

$$j^* = \arg\max_j \sum_{i'\in\mathcal{N}_i\cup\{i\}} \Theta_{i'}(\mathcal{P}_{i,j}). \tag{42}$$

Then, the power levels in $\mathcal{P}_{i,j^*}$ are employed at cells in $\mathcal{N}_i \cup \{i\}$. After every cell performs the power control process once, the association process and the ABS rate update process are performed sequentially.

As more additional research is required to combine the power control with the current proposed scheme, we would like to earmark the work of combining the power control process with the proposed scheme as a future research topic.

5 System-level simulation results

5.1 Simulation assumptions

To evaluate the average system performance, we employ most of the simulation assumptions suggested in [31]. In particular, we consider a wraparound structure consisting of 19 hexagonal macro-cell sites, each having three sectors. Note that each sector in a macro-cell site represents a macro-cell. In this configuration, the inter-site distance is 500 m and each macro-cell employs an 18-dBi gain directional antenna with 70° horizontal 3-dB beamwidth, 10° vertical 3-dB beamwidth, 15° down-tilt angle, and 32 m in height. In each macro-cell's geographical area, $L = 4$ small cells are uniformly distributed unless stated otherwise. Also, each small cell's minimum distances to its overlaid macro-cell and the other small cells are 75 and 40 m, respectively. Each small cell employs an omnidirectional antenna with 5-dBi antenna gain, and each user

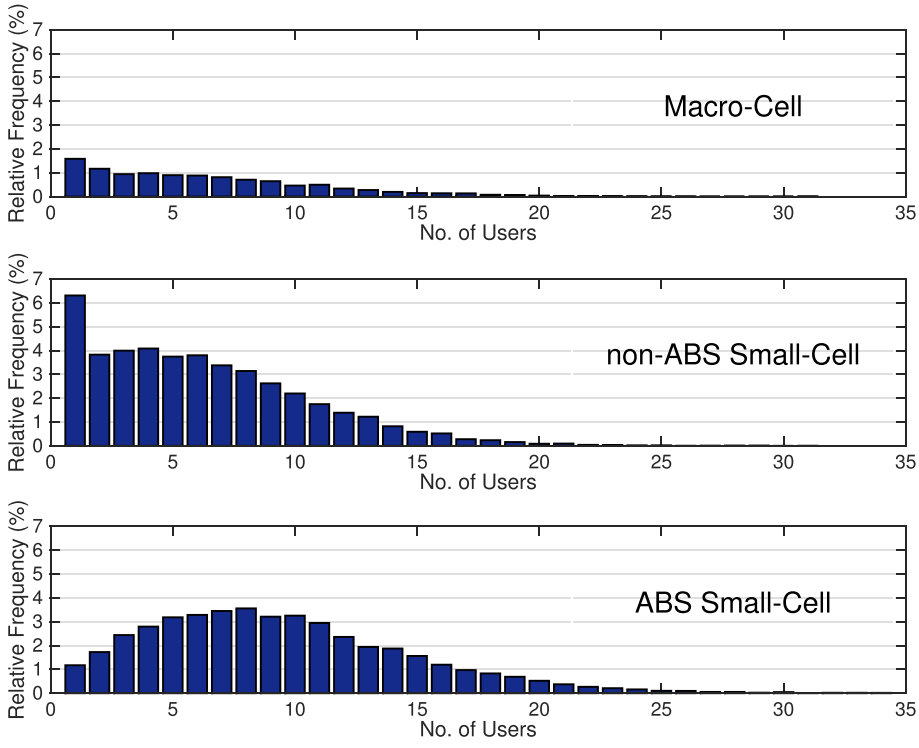

Figure 4 Histograms of numbers of users associated with cells for CRE bias of 15 dB.

employs an omnidirectional antenna with 0-dBi gain and 1.5 m in height.

Each cell employs an OFDMA system using a 9-MHz bandwidth and 10-ms frame length. Each PRB occupies 180 kHz in frequency [15]. Assume that each PRB experiences flat fading while it is i.i.d. between PRBs. We perform the scheduling process on each PRB independently as described in Section 3. The transmit powers of macro-cells and small cells are 46 and 37 dBm, respectively. The thermal noise density is −174 dBm/Hz, and the noise figure is 7 dB. The number of users per one geographical macro-cell area is 60.

The clustered user placement for small cells is employed as shown in [31]. In particular, $\left(\frac{1}{15}\right)$ of the users in each macro-cell area are uniformly distributed within the 40-m radius of each underlying small cell. In addition, these users are static and distributed in such a way as to satisfy the minimum distance to a macro-cell of 35 m and

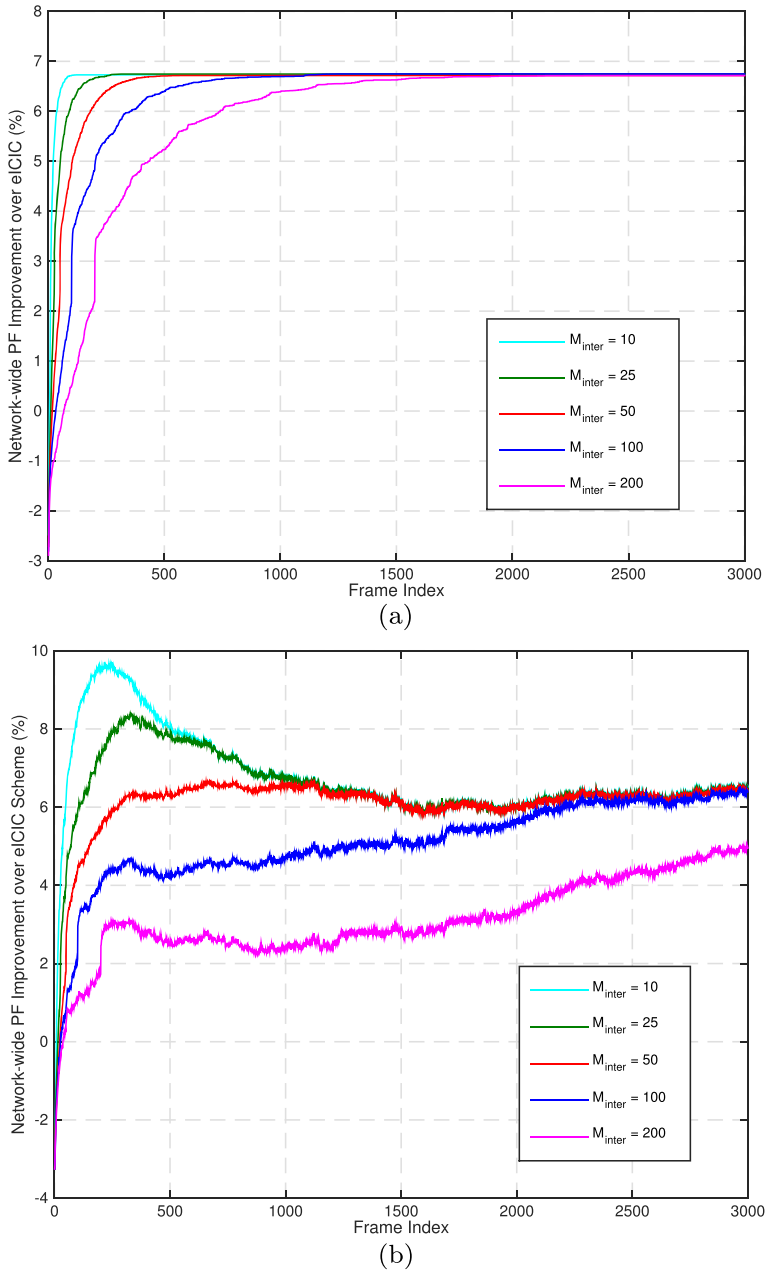

Figure 5 Network-wide PF improvement over eICIC scheme. With CRE bias of 15 dB under static condition **(a)** and mobility condition (3 km/h) **(b)**.

the minimum distance to a small cell of 10 m. The rest of the users, i.e., $\left(1-\frac{L}{15}\right)$ of all users, are assumed to be static and uniformly distributed in the whole area.

We employ the path-loss models in [31]. In particular, for a given distance d (in km) between a macro-cell and a user, the path-loss model is $\mathrm{PL}(d) = 128.1 + 37.6\log_{10}(d)$ (dB). In addition, for a given distance d between a small cell and a user, the path-loss model is $\mathrm{PL}(d) = 140.7 + 36.7\log_{10}(d)$ (dB). The shadowing standard deviation is 10 dB, and the penetration loss is 20 dB. Besides, the shadowing correlation between different cells is 0.5, while the shadowing between macro-cells at the same macro-cell site is identical. The intra-cell user association interval $M_{\mathrm{intra}} = 5$ is employed. Each user can be associated with one of $K = 5$ candidate macro-cells and $\tilde{K} = 5$ candidate small cells as suggested in (34). Table 1 summarizes the simulation assumptions.

To get the average performance, we numerically averaged over 100 different simulated realizations of the locations of small cells and users. To fairly compare the performance, we assumed that the eICIC scheme employs the same scheduling process described in Section 3 and determines the ABS rate such that the network-wide PF is maximized for a given value of CRE bias unless stated otherwise [9,10].

5.2 Numerical results

Performance metrics: We present the performance metrics to compare the performance of the proposed scheme and eICIC schemes as follows. The xth percentile user throughput indicates the value of user throughput below which a given x% of user throughputs fall. For example, the 5th percentile user throughput can be interpreted that 95% of the time, an arbitrary user can have the value of user throughput which is larger than that of the 5th percentile user throughput. Alternatively, 95% of users have the values of user throughputs larger that of the 5th percentile user throughput.

Recalling that a logarithmic function is employed as the utility function in Section 4.1, the network PF can be noticeably increased if the throughput of the users with low rates is improved rather than that of the users with high rates. Therefore, we are mainly interested in the low user-throughput performance improvement, i.e., the 5th percentile user throughput (i.e., the low user throughput) improvement. With this aim, we performed the simulations of the eICIC scheme for CRE bias values of $\{3, 6, 9, 12, 15, 18, 21, 24\}$ (in dB) for $L = 4$, and we found that the low user throughput of the eICIC scheme is maximized with a CRE bias of 15 dB. Therefore, we employ the CRE bias value of 15 dB for the eICIC scheme unless stated otherwise.

In addition, Jain's index is defined as [32]

$$\mathbb{J} = \frac{\left(\sum_{k\in\mathcal{U}} \gamma_k\right)^2}{|\mathcal{U}| \sum_{k\in\mathcal{U}} \gamma_k^2}. \tag{43}$$

Note that $\mathbb{J}$ ranges between $\frac{1}{|\mathcal{U}|}$ representing the least fair case and 1 representing the best fair case in which all users' throughputs are the same.

Numbers of users associated with cells: As discussed in Section 1, we argue that the numbers of victim and non-victim users are quite different from cell to cell. To support this argument, Figure 4 shows the histograms of the numbers of users associated with cells in the eICIC scheme with a CRE bias value of 15 dB. As expected, we see that the numbers of associated users are quite different from cell to cell. This means that a resource may be wasted when non-ABS and ABS are strictly assigned to non-victim and victim users, respectively.

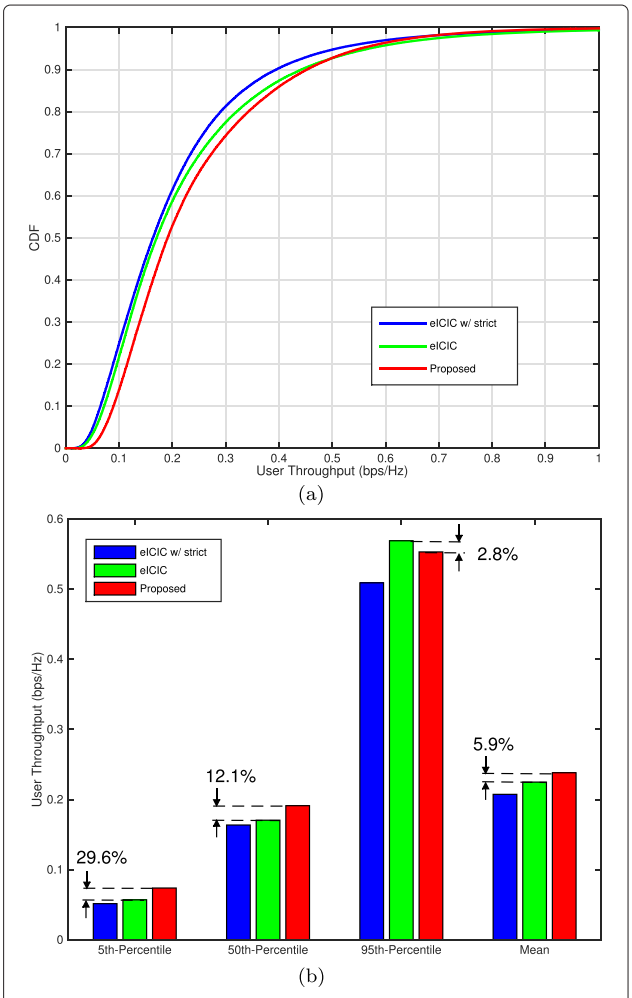

Figure 6 Empirical CDF of user throughputs **(a)** and various percentile throughputs **(b)**. The eICIC schemes employ a CRE bias of 15 dB.

Convergence properties: To investigate the number of frames required to converge the performance, Figure 5a,b shows the improvement of network-wide PF over the eICIC scheme with a CRE bias of 15 dB under static and mobility conditions, respectively. Note that, in the mobility condition, each user moves in a randomly chosen fixed direction at a speed of 3 km/h. In Figure 5a, the value of network PF converges quickly as M_{inter} decreases. For instance, approximately 500 frames are required to converge for $M_{\text{inter}} = 50$ under the static condition. The differences in the converged values are negligible regardless of M_{inter}. Consequently, considering the computation load and the convergence speed, we employ $M_{\text{inter}} = 50$ under the static condition in the following results.

In Figure 5b, it is observed that approximately two times longer time is required to converge the performance compared to those of the static case for $M_{\text{inter}} \leq 100$. However, the performance is maintained after converging the performance. On the other hand, for $M_{\text{inter}} = 200$, the performance degrades noticeably because M_{inter} is too

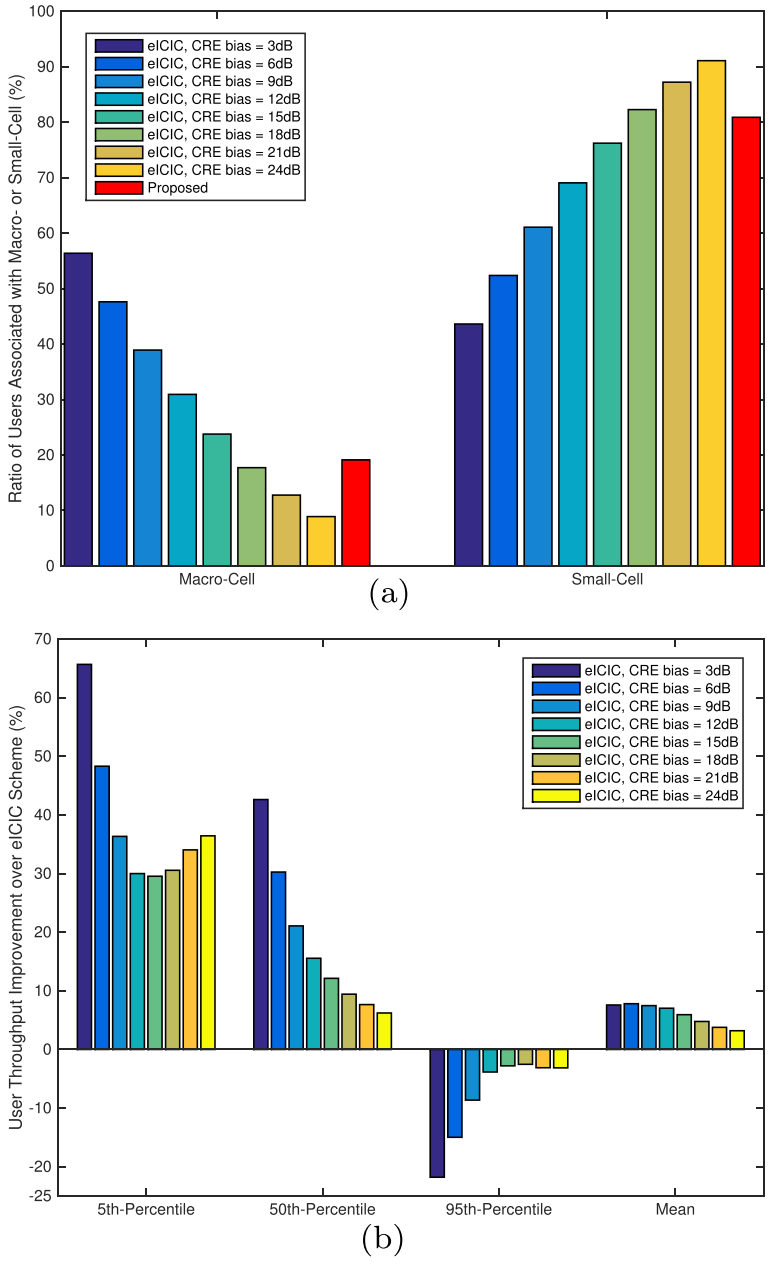

Figure 7 User association comparisons. Ratio of users associated with macro- or small cell **(a)** and user throughput improvement over eICIC schemes with various CRE biases **(b)**.

large to support the mobile users adequately. This means that the value of M_{inter} needs to be chosen carefully under the mobility condition.

User throughput improvements: Figure 6a,b shows the empirical CDF of user throughputs and various percentile throughputs, respectively. In these figures, the eICIC scheme employs a CRE bias of 15 dB and 'eICIC w/ strict' represents the eICIC scheme with the strict assignment policy in which ABS and non-ABS are strictly assigned to victim and non-victim users, respectively.

From these figures, we first notice that the eICIC scheme with the sharing assignment outperforms the eICIC scheme with the strict assignment. This implies that the user throughput performance is noticeably improved by simply allowing ABS to be shared among all users in a given small cell. In addition, we observe that the proposed scheme improves most users' throughput performances compared to those of eICIC schemes. In particular, in Figure 6b, the 5th percentile user throughput (i.e., the low user throughput) is improved by approximately 29.6% compared to that of the eICIC scheme with the sharing assignment. However, the 95th percentile user throughput (i.e., the high user throughput) is slightly decreased by approximately 2.8%. The reason is that the proposed scheme increases the network PF by mainly improving the low user-throughput performance.

User association comparisons: Figure 7a,b shows the ratio of users associated with macro-cells or small cells and the user throughput performance improvement over eICIC schemes with various CRE biases, respectively. Although a too large value of CRE bias for the eICIC scheme is unfeasible, we employ a wide range of CRE biases to investigate the performance trend. In Figure 7a, as the value of CRE bias increases, we find that more users from macro-cells are offloaded to small cells in the eICIC scheme. However, in Figure 7b, the 5th percentile user throughput (i.e., the low user throughput) is significantly degraded because too many macro-cell users are offloaded to small cells by simply comparing the received power differences.

In addition, we investigate fairness in the user throughput performance using Jain's index. Table 2 shows Jain's index improvement over eICIC schemes with the same setups in Figure 7a,b. From these results, we see that the proposed scheme also improves fairness in terms of user throughput performance.

Table 2 Jain's index improvement over eICIC schemes

	CRE bias for eICIC (dB)			
	3	**9**	**15**	**21**
Jain's index improvement (%)	57.0	27.0	15.4	11.5

Number of small cells per macro-cell: Figure 8 shows the user throughput with the number of small cells per macro-cell geographical area, L. Solid and dashed lines represent the proposed scheme and eICIC scheme, respectively. The eICIC scheme also employs a CRE bias value that maximizes the 5th percentile user throughput (i.e., the low user throughput) among $\{3, 6, \ldots, 24\}$ (in dB) for each value of L.

Compared to the eICIC scheme, it is observed that the performance of the 95th percentile user throughput (i.e., the high user throughput) is degraded. The reason is that the proposed scheme increases the network PF by mainly improving low user-throughput performance. However, it is observed that the proposed scheme improves the performance of the 5th percentile user throughput (i.e., the low user throughput) and the mean throughput regardless of L. This means that the proposed scheme finds suitable sets of user association and the ABS rate that improves the system performance under various system conditions.

6 Conclusions

In this paper, we proposed an LB scheme for two-tier cellular networks using the eICIC scheme under the CDF-based scheduling scheme that supports sharing ABS among all users in a given small cell. In particular, the user association between the cells and the ABS rate determination are recursively combined to improve the network-wide PF. System-level simulations demonstrated that the proposed scheme improves the network-wide PF compared to the eICIC scheme regardless of the number of macro-cells offloaded to the small cells.

We have considered the case where every cell transmits with a fixed transmission power. However, it is expected that the overall system performance could be

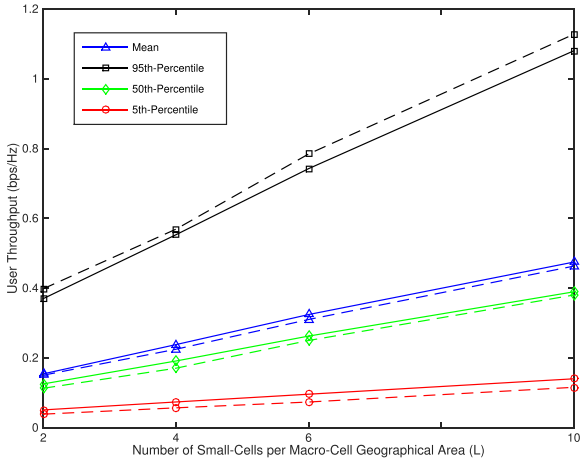

Figure 8 User throughput with the number of small cells per macro-cell geographical area. Solid and dashed lines represent the proposed scheme and eICIC scheme, respectively.

further improved if each cell's transmission power is adaptively controlled. Therefore, an LB algorithm that includes the transmit power control process is a topic of interest for future work. In addition, as it has been known that multiple antenna transmission can further improve throughput performance, it is necessary to extend this work in the future to the case when spatial diversity mechanisms to schedule multiple users simultaneously are employed along with exploiting multiuser diversity from the CDF-based scheduling. Finally, we have considered only the time-domain interference management scenario for two-tier cellular networks without CA. However, it was demonstrated in [6] that the overall throughput performance of a two-tier network with CA is significantly improved if joint interference management and cell association is performed in both time and frequency domains. Therefore, future work will involve the development of an LB algorithm to improve the network PF for two-tier cellular networks with CA.

Competing interests
The authors declare that they have no competing interests.

Acknowledgements
This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2012R1A1A1001769).

Received: 14 November 2014 Accepted: 18 March 2015
Published online: 09 April 2015

References

1. Cisco, Cisco visual networking index: global mobile data traffic forecast update, 2012-2017. Technical report (February 2013)
2. A Damnjanovic, J Montojo, Y Wei, T Ji, T Luo, M Vajapeyam, T Yoo, O Song, D Malladi, A survey on 3GPP heterogeneous networks. IEEE Wireless Commun. **18**(3), 10–21 (2011)
3. D López-Pérez, I Güvenc, G de la Roche, M Kountouris, TQS Quek, J Zhang, Enhanced intercell interference coordination challenges in heterogeneous networks. IEEE Wireless Commun. **18**(3), 22–30 (2011)
4. X Lin, JG Andrews, A Ghosh, Modeling, analysis and design for carrier aggregation in heterogeneous cellular networks. IEEE Trans. Commun. **61**(9), 4002–4015 (2013)
5. L Lindbom, R Love, S Krishnamurthy, C Yao, N Miki, V Chandrasekhar, Enhanced inter-cell interference coordination for heterogeneous networks in LTE-advanced: a survey (2011). CoRR abs/1112.1344
6. M Simsek, M Bennis, I Guvenc, Learning based frequency- and time-domain inter-cell interference coordination in HetNets. IEEE Trans. Vehicular Technol (2015). in press
7. Y Wang, B Soret, KI Pedersen, in *Proc. IEEE International Conference on Communications*. Sensitivity study of optimal eICIC configurations in different heterogeneous network scenarios (Ottawa, ON, 2012), pp. 6792–6796
8. J Pang, J Wang, D Wang, G Shen, Q Jiang, J Liu, in *Proc. IEEE Wireless Commun. and Networking Conference*. Optimized time-domain resource partitioning for enhanced inter-cell interference coordination in heterogeneous networks (Shanghai, China, 2012), pp. 1613–1617
9. S Vasudevan, RN Pupala, 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 Commun. **12**(10), 4956–4969 (2013)
10. S Lembo, P Lunden, O Tirkkonen, K Valkealahti, in *Proc. IEEE International Conference on Communications*. Optimal muting ratio for enhanced inter-cell interference coordination (eICIC) in HetNets (Budapest, Hungary, 2013), pp. 1145–1149
11. T Bu, L Li, R Ramjee, in *Proc. IEEE Conference on Computer Communications*. Generalized proportional fair scheduling in third generation wireless data networks (Barcelona, Spain, 2006), pp. 1–12
12. K Son, S Chong, G de Veciana, Dynamic association for load balancing and interference avoidance in multi-cell networks. IEEE Trans. Wireless Commun. **8**(7), 3566–3576 (2009)
13. A Weber, O Stanze, in *Proc. IEEE International Conference on Communications*. Scheduling strategies for HetNets using eICIC (Ottawa, Canada, 2012), pp. 6787–6791
14. KI Pedersen, Y Wang, S Strzyz, F Frederiksen, Enhanced inter-cell interference coordination in co-channel multi-layer LTE-advanced networks. IEEE Wireless Commun. **20**(3), 120–127 (2013)
15. H Holma, A Toskala, *LTE for UMTS-OFDMA and SC-FDMA Based Radio Access*. (John Wiley & Sons Ltd, West Sussex, UK, 2009)
16. AM Mourad, L Brunel, A Okazaki, U Salim, in *Proc. IEEE Vehicular Technology Conference*. Channel quality indicator estimation for OFDMA systems in the downlink (Dublin, Ireland, 2007), pp. 1771–1775
17. R Giuliano, F Mazzenga, Exponential effective SINR approximations for OFDM/OFDMA-based cellular system planning. IEEE Trans. Wireless Commun. **8**(9), 4434–4439 (2009)
18. D Park, H Seo, H Kwon, B Lee, Wireless packet scheduling based on the cumulative distribution function of user transmission rates. IEEE Trans. Commun. **53**(11), 1919–1929 (2005)
19. J-G Choi, S Bahk, Cell-throughput analysis of the proportional fair scheduler in the single-cell environment. IEEE Trans. Vehicular Technol. **56**(2), 766–778 (2007)
20. B Soret, H Wang, KI Pedersen, C Rosa, Multicell cooperation for LTE-advanced heterogeneous network scenarios. IEEE Wireless Commun. **20**(1), 27–34 (2013)
21. B Soret, Y Wang, KI Pedersen, in *Proc. IEEE International Conference on Communications*. CRS interference cancellation in heterogeneous networks for LTE-advanced downlink (Ottawa, Canada, 2012), pp. 6797–6801
22. M Sharif, B Hassibi, On the capacity of MIMO broadcast channels with partial side information. IEEE Trans. Inf. Theory. **51**(2), 506–522 (2005)
23. T Bonald, in *Proc. European Wireless*. A score-based opportunistic scheduler for fading radio channels (Barcelona, Spain, 2004), pp. 1–7
24. S Patil, G de Veciana, Measurement-based opportunistic scheduling for heterogenous wireless systems. **57**(9), 2745–2753 (2009)
25. P Svedman, SK Wilson, LJ Jr Cimini, B Ottersten, Opportunistic beamforming and scheduling for OFDMA systems. **55**(5), 941–952 (2007)
26. S Stanczak, M Wiczanowski, H Boche, *Fundamentals of Resource Allocation in Wireless Networks: Theory and Algorithms*. (Springer, Berlin Heidelberg, 2009)
27. R Knopp, PA Humblet, in *Proc. IEEE International Conference on Communications*. Information capacity and power control in single-cell multiuser communications, vol. 1 (Seattle, WA, 1995), pp. 331–335
28. HJ Kushner, PA Whiting, Convergence of proportional-fair sharing algorithms under general conditions. IEEE Trans. Wireless Commun. **3**(4), 1250–1259 (2004)
29. HS Dhillon, RK Ganti, F Baccelli, JG Andrews, Modeling and analysis of K-tier downlink heterogeneous cellular networks. IEEE J. Selected Areas Commun. **30**(3), 550–560 (2012)
30. RQ Hu, Y Qian, *Heterogeneous Cellular Networks*. (John Wiley & Sons Ltd, West Sussex, UK, 2013)
31. 3GPP TR 36.814 v9.0.0., Further advancement for E-UTRA physical layer aspects (Release 9). Technical report, 3GPP (2010)
32. R Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation and Modeling*. (Wiley, New York, 1991)