Dynamic Association for Load Balancing and Interference Avoidance in Multi-Cell Networks

Kyuho Son, *Student Member, IEEE*, Song Chong, *Member, IEEE*, and Gustavo de Veciana, *Fellow, IEEE*

***Abstract*—Next-generation cellular networks will provide higher cell capacity by adopting advanced physical layer techniques and broader bandwidth. Even in such networks, boundary users would suffer from low throughput due to severe inter-cell interference and unbalanced user distributions among cells, unless additional schemes to mitigate this problem are employed. In this paper, we tackle this problem by jointly optimizing partial frequency reuse and load-balancing schemes in a multi-cell network. We formulate this problem as a network-wide utility maximization problem and propose optimal offline and practical online algorithms to solve this. Our online algorithm turns out to be a simple mixture of inter- and intra-cell handover mechanisms for existing users and user association control and cell-site selection mechanisms for newly arriving users. A remarkable feature of the proposed algorithm is that it uses a notion of *expected throughput* as the decision making metric, as opposed to signal strength in conventional systems. Extensive simulations demonstrate that our online algorithm can not only closely approximate network-wide proportional fairness but also provide two types of gain, *interference avoidance* gain and *load balancing* gain, which yield $20{\sim}100\%$ throughput improvement of boundary users (depending on traffic load distribution), while not penalizing total system throughput. We also demonstrate that this improvement cannot be achieved by conventional systems using universal frequency reuse and signal strength as the decision making metric.**

***Index Terms*—Inter-cell interference (ICI), ICI avoidance, load balancing, association, network-wide proportional fairness, multi-cell network, network utility maximization.**

I. Introduction

To support higher data rates, several next-generation wireless broadband systems based on OFDMA (Orthogonal Frequency Division Multiple Access) are currently being standardized including: IEEE 802.16/ WiMAX (Wireless Interoperability for Microwave Access) [1] and 3GPP LTE (Long Term Evolution) [2]. In these promising systems, downlink signals originating from the same base station (BS) do not interfere with each other because subbands are allocated orthogonally across users. By contrast, signals from different BSs may interfere and as a consequence, inter-cell interference (ICI) is a major source of performance degradation. In particular, users at the cell edge (or simply, boundary users) may have low signal to interference plus noise ratio (SINR) because such locations suffer severely from ICI. In addition, in real-world systems, users are not evenly distributed across cells, yielding load imbalance between cells, which is the second major source of system-wide performance degradation. Especially, the performance of boundary users in a hot-spot cell is mostly affected by this load imbalance so that they might be unable to get services. To guarantee a reasonable system-wide quality of service (QoS) irrespective of users' geographical locations and enhance cell coverage, effective ICI mitigation and load balancing schemes are required.

There has been several researches on multi-cell networks, which can be classified into two types. The first is a traditional load balancing problem [3]–[6], and the second is ICI mitigation problem [7]–[11] attracting much attention recently. However, little work has been done to jointly consider both load balancing and ICI mitigation so far.

For load balancing, a cell breathing technique was investigated in [3] and [4]. It contracts (or expands) the coverage of congested (or under-loaded) cells by reducing (or raising) the power level, and therefore the load becomes more balanced. Sang *et al.* [5] proposed an integrated framework consisting of a MAC-layer cell breathing technique and load-aware handover/cell-site selection to deal with load balancing. Bu *et al.* [6] were the first to rigorously consider a formulation of network-wide proportional fairness (PF) [12] in a multi-cell network where associations between users and BSs are decision variables. They showed that the general problem is NP-hard and proposed a heuristic algorithm to approximately solve the problem. However, none of these works had considered ICI mitigation schemes in conjunction with their proposed load balancing schemes, which have an extra potential to further increase the system-wide performance.

For mitigating ICI, a brute-force approach is the use of traditional frequency reuse schemes [7] with a reuse factor greater than one. The more ICI is mitigated by using the higher reuse factor, the less resource is available at each cell. Frequency reuse will be effective in improving the throughput of low SINR users at the cell edge but less effective to high SINR users in the inner region of the cell so that it can waste frequency resources unless selectively applied.

Manuscript received October 15, 2007; revised March 10, 2008, October 2, 2008, and February 21, 2009; accepted April 16, 2009. The associate editor coordinating the review of this paper and approving it for publication was M. Chiang.

K. Son and S. Chong are with the School of Electrical Engineering and Computer Science at Korea Advanced Institute of Science and Technology (KAIST), Daejeon 305-701, Korea (e-mail: skio@netsys.kaist.ac.kr; song@ee.kaist.ac.kr).

G. de Veciana is with the Wireless Networking and Communications Group (WNCG), Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712 USA (e-mail: gustavo@ece.utexas.edu).

Some part of this work was presented at the WiOpt 2007 conference.

This research was supported in part by the Ministry of Knowledge Economy, Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Advancement) (IITA-2009-C1090-0902-0037) and the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MEST) (No. R01-2006-000-10753-0). The work of G. de Veciana was also supported in part by National Science Foundation Grant CNS-0721532.

Digital Object Identifier 10.1109/TWC.2009.071140

More elaborate work on mitigating ICI has been done by [8]–[11]. Li *et al.* [8] formulated an optimization problem to maximize the system throughput in a multi-cell OFDMA system. In their solution, a RNC (Radio Network Controller) coordinates the interference among multiple cells so that each cell utilizes not all but around 80% of its subbands to avoid the dominant ICI. Bonald *et al.* [9] examined the capacity gains achievable by inter-cell time resource sharing in CDMA/HDR systems. They formulated an optimization problem which coordinates the activity phases of BSs so as to provide higher data rate for boundary users by mitigating ICI. Even though cell selection for load balancing is also considered and studied in [11] in a three-cell system with an arbitrary traffic distribution, numerical results are only available for limited cases, that is to say, they do not give a clear answer to general multi-cell networks with heterogeneous traffic distribution. In both the [8], [9], it is noteworthy that using only partial resources (frequency and time, respectively) is essential to obtain potential performance gains associated with mitigating ICI.

In our work, partial frequency reuse (PFR), a practical ICI mitigation scheme is considered in conjunction with load balancing. Unlike traditional frequency reuse schemes, where all users share the same reuse factor, such as 1 (universal), 3, 4 or 7, PFR allows users in different channel conditions to enjoy different reuse factors. In this scheme, the entire system bandwidth is divided into two groups of subbands: *inner band* (with universal reuse factor) and *outer band* (with reuse factor greater than one). Each cell uses all the subbands in the inner band and a portion of the subbands in the outer band. For example, in type 1[1] cells in Fig. 1, users in the inner region of the cell are allowed to use the entire inner band and users at the cell edge are allowed to use a portion of the outer band, i.e., band O1. According to the recent technical report of 3GPP LTE [13], ICI mitigation approaches are classified into three types: 1) inter-cell interference randomization, 2) inter-cell interference cancelation and 3) inter-cell interference coordination/avoidance. The PFR scheme explained above belongs to the last category, i.e., aims at avoiding ICI by selectively restricting downlink frequency resources in a coordinated way between multiple cells.

In this paper, we extend the Bu's work [6] to multi-cell networks using PFR and jointly optimize load balancing and ICI coordination/avoidance to achieve network-wide proportional fairness. We assume that each BS has limited frequency resources based on an ICI pre-coordination of PFR and independently runs a proportional fair scheduler. In this setting, we concentrate on the following association (long-term binding relationship) problems:

- *Inter-cell association*: To which BS should each user be associated?
- *Intra-cell association*: Should a user be allocated to the inner or outer band?

The remainder of this paper is organized as follows. In Section II, we present our system model. In Section III, we present a general user assignment problem. In Section IV, we present a proportional fair user association problem and an offline optimal algorithm. In Section V, we present a practical online algorithm that uses the *expected throughput* as a key metric in making association decisions. In Section VI, we discuss our simulations for a two-tier cellular system, which demonstrate that our online algorithm can closely achieve network-wide PF. We also show that they perform better than conventional systems with universal frequency reuse where each user connects to the BS offering the best signal strength. In Section VII, we present further discussion and conclusions.

[1]Type k cell is the group of cells using k-th outer band

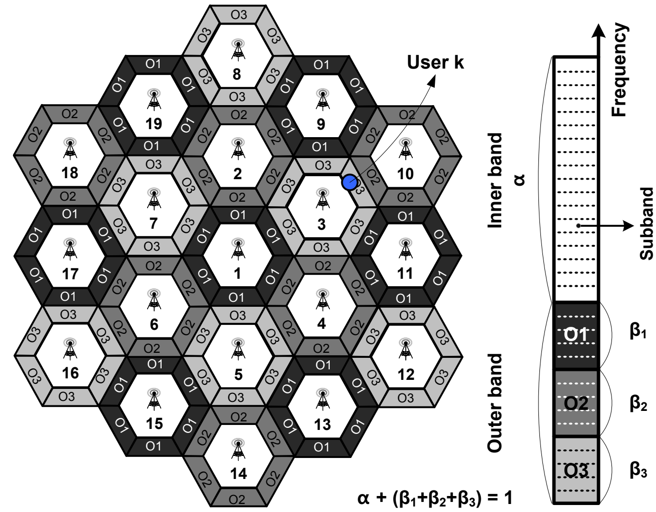


Fig. 1. Examples of frequency partitioning: the first-tier ICI coordination/avoidance

II. System Model

We consider a downlink network with PFR consisting of a set of BSs and users. We shall make the following assumptions for the remainder of this paper:

1) Each user generates persistent data traffic and has an infinitely backlogged queue.
2) Every time slot, each user can be associated with only one BS using either inner or outer band.
3) Each user knows instantaneous achievable rates of both inner and outer bands from all BSs to itself.
4) Each BS knows instantaneous achievable rates of both inner and outer bands from itself to all users.
5) Each time slot, each BS schedules two users, one for inner band and the other for outer band.
6) Each BS allocates power equally to all the subbands being used.

Remark: Assumptions 3) and 4) require network-wide channel estimations and feedbacks. As you will see in Section V, however, our proposed online algorithm substantially reduces this overhead by considering neighboring BSs only. The equal power assumption in 6) has been frequently used for implementation simplicity as well as analytical tractability in downlink resource allocation problems [8], [14]. Moreover, equal power allocation is near optimal in many cases especially in high SINR regime [15], [16].

A. Resource Partitioning

First of all, the PFR is a network-wide agreement in that all the cells in the network must be subject to its resource partitioning pattern. Fig. 1 depicts an example of frequency partitioning[2] in a regular multi-cell network using PFR. In this example, there are three types of cells and the entire bandwidth is divided into two subband groups: inner band (α) and outer band (β_1, β_2 and β_3), where $\alpha + (\beta_1+\beta_2+\beta_3) = 1$. Type i cells are the ones that serve their outer users using only β_i portion of subbands whereas all the cells in the network share the α portion of subbands in serving their inner users.

In this paper, we assume that a proper resource partitioning is given and fixed based on average statistics collected across cells and with time for a reasonably long period. One could also adapt the resource partitioning to the changes of this average statistics but this long-term adaptive resource partitioning problem is not the focus of our paper. Our focus is dynamic association problem that can maximize the long-term network-wide utility for a given resource partitioning no matter what the resource partitioning is. It is apparent that joint optimization of long-term adaptive resource partitioning and association strategy will enlarge the achievable region of long-term network-wide utility but we leave this as a future study.

B. Link Model

The sets of BSs and users in the network are denoted by N and K, respectively. The set of bands, consisting of inner and outer bands, is denoted by $B = \{in, out\}$. The received SINR at time slot t for user $k \in K$ from BS $n \in N$ on band $b \in B$ can be written as:

$$SINR_{n,k}^b(t) = \frac{g_{n,k}^b(t)p_n^b}{\phi\, g_{n,k}^b(t)p_n^b + \sum_{j \in L_n^b, j \neq n} g_{j,k}^b(t)p_j^b + N_0^b}, \quad (1)$$

where

- $g_{n,k}^b(t)p_n^b$ is the signal strength received by user k from BS n at time slot t with p_n^b and $g_{n,k}^b$ representing the transmit power of BS n on band b and the channel gain between BS n and user k on band b, respectively. The channel gain takes into account the path loss, log-normal shadowing and fast fading.
- L_n^b is the set of BSs allowed to use the same inner or outer band b as BS n.
- N_0^b is the additive white Gaussian noise (AWGN) on band b. Without loss of generality, we assume the noise level is the same for all users.
- ϕ is the orthogonality (or self-interference) factor that models transmitter and receiver non-linearities and limits the maximum SINR. For our simulations, we set ϕ to be 0.01 which correspondingly bounds the maximum SINR by 20dB.

[2]For ease of presentation, all boundaries of cells, inner and outer regions are depicted by straight lines. In reality, they would be irregular due to shadowing in the environment as well as the distribution of users. The inner and outer regions are logical concepts, which are dynamically changed depending on channel conditions and the distribution of users.

Given p_n^b, the instantaneous achievable rate at time slot t for user k from BS n on band b as given by $r_{n,k}^b(t) = BW^b \log_2\left(1 + SINR_{n,k}^b(t)\right)$ [bps], where BW^b is the bandwidth of band b. Note that once a resource partitioning is fixed, p_n^b is fixed due to Assumption 6). The set of instantaneous achievable rates of user k from BS n on band b, denote by $\left\{r_{n,k}^b(t),\ t \in \mathbb{Z}\right\}$ where $\mathbb{Z}$ is the set of nonnegative integers, is assumed to be a stationary ergodic process. These processes are independent, but not necessarily identically distributed across users, bands and BSs.

III. General User Assignment Problem

We shall start by presenting a general utility maximization problem in the multi-cell network setting. Our objective is to find the long-term throughput vector $\bar{\mathbf{R}} = \left(\bar{R}_k, k \in K\right)$ corresponding to resource allocation policy that maximizes the network-wide aggregate utility over a long-term achievable rate region $\mathcal{R}$:

$$\mathbf{P}: \qquad \max \quad \sum_{k \in K} U_k(\bar{R}_k) \quad (2)$$

$$\text{subsect to} \quad \bar{\mathbf{R}} \in \mathcal{R}, \quad (3)$$

where $U_k(\cdot)$ is an increasing, strictly concave and continuously differentiable utility function for user k. The set $\mathcal{R} \in \mathbb{R}_+^{|K|}$, the set of all achievable rate vectors over long-term, is shown to be a closed bounded convex set [17].

Now we describe a network-wide user assignment algorithm to achieve the optimal solution of **P**. We first denote by $\mathbf{I}(t) = \left(I_{n,k}^b(t) : n \in N, k \in K, b \in B\right)$ the user assignment indicator vector, i.e., $I_{n,k}^b(t) = 1$ when BS n assigns band b to user k at time slot t, and 0 otherwise. Since only one user can be selected in each BS n and band b for every time slot, we should have:

$$\sum_{k \in K} I_{n,k}^b(t) = 1, \quad \forall n \in N,\ \forall b \in B. \quad (4)$$

To find an optimal solution, we use a standard gradient-based algorithm that selects the achievable rate vector maximizing the sum of weighted rates where the weights are marginal utilities at each time slot. Then, it suffices to solve the following problem **P-assignment** at each time slot which determines the user assignment vector $\mathbf{I}(t)$:

P-assignment:

$$\max_{\mathbf{I}(t)} \quad \sum_{k \in K} U_k'(\bar{R}_k(t-1))R_k(t)$$

$$\text{subject to} \quad \sum_{k \in K} I_{n,k}^b(t) = 1, \quad \forall n \in N,\ \forall b \in B,$$

where $R_k(t) = \sum_{n \in N}\sum_{b \in B} I_{n,k}^b(t) r_{n,k}^b(t)$ be the data rate assigned to user k at time slot t and $\bar{R}_k(t) = \frac{1}{t}\sum_{\tau=1}^t R_k(\tau)$ $= \bar{R}_k(t-1) + \epsilon_t\left[R_k(t) - \bar{R}_k(t-1)\right]$ (by letting $\epsilon_t = 1/t$) is the long-term throughput for user k up to time slot t.

From the simple structure of the constraint in Eq. (4), we can develop the following user assignment algorithm, where each BS n selects the best user on each band b having the largest value of $U_k'(R_k(t-1))r_{n,k}^b(t)$ among all users $k \in K$ in the network. The proof of asymptotical convergence to the optimal solution is a slight extension to [17], [18] and is omitted here.

User assignment algorithm at the central node

$$I_{n,k}^b = \begin{cases} 1, & \text{if } k = \arg\max_{k \in K} U_k'(\bar{R}_k(t-1)) r_{n,k}^b(t), \\ 0, & \text{otherwise}, \end{cases} \tag{5}$$

Remark: In this algorithm, each user can be served from any BS on any band, and these may vary at every time slot. However, in the problem **P-association** described in the next section, we only focus on finding the long-term binding relationship, i.e., each user is served from a specific BS and band.

This optimal network-wide user assignment algorithm has implementation difficulties. Apart from the computational complexity of the algorithm, the central node running the algorithm needs to gather the following information from all users in the network - instantaneous data rates $r_{n,k}^b(t)$ of both inner and outer bands from all BSs $n \in N$. The total amount of feedbacks is quite large, i.e, $2|N||K|$, though they may be delivered along with wired backbone links. Furthermore, a series of tasks, including information feedback from users to the central node as well as the computation and the distribution of central node's decision, should be performed *within one time slot*, which makes its realization even more difficult in practice.

IV. Proportional Fair User Association Problem

In contrast to the centralized slot-by-slot user assignment algorithm in the previous section, *user scheduling in practice is typically undertaken by individual BSs independently* provided that associations between users and BSs are given. In this section, in order to take into account such autonomous feature in user scheduling, we only focus on finding optimal long-term user associations under the assumption that underlying intra-cell user schedulers are assumed to be given. Further, we limit our attention to the case of proportional fairness (PF), that is, all users have the same log utility function $U_k(\bar{R}_k) = \log(\bar{R}_k), \forall k$, and each BS independently executes the PF scheduler at every time slot.

Consider a network with a fixed number of users, i.e., no user arrivals or departures.[3] Denote by $\mathbf{X} = \left(X_{n,k}^b : n \in N, k \in K, b \in B\right)$ the association indicator vector, i.e., $X_{n,k}^b = 1$ when user k is associated with BS n on band b, and 0 otherwise. Since each user is associated with only one BS on either inner or outer band, we should have the following unique association constraint:

$$\sum_{n \in N} \sum_{b \in B} X_{n,k}^b = 1, \quad \forall k \in K. \tag{6}$$

For a given association $\mathbf{X}$, each BS n is assumed to run the following PF scheduler as an intra-cell scheduler to select the optimal user k_n^{b*} on each band $b \in B$ at every time slot t:

$$k_n^{b*}(t) = \arg\max_{k \in K_n^b} \frac{r_{n,k}^b(t)}{\bar{R}_k(t-1)}, \tag{7}$$

[3]Our online algorithm in the forthcoming section still works even if the system is dynamic where users are mobile, arrive and depart.

where $K_n^b = \left\{k \mid X_{n,k}^b = 1, k \in K\right\}$ is the set of all users who are associated with BS n on band b.

Following procedure analogous to that used in [18], the average long-term throughput $\bar{R}_k(t)$ as $t \to \infty$ corresponding to the above intra-cell scheduler can be written as follows:[4]

$$\bar{R}_k = \sum_{n \in N} \sum_{b \in B} X_{n,k}^b \left[\frac{G(Y_n^b)\ E[r_{n,k}^b]}{Y_n^b} \right], \tag{8}$$

where $E[r_{n,k}^b]$ is the expectation of $r_{n,k}^b$, i.e., the long-term average of instantaneous achievable data rate of user k from BS n on band b, and $Y_n^b = \sum_{k \in K} X_{n,k}^b$ is the number of users associated with BS n on band b; $G(y) = \sum_{k=1}^{y} \frac{1}{k}$ represent a multi-user diversity gain (scheduling gain) depending only on the number of users competing with the same resource. Note that the same result in Eq. (8) can be derived by another method [19] using the fact that PF scheduler assigns an equal fraction of slots to users, i.e., temporal fairness.

Therefore, we can formulate the following network-wide proportional fairness problem **P-association** for given intra-cell PF schedulers in Eq. (7):

P-association:

$$\begin{aligned} \max_{\mathbf{X}} \quad & \sum_{k \in K} \log\left(\bar{R}_k\right) \\ \text{subject to} \quad & \sum_{n \in N} \sum_{b \in B} X_{n,k}^b = 1, \quad k \in K, \\ & Y_n^b = \sum_{k \in K} X_{n,k}^b, \quad n \in N, b \in B, \\ & \bar{R}_k = \sum_{n \in N} \sum_{b \in B} X_{n,k}^b \left[\frac{G(Y_n^b) E[r_{n,k}^b]}{Y_n^b} \right], \ k \in K. \end{aligned}$$

Remark: This formulation is general enough to consider multiple bands (more than two, or even one), though we discuss only two bands, inner and outer bands, in this paper. In particular, if we consider a single band, i.e., $|B| = 1$, then this problem is reduced to that in Bu's work[6] that does not reflect the interference avoidance.

A. Offline algorithm

Now we present an offline algorithm that finds the optimal solution of the problem **P-association**. The Proposition 4.1 telling us the interesting property of the problem enable us to design an offline algorithm.

Proposition 4.1: If we fix the number of inner/outer users in each BS, then the problem **P-association** can be reduced to a maximum-weight bipartite matching (MWBM) problem which can be solved in polynomial time.

Proof: If $\mathbf{Y} = \left(Y_n^b : n \in N, b \in B\right)$ is fixed, then $w_{n,k}^b = G(Y_n^b)E[r_{n,k}^b]/Y_n^b$ is fixed. The problem is equivalent to finding a MWBM between virtual $2|N|$ BSs (each BS is split into two BSs, inner BS and outer BS) and $|K|$ users each with a nonnegative weight $w_{n,k}^b$. This can be solved in a polynomial time by the well-known Hungarian method $O(|K|^3)$ [20]. ■

[4]Mathematically, some assumptions are required to derived the Eq. (8): (i) all users have Rayleigh fading channels and (ii) the feasible rate is linear in the SINR.

If the number of BSs is a constant $|N|$, then the number of all possible configurations for $\mathbf{Y}$ is $O(|K|^{2|N|})$. For each $\mathbf{Y}$ configuration, we solve the above MWBM. Because both the number of all configurations and the computational complexity of MWBM are polynomial, the total running time of the proposed offline algorithm is also polynomial, $O(|K|^{2|N|+3})$. However, it is computationally too complex, e.g., $O(|K|^{2|N|+3}) = O(190^{41}) \approx 2.7 \times 10^{93}$ when the number of BSs is 19 (two-tier system) and each BS has only 10 users. In addition, the feedback overhead associated with collecting all users' average achievable data rates to a central node is excessive. In order to overcome these computational and feedback overheads, we consider the design of a heuristic online algorithm in the next section.

V. Online Algorithm - Dynamic User Association

The objective in solving **P-association** is to determine the association of each user, which is naturally related to handover and cell-site selection. Conventional algorithms for handover and cell-site selection are based on signal strength. Each user selects the best BS with the strongest mean channel quality, and it binds to the inner (or outer) band if the mean channel quality is higher (or lower) than a certain threshold. However, such a decision does not maximize the total utility since the satisfaction of a user depends on its actual throughput $\bar{R}_k$ rather than the signal strength. Moreover, $\bar{R}_k$ depends not only on the signal strength but on the population of users served by the BS. Even if the signal strength is high, the actual throughput may be low when many users are competing for the same resource. This observation motivates us to develop a new algorithm to bind users to BSs and bands. The following properties are essential in design of our online algorithm.

Proposition 5.1: (Intra-cell handover condition) Assume a user k is binding to BS n through band b and the numbers of inner/outer band users in BS n are large enough. Then changing the band currently being used to band $\bar{b}$ will improve the value of the network-wide objective function in Eq. (9) if

$$ET^b_{n,k} = \frac{G(Y^b_n)E[r^b_{n,k}]}{Y^b_n} < \frac{G(Y^{\bar{b}}_n + 1)E[r^{\bar{b}}_{n,k}]}{Y^{\bar{b}}_n + 1} = ET^{\bar{b}}_{n,k}, \tag{9}$$

where the overline represents changing the current associated band. Thus, $\bar{b}$ is equal to out-band if a user is currently associated with in-band, and in-band, otherwise; $ET^b_{n,k}$ is the *expected* long-term throughput according to Eq. (8) when user k is binding to BS n through band b.

Proposition 5.2: (Inter-cell handover condition) Assume a user k is binding to BS n and numbers of users in BS n and BS j are large enough. Then moving the user to another BS j will improve the value of the network-wide objective function in Eq. (9) if

$$ET^b_{n,k} = \frac{G(Y^b_n)E[r^b_{n,k}]}{Y^b_n} < \frac{G(Y^b_j + 1)E[r^b_{j,k}]}{Y^b_j + 1} = ET^b_{j,k}\,.^{5} \tag{10}$$

[5]The majority of inter-cell handovers will be between outer bands due to geographical adjacency.

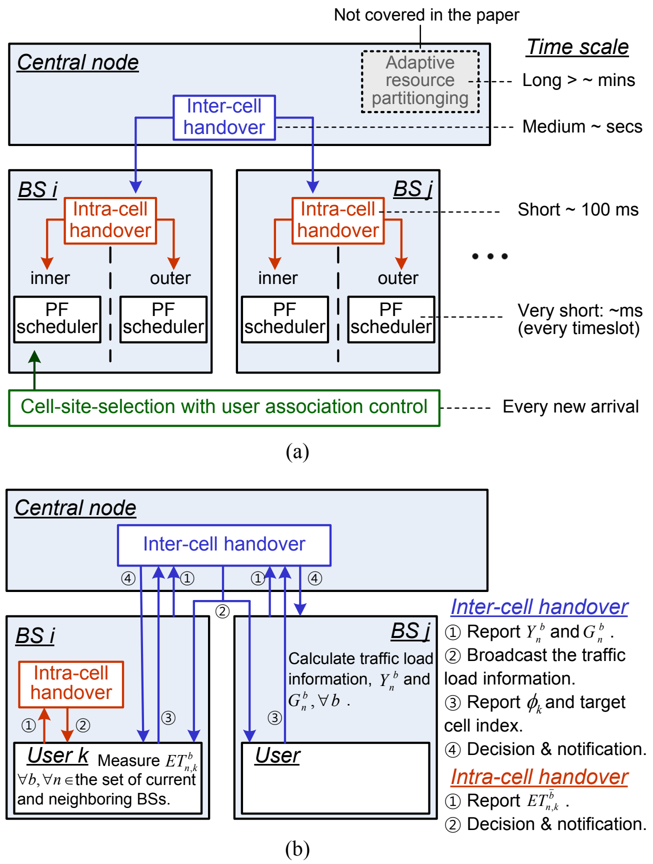


Fig. 2. Proposed framework in a hierarchy: (a) time scale of algorithms and (b) signaling for intra- and inter-cell handover.

Note that the left and right hand sides of Eqs. (9) and (10) are the *expected average throughput* of user k before and after the intra/inter-cell handover, respectively.

There is a reason why such simple conditions are obtained: when the number of users in each cell is large enough, the increment and decrement in the total utility (except user k) in cell i and j are almost the same and counterbalance each other. This makes the net increment of utility only depends on the handover of user k. Please refer to Appendix for detailed proofs. In a similar way, we can obtain a condition for user association control.

Proposition 5.3: (User association condition) Assume a new user k arrives to the system. Then admitting the user k to BS n on band b will improve the value of the network-wide objective function if

$$ET^b_{n,k} = \frac{G(Y^b_n + 1)E[r^b_{n,k}]}{Y^b_n + 1} > e, \tag{11}$$

where the constant e is base of the natural logarithm.

These propositions suggest the importance of the *expected throughput (load-aware metric)*. Based on this observation, we suggest a heuristic online algorithm for **P-association**. It is a simple mixture of intra/inter-cell handovers and cell-site selection with a user association control that use the *expected throughput* as a key metric in making association decisions instead of the signal strength. In our offline algorithm, whenever the arrival and departure of users occur or average channel gains are severely changed by users' mobility, we have to solve **P-association** again. Our online algorithm, however, keeps track of these dynamics and gradually changes users' associations following the steepest utility-increasing direction. For clarification, Fig. 2(a) and (b) depict the time scale of each component and signaling procedures for intra- and inter-cell handover, respectively. The following subsections describe the detailed procedure of each component.

A. Intra-cell Handover

Step 1. (Measurement & Report) Every user k binding to BS n periodically measures and reports the average achievable data rate in the band which is not in use. It is not necessary to report that of the band in use since the instantaneous achievable rate is sent to the BS each time slot to enable scheduling.
Step 2. (Decision) If many users change their serving bands at the same time, this may result in oscillations, thus BS n chooses only the user k_n^* that achieves the largest benefit by changing its band;[6]

$$k_n^* = \arg\max_k \phi_{n,k}, \tag{12}$$

where $\phi_{n,k} = ET_{n,k}^{\bar{b}}/ET_{n,k}^{b}$, $k = 1, \ldots, (Y_n^{in} + Y_n^{out})$ for all users in BS n and $\phi_{n,0} = \phi^{intra}$. We introduce a hysteresis $\phi^{intra} \geq 1$ to reduce possible ping-pong effects [21].
Step 3. (Notification) If $k_n^* = 0$, then either changing the band for any user cannot increase the value of the network-wide objective function or hysteresis precludes such a change. Thus, nothing occurs. Otherwise, the BS n notifies user k^* to change its intra-cell association.

The intra-cell handover is just the procedure to change the band currently being used, rather than a real handover, as such, it brings minimal system overhead. To accommodate channel variation due to mobility, the periodicity of intra-cell handovers should be performed on a short time scale (< 1 sec), and hysteresis, if used at all, should be small ($\phi^{intra} \approx 1$).

B. Inter-cell Handover

Step 1. (Measurement & Report) The central node periodically receives the following information from all BSs, and broadcasts this information so that every BS has the knowledge of its neighboring cells;

- Y_n^b and $G(Y_n^b)$: the number of users and multi-user diversity gain associated with BS n and band b.

Every BS n announces to all its associated users the above information for current and neighboring cells. Every user calculates expected throughputs from neighboring cells in addition to the current cell. Then, only users k, expecting higher throughput by changing BS n into j, report the highest ratio $\phi_k = ET_{j,k}/ET_{n,k}$ and the index of target cell j to the central node through the BS n.
Step 2. (Decision) To avoid the oscillation problems, the central node chooses the user k^* that achieves the largest benefit by changing its serving BS;

$$k^* = \arg\max_k \phi_k. \tag{13}$$

We introduce hysteresis with $\phi_0 = \phi^{inter} \geq 1$ to reduce possible ping-pong effects.
Step 3. (Notification) If $k^* = 0$, then either moving any user to another BS cannot increase the value of the network-wide objective function or hysteresis precludes such a switch. Otherwise, the central node notifies the user k^*, its original BS n and target BS j to handle the inter-cell handover.

In contrast to the intra-cell handover, the inter-cell handover is true a handover and brings additional system overheads. Thus the periodicity of inter-cell handovers should be large, i.e. time scales > 1 sec and also hysteresis should be implemented enough.

C. Cell-site Selection with User Association Control

Step 1. (Measurement & Report) A newly arriving user k measures average achievable data rates from several BSs. It reports this information to the central node through one of BSs offering the best signal strength.
Step 2. (Decision) The central node chooses the best BS n^* and band b^* that gives the highest expected throughput to the user k; [7]

$$(n^*, b^*) = \arg\max_{(n,b)} \ \phi_n^b, \tag{14}$$

where $\phi_n^b = ET_{n,k}^b, \ n \in N$ and $\phi_0^b = e$ according to the user association condition in Eq. (11).
Step 3. (Notification) If $i^* = 0$, the user k is rejected since admitting the user k will deteriorate the value of network-wide objective function. Otherwise, the central node notifies the user i to associate with the optimal BS i^* and band b^*.

D. Multi-user Diversity Gain Estimation

Under the assumptions made in Section III, the multi-user diversity gain can be written as $G(y) = \sum_{k=1}^{y} \frac{1}{k}$; it only depends on the number of users sharing the same resource. We make use of this property for mathematical tractability. However, in our simulations, we use the estimation of multi-user diversity gain to calculate the expected throughput more accurately. Below we describe a detailed procedure for estimating the multi-user diversity gain $G_n^b(t)$ at time t associated with BS n and band b.

1) A scheduler module, associated with BS n and band b, receives the instantaneous achievable rate $\{r_{n,k}^b(t)\}$ from all its associated users $k \in K_n^b$ at every time slot.
2) It takes an average of data rate for each user over fixed-length time window W:

$$\bar{r}_{n,k}^b(t) = \frac{1}{W} \sum_{\tau=t-W+1}^{t} r_{n,k}^b(\tau), \quad \forall k \in K_n^b. \tag{15}$$

[6]If more than one user achieves the same largest benefit, then a suitable random tie-breaking rule is used. The same is true for inter-cell handover in the next subsection.

[7]If more than one (i, b) achieves the same largest benefit, then a suitable random tie-breaking rule is used.

3) It also takes an average of the data rate for the same time window only if each user is selected/served by the scheduler:

$$\bar{r}_{n,k}^{b\,*}(t) = \frac{\sum_{\tau=t-W+1}^{t} r_{n,k}^b(\tau) I_{n,k}^b(\tau)}{\sum_{\tau=t-W+1}^{t} I_{n,k}^b(\tau)}, \quad \forall k \in K_n^b. \tag{16}$$

4) We obtain the multi-user diversity gain by taking an average of the ratio of Eq. (16) to Eq. (15) for all its associated users:

$$G_n^b(t) = \frac{1}{Y_n^b} \sum_{k \in K_n^b} \frac{\bar{r}_{n,k}^{b\,*}(t)}{\bar{r}_{n,k}^b(t)}. \tag{17}$$

E. Considerations for Implementation in Practice

There are several system parameters to be determined for the online algorithm, such as the periodicity of handover decision, the size of candidate set of BSs, the maximum number of association changes at a time and etc. In particular, the periodicity of handover decision balances between signaling overhead and the efficiency/accuracy of the algorithm (the shorter period, the faster response to change in the network; but the larger overhead). The other parameters play similar roles in the algorithm. We think the optimization of these parameters is more system dependent, so that we will leave this topic to system designers. In the paper, instead we have performed many simulations by varying these parameters and chosen them properly.

Using our online algorithm, additional inter-cell handover events may occur. If the system designers are concerned about this overhead, then they may use another passive approach using our load-aware metric only for new arrivals or during conventional handovers due to the mobility, rather than considering periodically the possibility of inter-cell handover for all users. It performs better than the conventional approach while maintaining the number of inter-cell handover events almost same, but does not guarantee the achievement of our long-term objective in Eq. (9).

VI. Performance Evaluation

A. Simulation Setup

The two-tier multi-cell network composed of 19(=N) hexagonal cells shown in Fig. 1 was considered, where the distance between BSs is 2km. In order to provide more realistic simulation results, we have investigated the performance of our scheme in a dynamic setting. Users arrive in any cell at tier t according to a Poisson process with rate λ_t at uniformly distributed locations and depart from the system after a holding time that is exponentially distributed with mean of $1/\mu_t = 60$sec. During their lifetime, we assume that users have infinitely backlogged queues and move based on the random waypoint model in which we fix the speed at 3km/h. The traffic load of tier t, $\rho_t = \frac{\lambda_t}{\mu_t}$, i.e., the average number of users in the system, can be changed by choosing a proper arrival rate λ_t.

All BSs have the same maximum transmission power p^{max}=20W and use up the power by allocating the power evenly to all the subbands being used. Therefore, the transmit powers for inner and outer bands are proportional to the size of each band so that they are given by $p^{in} = \frac{\alpha}{\alpha+(1-\alpha)/3} p^{max}$ and $p^{out} = p^{max} - p^{in}$, respectively. In modeling the propagation environment, a path loss $\Gamma(d[km]) = -130 - 35\log_{10}(d[km])$, log-normal shadowing with a standard deviation σ_s=8dB, and Jakes' Rayleigh fading [22] models were adopted. There is a shadowing correlation of 0.5 among paths from several BSs. We have assumed that each user sees interference from other cells up to two-tiers using a wrap-around technique [23]. The system bandwidth is 10MHz and the time slot is 5ms, this conforms with the IEEE 802.16e standard. The periods for intra- and inter-cell handovers are t^{intra}=0.1sec and t^{inter}=1sec, and ϕ^{intra}=1.01 and ϕ^{inter}=1.10 are used for hysteresis, respectively. For each given parameter set, we ran simulations over 720000 time slots (3600 sec).

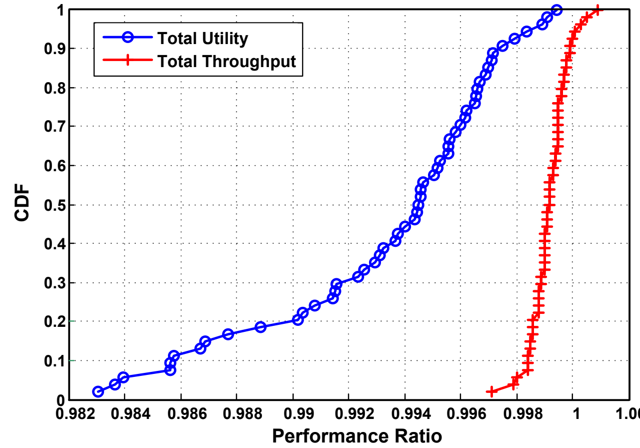


Fig. 3. Performance ratios between the optimal and the online algorithm: total utility and total throughput.

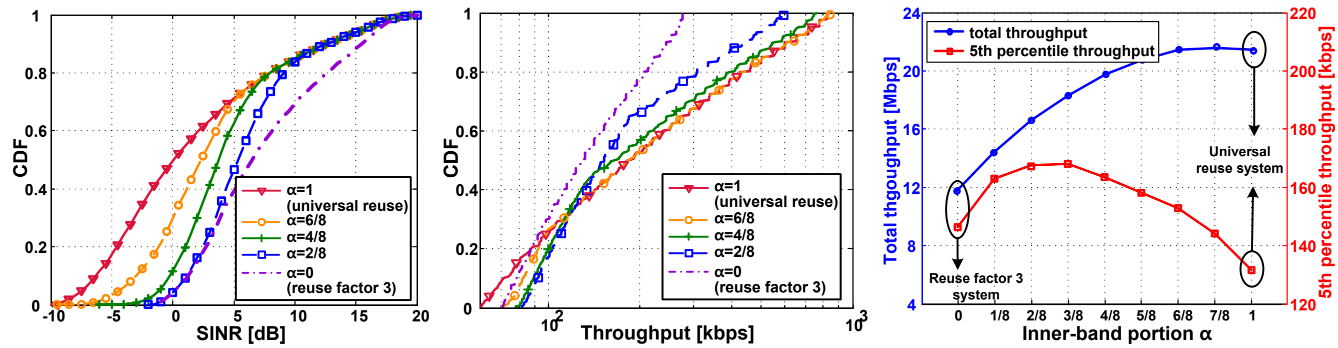
B. Comparison of Online Algorithm with Optimal

We randomly picked 100 different static (no user arrivals/departures or mobility) scenarios varying the resource partitioning and the number of users in each cell. For each scenario the optimal association was obtained by the offline algorithm, and we evaluated the heuristic-based online algorithm. Fig. 3 exhibits the CDF for the performance ratios, which are defined as the ratio between performance values obtained from the online algorithm and that from the optimal offline algorithm. As can be seen, the performance ratios of the total utility exceed 98% for all scenarios. Similarly, our online algorithm achieves a total throughput which is identical to that of optimal algorithm.[8] Thus we can conclude that *our online algorithm, which is efficient and easy to implement, is a good approximation of the offline algorithm.*

C. Interference Avoidance Gain

Next we move to dynamic scenarios, where users arrive and depart at/from uniformly distributed locations and also have

[8]Note occasionally the total throughput achieved by an online algorithm exceeds that of the offline algorithm, however, the total utility of the throughput is always lower because this is not an optimal point.



(a) (b) (c)

Fig. 4. Impact of the inner band portion α on system performances: (a) CDF of SINR, (b) CDF of throughput and (c) 5th percentile and total throughput.

mobility during their lifetime. We assume that every cell in the network has the same traffic load $\rho = 40$. We evaluate the system performance by varying the portion α of the spectrum devoted to the inner band. If $\alpha = 1$, then the system is operating under universal frequency reuse since each cell uses all resources. Also if $\alpha = 0$, then the system, operates under a reuse factor of 3 since each cell uses only $1/3$ of the total spectrum. So these two points can be obtained by traditional frequency reuse schemes. The middle points ($0 < \alpha < 1$) are newly achieved by employing PFR and our intra-cell handover to determine the optimal user binding to inner and outer bands.

Fig. 4(a) plots the CDF of the SINRs seen by users. The SINR curve moves consistently to the right as the inner band portion α decreases. Note that only the SINR of boundary users using the ICI mitigated outer band is improved about 4∼8dB using PFR. The lower α, the more ICI avoidance gain can be obtained. However, the amount of resource available in each cell is also reduced. To observe the real gains of PFR, we examine the throughput seen by each user in Fig. 4(b). As α decreases, there is a decreasing and increasing trend for the throughput of inner region users (high throughput users) and for the throughput of boundary users (low throughput users), respectively. It is worthy of note that the increment of throughput at the cell boundary becomes smaller as α decreases. In the extreme case ($\alpha = 0$), the throughput of edge users is even lower than $\alpha = 6/8$ case due to the lack of total resource available. Meanwhile, the increment of throughput in the inner region of cell becomes larger as α decreases. Therefore, it is very critical to choose a proper α, such that the performance of boundary users will be improved as much as possible while that of users associated with the inner region of cell is not excessively degraded.

Fig. 4(c) shows the total and 5th percentile average throughput together. The 5th percentile average throughput[9] is equal to the average of the lowest 5% throughput of users. This can be regarded as a representative performance metric of boundary users. As explained above, the decrease of α from $\alpha = 1$ increases the 5th percentile throughput until $\alpha = 3/8$, but decreases in the region $\alpha < 3/8$. At the same time, the total throughput is maximized at a certain point ($\alpha = 7/8$) and decreases after this point, i.e., concave function over α. From several results in Fig. 4, we choose $\alpha = 6/8$ as the resource partitioning scheme in the following simulation study.

D. Load Balancing Gain

Now let us consider a network with a heterogeneous user distribution. Cells in Tier 1 and 2 have $\rho = 40$, while the cell in Tier 0 has ρ_0 in [20,100]. The following four schemes are evaluated to determine the performance gains: *interference avoidance (IA)* and *load balancing (LB)*.

1) N/A: Universal frequency reuse ($\alpha = 1$); The inter-cell association is determined based on the best signal strength BS. This is a conventional scheme.
2) LB: Universal frequency reuse ($\alpha = 1$); Our proposed cell-site selection and inter-cell handover are used for balancing load. Note that this LB case, considering only load balancing scheme, can be regarded as Bu's work [6] because our [P2] without outer band reduces to their problem.
3) IA: PFR with $\alpha = 6/8$ for ICI coordination/avoidance; The inter-cell association is determined based on the best signal strength BS. Our proposed intra-cell handover is used for optimal band selection.
4) IA + LB: PFR with $\alpha = 6/8$ for ICI coordination/avoidance; Our proposed intra- and inter-cell handovers as well as cell-site selection are used.

Figs. 5(a)(b) show the 5th percentile throughput. At the same time, the average number of users in each tier is attached in the middle of graphs, i.e., (the average number of users without LB → the average number of users with LB). When Tier 0 is under-loaded ($\rho_0 < 40$), the LB moves edge users from Tier 1 to Tier 0. On the other hand, when Tier 0 is over-loaded ($\rho_0 > 40$), LB moves edge users from Tier 0 to Tier 1. For example, when $\rho_0 = 100$ in Fig. 5(a), it shifts about 20 users to the Tier 1. Thus, the average numbers of users are 78(≃100-22) and 43(≃ 40+22/6) for Tier 0 and Tier 1, respectively. Now let us examine the LB gain. The LB scheme shifts edge users in the hot-spot cell to under-loaded neighbor cells. This gives two advantages to the hot-spot cell: 1) the number of users competing for the same resource in the hot-spot cell is reduced; 2) less-poor channel users become

[9] Actually, *the cell boundary* is not clearly defined because a user, located closer to a BS n than another user associated with BS n, may be associated with another BS j due to shadowing. Moreover, if we adopt load balancing, then the boundary may be load dependent. Nevertheless, low throughput users in each cell are likely to locate in near its boundary. This is the reason why we regard the 5th percentile throughput as a representative for the performance of boundary users.

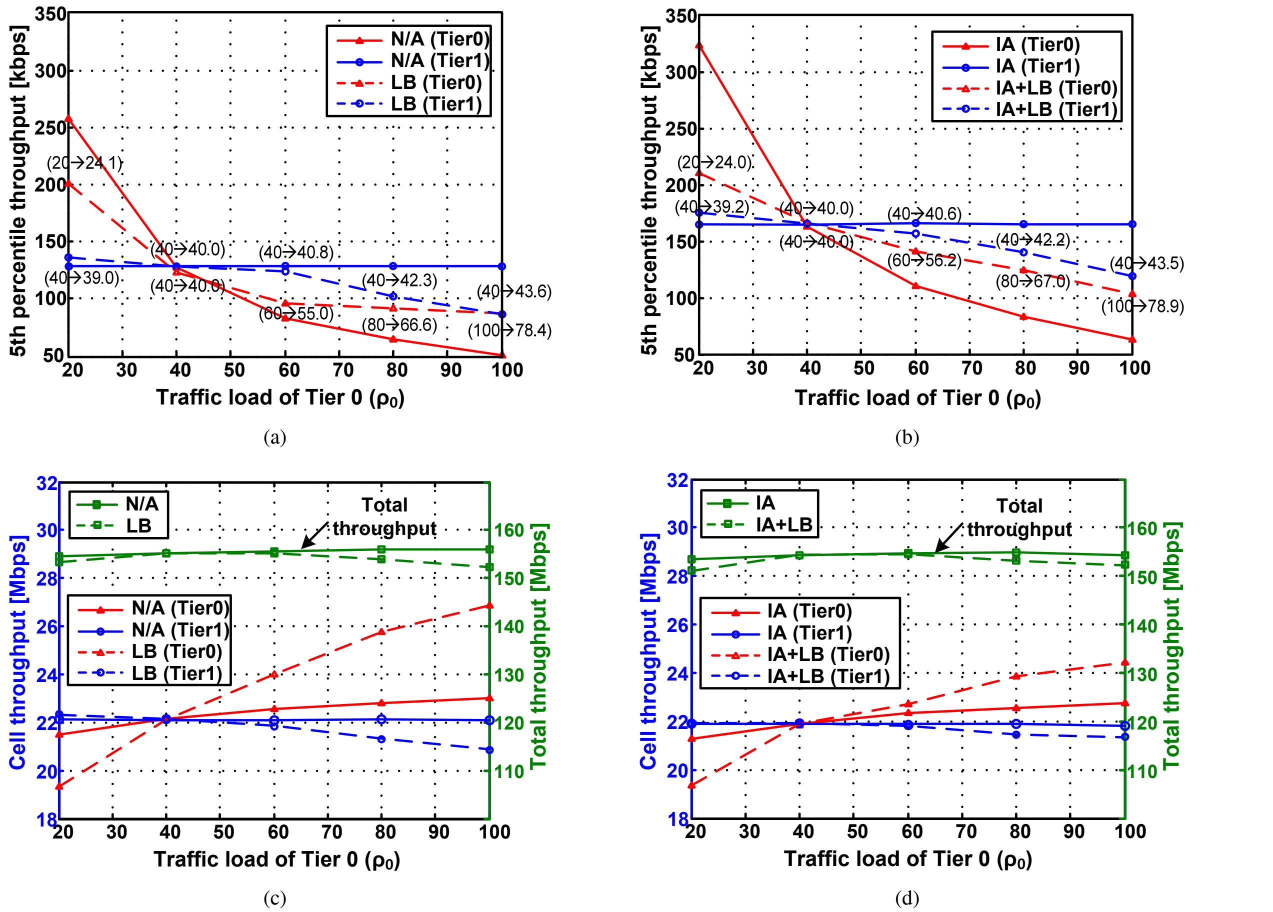


Fig. 5. System performances with different four schemes (N/A, LB, IA, IA+LB): (a),(b) 5th percentile throughput and (c),(d) cell and total throughput.

boundary users of the hot-spot cell because the coverage (not the physical signal range of the BS, but the area of users binding to the BS) of the hot-spot cell shrinks. Thus, as illustrated in Figs. 5(a)(b), the LB reduces the gap between 5th percentile throughputs and thus improves boundary (the worst) user performance in the system by a factor of $10{\sim}80\%$. The gain realized by LB increases with the heterogeneity of user distribution. Now let us consider IA scenario: By comparing Fig. 5(a) with (b), one can easily notice the difference of relative level. The 5th percentile throughput in the latter case, adopting load balancing, is 20% better than that in the former case. This gain depends on the value of α as we explained in the previous subsection.

Figs. 5(c)(d) show the cell throughput for each tier and total throughput which is defined as the sum of throughputs of all cells in Tier 0 and 1. When the LB is used for the over-loaded situation ($\rho_0 > 40$), the cell throughputs of Tier 0 and 1 are improved and degraded, respectively, compared to the cases where LB is not used (N/A, IA). This is because LB shifts edge users from Tier 0 to 1, which upgrades and degrades the average channel quality in Tier 0 and 1, respectively. This relation is reversed for the under-loaded situation ($\rho_0 < 40$). Meanwhile, the total throughput deteriorates a little as the heterogeneity increases. On the whole, however, total throughput remains almost same for all cases. In brief, *IA and LB gains improve the performance of cell boundary users while keeping the total throughput almost unchanged.*

E. QoS Violation Probability

To strengthen our claim that IA and LB help users at the cell edge, let us compute the QoS violation probability with a minimum throughput requirement, $r_k \geq m$. In Fig. 6, we plot the percentage of users whose average throughput is lower than a threshold for m=100kbps. As expected, the violation percentage decreases when IA or LB schemes are used. When both schemes are used, it decreases significantly. This means that the system can accommodate more satisfied users which boost the revenue of service provider.

VII. Conclusion

Next-generation broadband systems can provide the higher capacity, but users at cell edge still suffer from low throughput due to severe ICI and load imbalance. Therefore, to guarantee a QoS for boundary users and more balanced data rate among all users, PFR and load-balancing are considered in this paper. The main contributions of the paper can be summarized as follows.

We have formulated the proportional fair association problem **P-association** in multi-cell networks. We proposed an offline algorithm achieving the network-wide proportional fairness, and also proposed a practical online algorithm with less computational and feedback overheads. A remarkable feature of the proposed online algorithm is that it uses a notion of *expected throughput* as the decision making metric, instead of the signal strength in conventional systems. Extensive simulations demonstrate that our online algorithm brings two

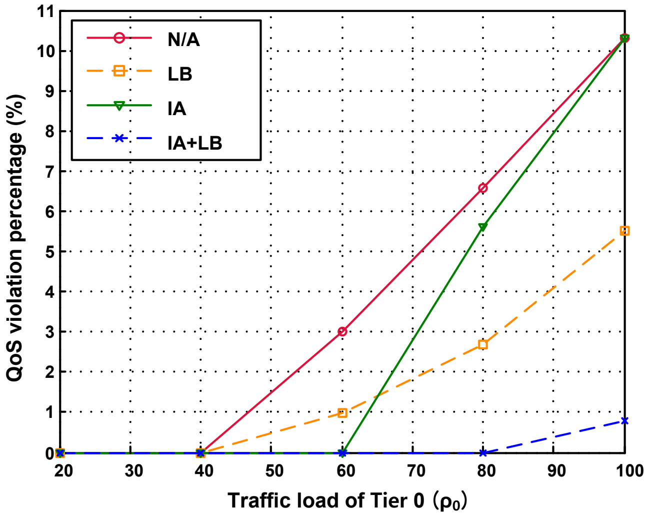


Fig. 6. QoS violation percentage with different four schemes (N/A, LB, IA, IA+LB).

types of performance gain: IA and LB gains, which improve the performance of users at the cell edge while not penalizing total system throughput.

APPENDIX

A. Proof of Proposition 5.1 and 5.2: Inter/Intra-handover conditions

Consider the net increment of utility for the inter-cell handover when a user k moves from BS n to j.

$$\begin{aligned}\triangle U = & \left[\log \frac{G(Y_j^b+1)E[r_{j,k}^b]}{Y_j^b+1} - \log \frac{G(Y_n^b)E[r_{n,k}^b]}{Y_n^b}\right] \\ & + \sum_{l \in k_n^b - \{k\}} \left[\log \frac{G(Y_n^b-1)E[r_{i,l}^b]}{Y_n^b-1} - \log \frac{G(Y_n^b)E[r_{i,l}^b]}{Y_n^b}\right] \\ & + \sum_{l \in K_j^b} \left[\log \frac{G(Y_j^b+1)E[r_{j,l}^b]}{Y_j^b+1} - \log \frac{G(Y_j^b)E[r_{j,l}^b]}{Y_j^b}\right],\end{aligned} \tag{18}$$

where $k_n^b = \{k \,|\, x_{n,k}^b = 1, k \in K\}$. The first term of Eq. (18) means the utility increment for the user k by changing the serving BS. And the second and the last term mean the aggregate utility increment of BS n by losing the user k and the decrement of BS j by adding the user k, respectively. We can rewrite Eq. (18) by deleting $E[r_{j,l}^b]$ and $E[r_{i,l}^b]$ because the handover does not affect them and all the users have log utility functions.

$$\begin{aligned}\triangle U = & \left[\log \frac{G(Y_j^b+1)E[r_{j,k}^b]}{Y_j^b+1} \frac{Y_n^b}{G(Y_n^b)E[r_{n,k}^b]}\right] \\ & + \log \left[\frac{G(Y_n^b-1)}{G(Y_n^b)} \frac{Y_n^b}{(Y_n^b-1)}\right]^{Y_n^b-1} \\ & + \log \left[\frac{G(Y_j^b+1)}{G(Y_j^b)} \frac{Y_j^b}{(Y_j^b+1)}\right]^{Y_j^b},\end{aligned} \tag{19}$$

Using the $\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e$ and the Euler's approximation to harmonic series $G(M) = \sum_{m=1}^{M} \frac{1}{m} \simeq \gamma + \log(M)$ where $\gamma = 0.5772\cdots$ is the Euler-Mascheroni constant, we can obtain the following equations for $Y_n^b \gg 1$ and $Y_j^b \gg 1$:

$$\begin{aligned}& \left(\frac{Y_n^b}{Y_n^b-1}\right)^{Y_n^b-1} \simeq e, \quad \left(\frac{Y_j^b}{Y_j^b+1}\right)^{Y_j^b} \simeq \frac{1}{e}, \\ & \left(\frac{G\left(Y_n^b-1\right)}{G\left(Y_n^b\right)}\right)^{Y_n^b-1} \simeq 1, \quad \left(\frac{G\left(Y_j^b+1\right)}{G\left(Y_j^b\right)}\right)^{Y_j^b} \simeq 1 .\end{aligned} \tag{20}$$

By putting Eq. (20) into Eq. (19), the aggregate utility increment of BS n decrement of BS j become +1 and -1 regardless of the numbers of users as long as they are large, and they counterbalance each other and do not contribute to the net change in the network-wide aggregate utility; only the first term survives in Eq. (19).

Note that the net change in the network-wide aggregate utility is affected only by the handover user k neglecting the impact of handover event on the other users. Thus, we have the inter-cell handover condition Eq. (10) (i.e., $\triangle U > 0$). Following the same procedure, we can derive the intra-cell handover condition Eq. (9) as well. ■

B. Proof of Proposition 5.3: User association condition

Let us consider the net increment of utility when a newly user k arrives to BS n.

$$\begin{aligned}\triangle U = & \log \frac{G(Y_n^b+1)E[r_{n,k}^b]}{Y_n^b+1} \\ & + \sum_{k \in k_n} \left[\log \frac{G(Y_n^b+1)E[r_{n,k}^b]}{Y_n^b+1} - \log \frac{G(Y_n^b)E[r_{n,k}^b]}{Y_n^b}\right].\end{aligned} \tag{21}$$

By putting Eq. (20) into Eq. (21), we have the user association condition Eq. (11) (i.e., $\triangle U > 0$). ■

ACKNOWLEDGMENTS

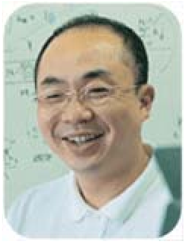
The authors would like to thank Jeong-woo Cho, Hyang-won Lee and Bang Chul Jung for their valuable discussions and help with our simulation. We would also like to thank the anonymous reviewers for their suggestions and comments.

REFERENCES

[1] IEEE Std. 802.16e-2005, "Part 16: air interface for fixed and mobile broadband wireless access systems," Feb. 2006.
[2] 3GPP TR 25.912 v7.2.0, "Feasibility study for evolved universal terrestrial radio access (UTRA) and universal terrestrial radio access network (UTRAN)," June 2007.
[3] S. V. Hanly, "An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity," *IEEE J. Select. Areas Commun.*, vol. 13, no. 7, pp. 1332-1340, Sept. 1995.
[4] S. Das, H. Viswanathan, and G. Rittenhouse, "Dynamic load balancing through coordinated scheduling in packet data systems," in *Proc. IEEE Infocom*, Mar. 2003, pp. 786-796.
[5] A. Sang, X. Wang, M. Madihian, and R. Gitlin, "Coordinated load balancing, handoff/cell-site selection, and scheduling in multi-cell packet data systems," in *Proc. ACM Mobicom*, Sept. 2004, pp. 302-314.
[6] T. Bu, L. Li, and R. Ramjee, "Generalized proportional fair scheduling in third generation wireless data networks," in *Proc. IEEE Infocom*, Apr. 2006, pp. 1-12.
[7] V. H. MacDonald, "The cellular concept," *Bell System Technical J.*, vol. 58, no. 1, pp. 15-41, Jan. 1979.
[8] G. Li and H. Liu, "Downlink radio resource allocation for multi-cell OFDMA system," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3451-3459, Dec. 2006.
[9] T. Bonald, S. Borst, and A. Proutiere, "Inter-cell scheduling in wireless data networks," in *Proc. European Wireless*, Apr. 2005.

[10] J. Cho, J. Mo, and S. Chong, “Joint network-wide opportunistic scheduling and power control in multi-cell networks," in *Proc. IEEE WoWMoM 2007*, June 2007, pp. 1-12.
[11] S. Liu and J. Virtamo, “Inter-cell coordination with inhomogeneous traffic distribution," in *Proc. NGI 2006*, Apr. 2006, pp. 64-71.
[12] F. Kelly, A. Maullo, and D. Tan, “Rate control in communication networks: shadow prices, proportional fairness and stability," *J. Operational Research Society*, vol. 49, pp. 237-252, July 1998.
[13] 3GPP TR 25.814 v7.1.0, “Physical layer aspect for evolved universal terrestrial radio access (UTRA)," Sept. 2006.
[14] Z. Zhang, Y. He, and E. K. P. Chong, “Opportunistic downlink scheduling for multiuser OFDM systems," in *Proc. IEEE WCNC*, Mar. 2005, pp. 1206-1212.
[15] J. Jang and K. B. Lee, “Transmit power adaptation for multiuser OFDM system," *IEEE J. Select. Areas Commun.*, vol. 21, no. 2, pp. 171-178, Feb. 2003.
[16] P. Viswanath, D. N. C. Tse, and V. Anantharam, “Asymptotically optimal water-filling in vector multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 1, pp. 241-267, Jan. 2001.
[17] A. L. Stolyar, “On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation," *Operations Research*, vol. 53, no. 1, pp. 12-25, Jan. 2005.
[18] H. J. Kushner and P. A. Whiting, “Convergence of proportional-fair sharing algorithms under general conditions," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1250-1259, July 2004.
[19] S. Borst, “User-level performance of channel-aware scheduling algorithms in wireless data networks," in *Proc. Infocom*, Mar. 2003, pp. 321-331.
[20] C. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, 1982.
[21] G. P. Pollini, “Trends in handover design," *IEEE Commun. Mag.*, vol. 34, no. 3, pp. 82-90, Mar. 1996.
[22] W. C. Jakes, *Microwave Mobile Communication*. New York: Wiley, 1974.
[23] J. Zander and S.-L. Kim, *Radio Resource Management in Wireless Networks*. Boston & London: Artech House, 2001.

Kyuho Son (S’03) received his B.S. and M.S. degrees in Electrical Engineering and Computer Science from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2002 and 2004, respectively. He is currently a Ph.D. candidate in the same university. His current research interests include interference management and load balancing in multi-cell networks, and spectrum sharing in cognitive radio networks. He is currently the Web Chair of the 7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt 2009).



Song Chong (S’93-M’95) received the B.S. and M.S. degrees in Control and Instrumentation Engineering from Seoul National University, Seoul, Korea, in 1988 and 1990, respectively, and the Ph.D. degree in Electrical and Computer Engineering from the University of Texas at Austin in 1995. Since March 2000, he has been with the School of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, where he is a Professor and the Director of Communications and Computing Group of the school. Prior to joining KAIST, he was with the Performance Analysis Department, AT&T Bell Laboratories, New Jersey, as a Member of Technical Staff. His current research interests include wireless networks, future Internet, human mobility and performance evaluation. He has published more than 70 papers in international journals and conferences and holds three U.S. patents in these areas.

He is an Editor of the JOURNAL OF COMMUNICATIONS AND NETWORKS and has served on the Technical Program Committee of a number of key international conferences including IEEE INFOCOM, ACM CoNEXT, ITC etc. He is the General Chair of WiOpt 2009. He is currently the Chair of Wireless Working Group of the Future Internet Forum and the Vice President of Information and Communication Society of Korea.

Gustavo de Veciana (S’88-M’94-SM’01-F’09) received his B.S., M.S, and Ph.D. in electrical engineering from the University of California at Berkeley in 1987, 1990, and 1993 respectively.

He is currently a Professor at the Department of Electrical and Computer Engineering at the University of Texas at Austin. He served as the Associate Director and then Director of the Wireless Networking and Communications Group (WNCG) 2004-2008. His research focuses on the design, analysis and control of telecommunication networks. Current interests include: measurement, modeling and performance evaluation; wireless and sensor networks; architectures and algorithms to design reliable computing and network systems.

Dr. de Veciana has served as editor for the IEEE/ACM TRANSACTIONS ON NETWORKING, and as co-chair of ACM CoNEXT 2008. He is the recipient of General Motors Foundation Centennial Fellowship in Electrical Engineering, an NSF Foundation CAREER Award 1996, co-recipient of the IEEE William McCalla Best ICCAD Paper Award 2000, and co-recipient of the Best Paper in ACM Transactions on Design Automation of Electronic Systems, 2002-2004.