J Netw Syst Manage
DOI 10.1007/s10922-009-9149-x

Modeling of Adaptive Load Balancing with Hybrid F/CDMA and Sectorization Schemes in Mobile Communication Networks

Kuo-Chung Chu · Frank Yeong-Sung Lin · Chun-Sheng Wang

Abstract In this paper, we investigate the load balancing problem in mobile communications by considering sectorization and a hybrid F/CDMA scheme (HFCS) jointly in the scenario of uneven traffic distributions. The problem is formulated as a combinatorial optimization model, subject to quality of service (QoS) requirements, and solved by the Lagrangean relaxation approach. In addition, Lagrangean multipliers are used to conduct sensitivity analysis. The model's objective is to minimize the weighted call blocking rate in terms of the distribution diversity. The model's performance is evaluated by the proposed HFCS, which is an adaptive scheme (AS). We compare the performance of AS with that of a non-adaptive (NA) scheme, which is a common power control method. Experiment results show that combining sectorization with the bandwidth segmentation scheme provides effective adaptive load balancing (ALB). The performance improvement achieved by the proposed adaptive scheme over the common power control scheme is as high as 68%. Moreover, under the scheme, the performance improves as the traffic load increases. Load balancing improves even further when AS is combined with the sectorization.

K.-C. Chu
Department of Information Management, National Taipei College of Nursing,
No. 365, Ming-te Road, Taipei 112, Taiwan
e-mail: kcchu@ntcn.edu.tw; kcchu8992@gmail.com

F. Y.-S. Lin
Department of Information Management, National Taiwan University,
No.1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan
e-mail: yslin@im.ntu.edu.tw

C.-S. Wang (✉)
Department of Information Management, Jinwen University of Science and Technology,
No. 99, An-chung Road, Taipei 231, Taiwan
e-mail: seanwang@just.edu.tw

Keywords Adaptive load balancing (ALB) · Hybrid F/CDMA scheme (HFCS) · Mobile communications · Performance modeling · Quality of service · Sensitivity analysis

1 Introduction

Code division multiple access (CDMA) is a promising approach for 3G wireless communication systems and beyond because all users share the same system bandwidth, i.e., reuse of unity, soft channel capacity, and so forth. Users transmitting on the same frequency band are identified by a user-specific code, but interference occurs due to imperfect code orthogonality. In a multi-cellular environment, when a particular CDMA cell becomes heavily loaded as the number of users increases, it will unavoidably affect all system users, i.e., users in the home cell as well as those in neighboring cells, especially in the scenario of a hot-spot and uneven traffic distribution. In CDMA systems, multi-access interference is a function of the number of users and is a limiting factor in ensuring quality of service (QoS), so there must be a tradeoff between the system capacity and the level of communication quality. Hence, the system's performance deteriorates as the number of active users increases.

Many studies have evaluated the capacity of CDMA systems. Some of them, e.g., [1, 2], assume uniform spatial traffic distribution because it best fits CDMA's characteristics if all signals share the whole spectral resource. Nevertheless, uniform traffic between cells (equal cell loads) is very uncommon in practice. For example, if the whole bandwidth is divided equally among the cells, the heaviest loaded cells have the same frequency resources as any other cell. Uneven traffic distribution will reduce the system's capacity. Though planned with sufficient capacity, uneven traffic distribution may occur in a cellular system thereby creating a "hot spot" that exceeds the pre-determined capacity, and introduces a large call blocking probability (CBP). For uneven traffic distribution, sectorization is an effective way to maximize a network's capacity [3, 4]. Meanwhile, soft handoff enforced by power control has been proposed as another possible solution to local traffic imbalances among cells [5–7].

If there is imbalance in the traffic distribution, the communications quality expressed by the signal to interference ratio (SIR) will differ between cells, and degrade efficient spectrum reuse in the whole system. For example, Fig. 1 shows two uneven traffic distributions in a 5 × 5 two-dimensional array with hexagonal cells. The shadowed cells indicate an uneven load; that is, their loads are either heavier or lighter than normal cells (without shadow). The user density is assumed to be uniform inside both uneven and normal cells.

One solution to the interference problem in a uniform traffic distribution environment is to use a power control mechanism, which attempts to achieve constant received mean power from each mobile station (MS) within a cell [8–10]. An adaptive load-shedding scheme combines the power control mechanism and the soft handoff function to force the users farthest away from the sector/cell (the two terms are used interchangeably hereafter) to enter forced soft handoff, and transfer

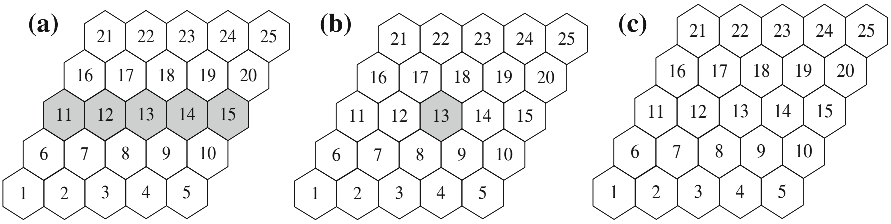


Fig. 1 Scenarios of traffic distributions. **a** Linear model. **b** Hot spot. **c** Uniform model

some of their loads to neighboring cells that are lightly loaded. In this way, heavily loaded cells dynamically down size their coverage area in order to service traffic, while adjacent cells that are less heavily loaded increase their coverage to accommodate the extra traffic, as shown in Fig. 2. However, with a hot-spot cell, powering up all users in the cell results in excessive interference for users in neighboring cells, so they cannot maintain sufficient SIR levels at their cell sites. Channel borrowing is a popular scheme to achieve load balancing, for example, [11] used fuzzy control to deal with the load imbalance; while [12] proposed a hybrid scheme of channel borrowing scheme and load transfer, it allows borrowing channels from light load cells, and ongoing calls can be transferred from heavy load cells into the overlapping cells they are light load. In addition, to improve global resource utilization and reduce regional congestion given heterogeneous arrivals, [13] requires load balancing among multiple cells. However, their works cannot be applied to CDMA systems because a lot of issues are different from other systems, e.g., channel definition, interferences, soft handoff.

To address this problem, [14] proposes a hybrid F/CDMA scheme (HFCS) that mitigates interference moderately. In [15, 16], HFCS is extended by using capacity analysis in a multi-band overlaid CDMA to achieve maximum bandwidth utilization. Specifically, the multi-band spectrum meets heterogeneous service requirements with sub-bands [16]. In this paper, we propose an adaptive load balancing (ALB) model that combines HFCS with a sectorization mechanism; the former effectively allocates the bandwidth to cells by applying a multi-band system, while the latter adjusts the size of each cell to accommodate users based on the

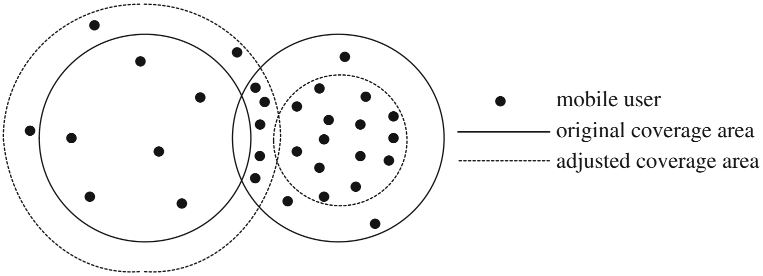


Fig. 2 Power control-based load balancing scheme

traffic load on the cell. The ALB problem is formulated as a mathematical optimization model. To address the load balancing problem, we try to maximize the system capacity by considering sectorization and HFCS jointly, and then allocate the appropriate sub-spectrum in a cell. We employ Lagrangean relaxation (LR) as our solution approach to efficiently solve complicated integer programming problem. The remainder of this paper is organized as follows. Sect. 2 reviews the background of CDMA. In Sect. 3, we present the ALB model as well as the solution approach. In Sect. 4, we describe the computational experiments and the sensitivity analysis. Then, in Sect. 5, we summarize our findings.

2 CDMA Background

2.1 Sectorization

CDMA capacity is limited by interference between users. An effective way to reduce the interference is to sectorize a cell by using directional antennas. This approach utilizes the spatial domain to introduce orthogonalization to the system. Sectorization employs directional antennas, so each user's signal is only received by one antenna. Since each antenna only receives a subset of users, the interference that each user experiences is less than that in a single antenna system. If the sector configuration is known, the interference can be estimated accurately by introducing an indicator function (Ω) of interference between sectors.

To pre-calculate the indicator function, the sector candidates to be configured in each base station (BS) must be defined. Denote K and B as the set of sector configurations and the set of BSs, respectively. In this paper, two probable configurations are given for a BS ($|K| = 2$), namely a one sector configuration (360° with an omni-directional antenna) and a three-sector configuration (120°per sector); and k ($k \in K$) is assigned as the identification (I.D.) of each configuration. Moving in an anti-clockwise direction, the sector I.D. i identifies a sector in the configuration k. Table 1 summarizes the sector candidates for each BS with a combination of k and i. Let S be the set of sectors. Each sector $s_{k,i}$ ($\forall s_{k,i} \in S$) is defined by both the sector configuration and the sector I.D. The coverage (in degrees) of each sector is listed in Table 2, where ϕ is the degree of the baseline. Generally speaking, ϕ can be assigned arbitrarily, but it is given $-30°$ in a cellular structure, as shown in Fig. 3. For simplicity, we omit subscripts k and i, then sector $s_{k,i}$ is replaced by s. Furthermore, sector s in BS j is denoted by sector_{js}, as shown in Fig. 4.

Table 1 Sector candidates for a BS

Candidate $s_{k,i}$	Configuration I.D. k	Sector I.D. i
1	1	1
2	2	1
3	2	2
4	2	3

Table 2 Coverage of candidate sectors

Candidate $S_{k,i}$	Sector I.D. i	Coverage of $S_{k,i}$
1	1	$(\phi, \phi + 360^\circ)$
2	1	$(\phi, \phi + 120^\circ)$
3	2	$(\phi + 120^\circ, \phi + 240^\circ)$
4	3	$(\phi + 240^\circ, \phi + 360^\circ)$

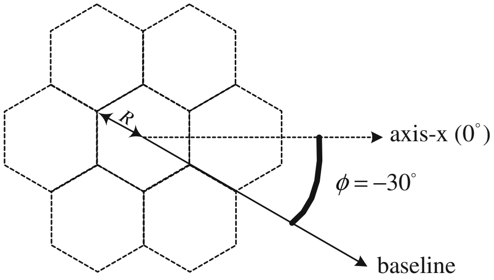


Fig. 3 The baseline deployed in a cellular structure

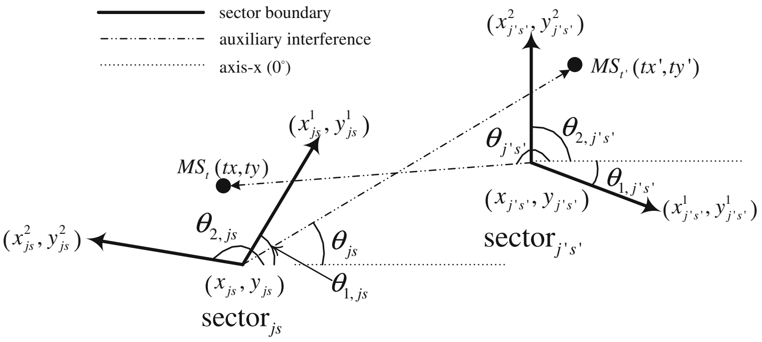


Fig. 4 Mutual interference between sectors

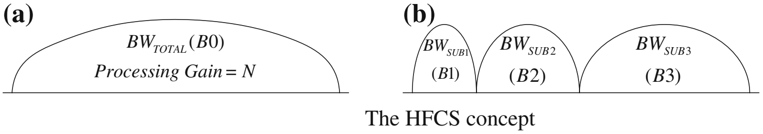


Fig. 5 The HFCS concept. **a** The spectrum of traditional CDMA. **b** The spectrum of the HFCS

2.2 Hybrid F/CDMA Scheme (HFCS)

As mentioned earlier the hybrid F/CDMA scheme (HFCS) [14] is also used to mitigate interference. Specifically, the available wideband spectrum, see Fig. 5a, is divided into a number of subspectra with smaller bandwidths, as shown in Fig. 5b.

Each subspectrum employs direct sequence spectrum spreading with reduced processing gain, which is transmitted in one and only one subspectrum. Assume BW_{WHOLE}, which is the whole frequency bandwidth on both the uplink (UL) and the downlink (DL), consists of a number of frequency units (FUs), each of which has a bandwidth BW_{FU}. Denote F_U as the set of FUs; then, $|F_U| = BW_{WHOLE}/BW_{FU}$. Furthermore, several frequency segments (FSs) in consecutive FUs can be combined. We denote F_S as the set of FSs.

For example, if the bandwidth allocated to the DL is decomposed into five FUs, $|F_U| = BW_{WHOLE}/BW_{FU} = 5$, as shown in Fig. 6, the FSs can be categorized, based on their length, into the five groups listed in Table 3. The notation (•) represents an FS and the notation • represents an FU. Based on $F_U = \{1, 2, 3, 4, 5\}$, $F_S = \{(1), (2), (3), (4), (5), (1,2), (2,3), (3,4), (4,5), (1,2,3), (2,3,4), (3,4,5), (1,2,3,4), (2,3,4,5), (1,2,3,4,5)\}$. The total number of FSs is $|F_S| = |F_U| \times (|F_U| + 1)/2 = 5 \times 6/2 = 15$. The FSs and their specific IDs are enumerated in Table 4. The capacity of HFCS is calculated as the sum of the capacities of the subspectra.

We investigate the load balancing problem by considering sectorization and the hybrid F/CDMA scheme jointly in a uneven environment. If there are four frequency segments (FS0, FS1, FS2, and FS3) to be assigned in a cell/sector, the FS assignment under different traffic distribution models would probably vary, as

Fig. 6 Example of bandwidth decomposition

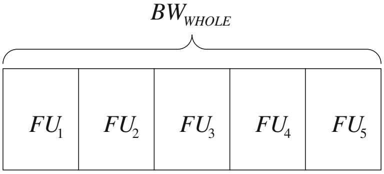


Table 3 Categorization of the FSs in the above example

Length of FS	Combination of FU	Number of FS
1	(1) (2) (3) (4) (5)	5
2	(1, 2) (2, 3) (3, 4) (4, 5)	4
3	(1, 2, 3) (2, 3, 4) (3, 4, 5)	3
4	(1, 2, 3, 4) (2, 3, 4, 5)	2
5	(1, 2, 3, 4, 5)	1

Table 4 The FSs in the above example

FS I.D.	Combination of FU	FS I.D.	Combination of FU	FS I.D.	Combination of FU
1	(1)	6	(1, 2)	11	(2, 3, 4)
2	(2)	7	(2, 3)	12	(3, 4, 5)
3	(3)	8	(3, 4)	13	(1, 2, 3, 4)
4	(4)	9	(4, 5)	14	(2, 3, 4, 5)
5	(5)	10	(1, 2, 3)	15	(1, 2, 3, 4, 5)

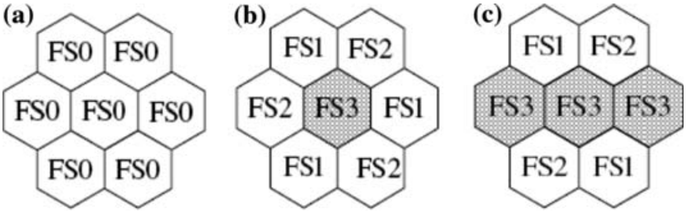


Fig. 7 The HFCS applied in different traffic distribution models. **a** Uniform model **b** hot spot **c** linear model

shown in Fig. 7, where the shadowed cells indicate heavy loads, because of the nature of uneven traffic distribution, the bandwidth that each cell requires to satisfy SIR would be different. The proposed scheme, called adaptive load balancing (ALB), optimally assigns FSs based on the traffic loads, and mitigates interference between cells/sectors by bandwidth segmentation.

2.3 Interference Model

Denote BW_{FU}^{UL} and BW_{FU}^{DL} as the FU assigned in the UL and DL directions, respectively. We assume that the both directions have the same number of FSs, i.e., the same $|F_S|$, and build SIR models with the HFCS. On the UL and DL, we denote y_{js} as the set of decision variables (DVs), which is m if sector$_{js}$ deploys FS m, $m \in F_S$. Then, the bandwidth allocated to the UL (W_{js}^{UL}) and the DL (W_{js}^{DL}) is calculated by (1) and (2), respectively, where $L(y_{js})$ is the length of FS m.

$$\mathrm{W}_{js}^{UL} = L(y_{js}) \cdot BW_{FU}^{UL}\, \forall j \in B, s \in S \tag{1}$$

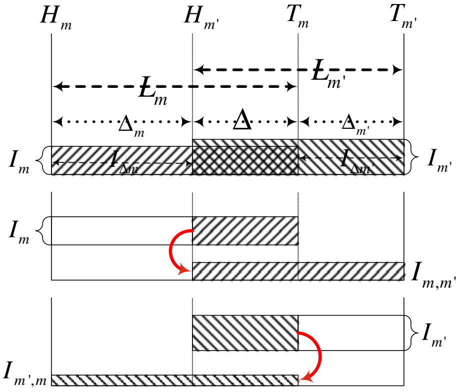
$$\mathrm{W}_{js}^{DL} = L(y_{js}) \cdot BW_{FU}^{DL}\, \forall j \in B, s \in S \tag{2}$$

The indicator function, $\Psi(y_{js}, y_{j's'})$, which is interference between sector$_{js}$ (using FS $y_{js} = m$) and sector$_{j's'}$ (using FS $y_{j's'} = m'$) can be pre-calculated. To better describe the calculation of indicator function $\Psi(y_{js}, y_{j's'})$, we denote $I_{m,m'} = \Psi(y_{js}, y_{j's'})$ and $I_{m',m} = \Psi(y_{j's'}, y_{js})$ as the interference from FS m to FS m' and from FS m' to FS m, respectively, and illustrate the mutual interference between FSs in Fig. 8.

Denote L_m as the length of FS m, where H_m and T_m are the first and last FUs of FS m, respectively. Then, $L_m = |T_m - H_m + 1|$, and $\Delta = |T_m - H_{m'} + 1|$ is the length of the overlap between FS m and FS m'. To calculate $I_{m,m'}$ and $I_{m',m}$, we only consider segment Δ, since the residual part I_{Δ_m} of I_m and the residual part $I_{\Delta_{m'}}$ of $I_{m'}$ will not interfere with each other. Thus, the interference strength in length Δ of FS m will be converted to the interference strength $I_{m,m'}$ with length $L_{m'}$ in (3), so we get $I_{m,m'}$ in (4). The calculation of $I_{m',m}$ is similar to that of $I_{m,m'}$, and is defined in (5) and (6).

$$\Delta = L_{m'} \cdot I_{m,m'} \tag{3}$$

Fig. 8 Mutual interference between FSs



$$I_{m,m'} = \frac{\Delta}{L_{m'}} \tag{4}$$

$$\Delta = L_m \cdot I_{m',m} \tag{5}$$

$$I_{m',m} = \frac{\Delta}{L_m} \tag{6}$$

To express the SIR models when HFCS is combined with sectorization, the interference scenario is shown in Fig. 9. Assuming the power of both the UL and DL are perfectly controlled, the received power in sector$_{js}$ from MS t ($t \in$ T, where T is a set of MSs) with constant value $P^{UL}_{c(t)}$ will be in the same traffic class-$c(t)$ in the UL, and the received power at MS t from sector$_{js}$with constant value $P^{DL}_{c(t)}$ will be in same traffic class-$c(t)$ in the DL. Denote $d^{UL}_{c(t)}$ ($d^{DL}_{c(t)}$) as the information rate in the UL (DL), and z_{jst} as a DV, which is 1 if MS t is admitted by sector$_{js}$, and 0 otherwise. The SIR values $SIR^{UL}_{js,c(t)}$ and $SIR^{DL}_{js,c(t)}$ in the UL and the DL are defined in (7) and (8), respectively, where ρ^{UL} (ρ^{DL}) is the UL (DL) orthogonality factor.

For $SIR^{UL}_{js,c(t)}$, $I^{UL}_{jst,intra}$ is intra-sector interference defined in (9), where $\alpha^{UL}_{c(t)}$ is an activity factor of traffic class-$c(t)$; while $I^{UL}_{jst,inter}$ is inter-sector interference defined in (10), where $\Omega^{UL}_{j's'js}$ indicates the interference between sectors and $\Psi(y_{js},\ y_{j's'})$

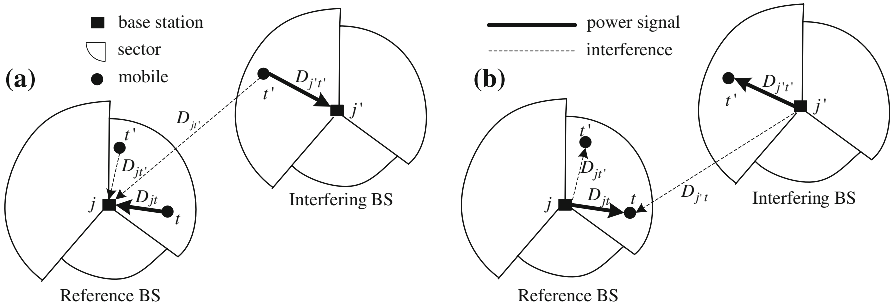


Fig. 9 The modified interference scenario: **a** UL interference; **b** DL interference

indicates the interference between FSs used in sector$_{js}$ and sector$_{j's'}$. If D_{jt} is the distance from MS t to sector$_{js}$, and r_{js} is the radius distance of sector$_{js}$, and τ is an attenuation factor; the $I^{UL}_{jst,inter}$ can be expressed by a reciprocal of distance from MS t to sector$_{js}$. For $SIR^{DL}_{js,c(t)}$, $I^{DL}_{jst,intra}$ is intra-sector interference defined in (11), where $\alpha^{DL}_{c(t)}$ is an activity factor of traffic class-$c(t)$; while $I^{DL}_{jst,inter}$ is inter-sector interference defined in (12), where $\Omega^{DL}_{j's'js}$ indicates the interference between sectors.

Equations (7) and (8) give a very large artificial constant value V in the numerator in order to satisfy the SIR constraints. This is because the SIR value must be larger than a pre-defined threshold, say the bit energy to noise ratio (BENR), if MS t is to be admitted by sector$_{js}$($z_{jst} = 1$); in other words, the constraint BENR $\leqq$ SIR must be satisfied. For example, in the UL in Eq. (7), if MS t is to be admitted by sector$_{js}$ ($z_{jst} = 1$), the SIR value $SIR^{UL}_{js,c(t)}$ is calculated by $\left(W^{UL}\Big/d^{UL}_{c(t)}\right)\cdot P^{UL}_{c(t)}\Big/\left((1-\rho^{UL})I^{UL}_{jst,intra}+I^{UL}_{jst,inter}\right)$ to determine whether the SIR constraint can be satisfied. In contrast, if MS t ($z_{jst} = 0$) is rejected, the SIR value is always larger than BENR (BENR $\ll$ SIR) because the value V is dominant $P^{UL}_{c(t)}$; thus, $SIR^{UL}_{js,c(t)}$ is calculated as a very large value. This implies that the constraint BENR $\leqq$ SIR can be ignored, as it is always satisfied.

$$\begin{aligned} SIR^{UL}_{js,c} &= \frac{W^{UL}_{js}}{d^{UL}_{c(t)}}\cdot\frac{P^{UL}_{c(t)}+(1-z_{jst})V}{(1-\rho^{UL})I^{UL}_{jst,intra}+I^{UL}_{jst,inter}}\\ &= \frac{L(y_{js})\cdot BW^{UL}_{FU}}{d^{UL}_{c(t)}}\cdot\frac{P^{UL}_{c(t)}+(1-z_{jst})V}{(1-\rho^{UL})I^{UL}_{jst,intra}+I^{UL}_{jst,inter}} \end{aligned} \tag{7}$$

$$\begin{aligned} SIR^{DL}_{js,c} &= \frac{W^{DL}_{js}}{d^{DL}_{c(t)}}\cdot\frac{P^{UL}_{c(t)}+\left(1-z_{jst}\right)V}{(1-\rho^{UL})I^{UL}_{jst,intra}+I^{UL}_{jst,inter}}\\ &= \frac{L\left(y_{js}\right)\cdot BW^{DL}_{FU}}{d^{DL}_{c(t)}}\cdot\frac{P^{DL}_{c(t)}+\left(1-z_{jst}\right)V}{(1-\rho^{DL})I^{DL}_{jst,intra}+I^{DL}_{jst,inter}} \end{aligned} \tag{8}$$

$$I^{UL}_{jst,intra} = \sum_{\substack{t'\in T\\ t'\neq t}} \alpha^{UL}_{c(t')}P^{UL}_{c(t')}z_{jst'} \tag{9}$$

$$I^{UL}_{jst,inter} = \sum_{\substack{j'\in B\\ j'\neq j}}\sum_{\substack{s'\in S\\ s'\neq s}}\sum_{\substack{t'\in T\\ t'\neq t}} \Omega^{UL}_{j's'js}\,\alpha^{UL}_{c(t')}P^{UL}_{c(t')}\left(\frac{D_{j't'}}{D_{jt'}}\right)^{\tau} z_{j's't'}\Psi\left(y_{j's'},y_{js}\right) \tag{10}$$

$$I^{DL}_{jst,intra} = \sum_{\substack{t'\in T\\ t'\neq t}} \alpha^{DL}_{c(t')}P^{DL}_{c(t')}\left(\frac{D_{jt'}}{D_{jt}}\right)^{\tau} z_{jst'} \tag{11}$$

$$I^{DL}_{jst,inter} = \sum_{\substack{j'\in B\\ j'\neq j}}\sum_{\substack{s'\in S\\ s'\neq s}}\sum_{\substack{t'\in T\\ t'\neq t}} \Omega^{DL}_{j's'js}\,\alpha^{DL}_{c(t')}P^{DL}_{c(t')}\left(\frac{D_{j't'}}{D_{j't}}\right)^{\tau} z_{j's't'}\Psi\left(y_{j's'},y_{js}\right) \tag{12}$$

3 Adaptive Load Balancing Model

3.1 Performance Measure

The adaptive load balancing (ALB) model considers multiple traffic classes, where each class is denoted as $c \in C$, and C is a set of traffic classes. The Kaufman model [17] is used as a performance measure to analyze the call blocking probability (CBP) of each traffic class. Assume M channels are shared by all traffic requirements. The traffic arrival rate is a stationary Poisson process with mean rate λ, and the channel requirement b is an arbitrary discrete random variable ($P\{b = b_c\} = q_c$, $c \in C$). A call request with channel requirement b_c has a holding time with mean $1/\mu_c$. Thus, traffic with channel requirement b_c is generated by a Poisson arrival process with mean rate $\lambda_c = \lambda q_c$, and the offered load of traffic class-c is $a_c = \lambda_c/\mu_c$. The CBP of traffic class-c is defined in (13), where the distribution of q(•), i.e., the number of channels required by the complete sharing policy, satisfies Equation (14) [17].

$$B^c(a,b) = \sum_{i=0}^{b_c-1} q(|M| - i)\, \forall c \in C \tag{13}$$

$$\sum_{c \in C} a_c b_c q(j - b_c) = jq(j)\, j = 0, 1, \ldots, M \tag{14}$$

where $q(x) = 0$ for $x < 0$, and $\sum_{j=0}^{M} q(j) = 1$. To better describe the concept of Kaufman's model, we use an example given in [17]. If we define $\hat{q}(j) = q(j)/q(0)$, then $\hat{q}(0) = 1$; and if we re-write (14) in the form of (15), it recursively generates $\hat{q}(j)$, $j = 1, \ldots, M$.) Then, (16) follows from (15) such that $q(j) = q(0)\hat{q}(j), j = 1, \ldots, M$.

$$\hat{q}(j) = \frac{1}{j}\left\{ \sum_{c \in C} a_c b_c \hat{q}(j - b_c) \right\} \tag{15}$$

$$q(0) = \left[\sum_{j=0}^{M} \hat{q}(j) \right]^{-1} \tag{16}$$

Suppose that $(a_1, a_2) = (1/2, 1/3)$, $(b_1, b_2) = (2, 3)$, $M = 5$. Then,

$$\hat{q}(0) = 1 \tag{17}$$

$$\hat{q}(j) = \frac{1}{j}\{\hat{q}(j-2) + \hat{q}(j-3)\}. \tag{18}$$

We then calculate that

$$\hat{q} = \left(1,\ 0,\ \frac{1}{2},\ \frac{1}{3},\ \frac{1}{8},\ \frac{1}{6}\right)$$

$$q(0) = \frac{24}{51},\ q = \frac{1}{51}(24,\ 0,\ 12,\ 8,\ 3,\ 4)$$

and

$$B^1 = \frac{7}{51}, B^2 = \frac{15}{51}$$

3.2 Model Formulation

Generally speaking, traffic load imbalances are unavoidable in a multi-cellular environment. Thus the goal of HFCS is to provide the best possible load balancing operation. Since the capacity of each cell/sector is calculated subject to SIR requirements, a lightly loaded cell/sector that will probably incur more interference than a heavily loaded cell. This results in an increased CBP in the lightly loaded cell. By allocating an appropriate FS to each sector, it is possible to mitigate interference between cells in a uneven traffic environment, as mentioned in Sect. 2.2. The notations and decision variables (DVs) used in the load balancing model are defined in Table 5.

The proposed model examines the load balance among all cells/sectors in terms of the call blocking rate. In each sector$_{js}$, we denote the vector of traffic intensities $g_{js}^{c_i}$ as $VG_{js} = (g_{js}^{c_1}, g_{js}^{c_2}, \ldots, g_{js}^{c_{|C|}})$ and as the vector of channels required for traffic class-c_i as $VM_{js} = (m_{js}^{c_1}, m_{js}^{c_2}, \ldots, m_{js}^{c_{|C|}})$. The weighted factor w_{js} is expressed by $w_{js} = \sum_{c \in C} g_{js}^c \Big/ \sum_{j \in B} \sum_{s \in S} \sum_{c \in C} g_{js}^c$. The objective function (IP) is to minimize the weighted call blocking rate (CBR):

$$Z_{IP} = \min \sum_{t \in T} K^{c(t)} \sum_{j \in B} \sum_{s \in S} w_{js} g_{js}^{c(t)} CB_{js}^{c(t)} \left(VG_{js}, VM_{js} \right), \tag{IP}$$

subject to

$$\left(\frac{E_b}{N_{TOTAL}} \right)_{c(t)}^{UL} \leq \frac{L\left(y_{js}\right) \cdot BW_{FU}^{UL}}{d_{c(t)}^{UL}} \cdot \frac{P_{c(t)}^{UL} + \left(1 - z_{jst}\right) V}{(1 - \rho^{UL}) \sum_{\substack{t' \in T \\ t' \neq t}} \alpha_{c(t')}^{UL} P_{c(t')}^{UL} z_{jst'} + \sum_{\substack{j' \in B \\ j' \neq j}} \sum_{\substack{s' \in S \\ s' \neq s}} \sum_{\substack{t' \in T \\ t' \neq t}} \Omega_{j's'js}^{UL} \alpha_{c(t')}^{UL} P_{c(t')}^{UL} \left(\frac{D_{j't'}}{D_{jt'}} \right)^{\tau} z_{j's't'} \Psi\left(y_{j's'}, y_{js}\right)}$$

$$\forall j \in B, s \in S, t \in T \tag{19}$$

$$\left(\frac{E_b}{N_{TOTAL}} \right)_{c(t)}^{DL} \leq \frac{L(y_{js}) \cdot BW_{FU}^{DL}}{d_{c(t)}^{DL}} \frac{P_{c(t)}^{UL} + (1 - z_{jst}) V}{\left((1 - \rho^{DL}) \sum_{t' \neq t^{t' \in T}} \alpha_{c(t')}^{DL} P_{c(t')}^{DL} \left(\frac{D_{jt'}}{D_{jt}} \right)^{\tau} z_{jst'} + \sum_{\substack{j' \in B \\ j' \neq j}} \sum_{\substack{s' \in S \\ s' \neq s}} \sum_{\substack{t' \in T \\ t' \neq t}} \mathit{\Omega}_{j's'js}^{DL} \alpha_{c(t')}^{DL} P_{c(t')}^{DL} \left(\frac{D_{j't'}}{D_{j't}} \right)^{\tau} z_{j's't'} \mathit{\Psi}\left(y_{j's'}, y_{js}\right) \right)}$$

$$\forall j \in B, s \in S, t \in T \tag{20}$$

Table 5 The notations used in the load balancing model

Notation	Description
$\Omega_{jsj's'}^{dir}$	Indicator function of interference in the direction *dir* from sector$_{js}$ to sector$_{j's'}$
Φ_{js}^{c}	Threshold of the service rate for class-*c* in sector$_{js}$
τ	Attenuation factor, set at 4 throughout the paper
$\alpha_{c(t)}^{dir}$	Activity factor (AF) of traffic class-*c*(*t*) in direction *dir*, in which *dir* means the UL only, and *c*(*t*) means voice call only
ρ^{dir}	Orthogonality factor in *dir*
BW_{FU}	Bandwidth of the FU
BW_{WHOLE}	Whole bandwidth
λ_v/λ_d	Mean arrival rate of voice/data class traffic, which is a function of MS *t*, *c*(*t*)
$1/\mu_v$, $1/\mu_d$	Mean call holding time of voice/data traffic
B	The set of BSs
CB_{js}^{c}	CBP of Kaufman's function for traffic class-*c* in sector$_{js}$, as defined in (13)
D_{jt}	Distance between BS *j* and MS *t*
$(E_b/N_{total})_{c(t)}^{dir}$	BENR value for class-*c*(*t*) in *dir*
F_{LB}	Load balancing factor
F_S	The set of FSs
F_U	The set of FUs
g_{js}^{c}	Aggregate traffic intensity of traffic class-*c* in sector$_{js}$
g_{js}	Aggregate traffic intensity in sector$_{js}$
K^c	Balancing coefficient for traffic class-*c*, where $\sum_{c\in C} K^c = 1$
m_{js}^{c}	The number of channels required for traffic class-*c*(*t*) in sector$_{js}$
m_{js}	The total number of channels allocated to sector$_{js}$
VG_{js}	The vector of traffic intensities, g_{js}^{c}
VM_{js}	The vector of channels required for traffic classes, m_{js}^{c}
w_{js}	The weight of the traffic load in sector$_{js}$
$L(y_{js})$	DV, which is a length of BW_{FU}, of FS indicated by y_{js}
$P_{c(t)}^{dir}$	The signal strength received at BS *j*/sector$_{js}$ with traffic class-*c*(*t*) in *dir*
R_{js}	Upper bound (UB) on the radius of power transmission in sector$_{js}$
r_{js}	DV of the transmission power radius in sector$_{js}$
T	The set of MSs
u_{jt}/u_{jst}	Indicator function, which is 1 if MS *t* is covered by BS *j*/sector$_{js}$, and 0 otherwise
V	A very large constant value
Y	The set of power transmission radii
y_{js}	DV, which is *i* if sector$_{js}$ deploys FS *i*
z_{jst}	DV, which is 1 if MS *t* is admitted by sector$_{js}$ and 0 otherwise

$$\sum\nolimits_{t\in T} z_{jst}\lambda_{c(t)} \Big/ \mu_{c(t)} = g_{js}^{c(t)} \quad \forall j \in B, s \in S \tag{21}$$

$$\sum\nolimits_{t\in T} z_{jst} m^{c(t)} = m_{js}^{c(t)} \quad \forall j \in B, s \in S \tag{22}$$

$$z_{jst} D_{jt} \le r_{js} u_{jst} \quad \forall j \in B, s \in S, t \in T \tag{23}$$

$$0 \le r_{js} \le R_{js} \quad \forall j \in B, s \in S, r_{js} \in Y \tag{24}$$

$$\sum_{j \in B} \sum_{s \in S} z_{jst} \le 1 \quad \forall t \in T \tag{25}$$

$$\Phi^c_{js} \le \frac{\sum_{t \in T} z_{jst}}{\sum_{t \in T} u_{jst}} \quad \forall j \in B, s \in S, c \in C \tag{26}$$

$$z_{jst} = 0 \text{ or } 1 \quad \forall j \in B, s \in S, t \in T \tag{27}$$

$$y_{js} \in F_S \quad \forall j, \in B, s \in S \tag{28}$$

The SIR models, defined in (7) and (8), must be greater than the pre-defined threshold E_b/N_{TOTAL}. The SIR constraints for the UL and the DL are expressed by (19) and (20), respectively. The traffic intensity of class-c in sector$_{js}$ is calculated by (21). Allocation of the number of channels is constrained by (22). MS t must be serviced in the coverage of sector$_{js}$ by (23). The power transmission radius is assigned in a range by Constraint (24). Constraint (25) requires that each MS can be homed to only one sector$_{js}$. A pre-defined service rate Φ^c_{js} for class-c is given in (26). Constraint (27) defines the integer properties of the DVs. Under HFCS, only one FS can be deployed on both the UL and the DL by DV y_{js} in Constraint (28).

3.3 Solution Approach

Lagrangean relaxation (LR) is a solution approach for solving mathematical optimization problems, and is used to decompose such problems to exploit their special structure. Even though LR is a standard solution technique that can solve a wide range of combinatorial optimization problems, the following nontrivial tasks must also be considered in order to solve those problems efficiently and effectively. (1) In the problem formulation stage, the first step is to find a suitable formulation that can be decomposed into subproblems and solved by LR. This may require many attempts to reformulate the problem by trial and error. (2) In the solution procedure stage, it is necessary to decide which constraints should be relaxed and how a number of critical parameters can be carefully determined, so that the solution optimality and a high convergence rate can be achieved. These remain open issues. (3) How to apply Lagrangean multipliers and develop an efficient algorithm to get primal feasible solutions is a challenging issue. We use Lagrangean multipliers for sensitivity analysis so that the corresponding constraints can be evaluated for decision support.

The problem (IP) is transformed into an LR problem (LR) by relaxing constraints (19), (20), and (23), and multiplying them by the corresponding Lagrangean multipliers vector $V = (v^1_{jst}, v^2_{jst}, v^3_{jst}) \ge 0$. Then, we add them to the primal objective function as follows:

$$
\begin{aligned}
Z_D\left(v^1_{jst}, v^2_{jst}, v^3_{jst}\right) = \min \sum_{t\in T} K^{c(t)} \sum_{j\in B}\sum_{s\in S} w_{js} g^{c(t)}_{js} CB^{c(t)}_{js}\left(VG_{js}, VM_{js}\right) + \sum_{j\in B}\sum_{s\in S}\sum_{t\in T} v^1_{jst} \Bigg[(E_b/N_{TOTAL})^{UL}_{c(t)} \\
d^{DL}_{c(t)} \Bigg((1-\rho^{DL}) \sum_{\substack{t'\in T \\ t'\neq t}} \alpha^{DL}_{c(t')} P^{DL}_{c(t')} \left(\frac{D_{jt'}}{D_{jt}}\right)^{\tau} z_{jst'} + \sum_{\substack{j'\in B \\ j'\neq j}} \sum_{\substack{s'\in S \\ s'\neq s}} \sum_{\substack{t'\in T \\ t'\neq t}} \Omega^{DL}_{j's'js} \alpha^{DL}_{c(t')} P^{DL}_{c(t')} \left(\frac{D_{j't'}}{D_{j't}}\right)^{\tau} z_{j's't'} \Psi\left(y_{j's'}, y_{js}\right) \Bigg) \\
-L\left(y_{js}\right) \cdot BW^{UL}_{FU} \cdot \left(P^{UL}_{c(t)} + \left(1 - z_{jst}\right)V\right) \Bigg] + \sum_{j\in B}\sum_{s\in S}\sum_{t\in T} v^2_{jst} \Bigg[(E_b/N_{TOTAL})^{DL}_{c(t)} \\
d^{DL}_{c(t)} \Bigg((1-\rho^{DL}) \sum_{\substack{t'\in T \\ t'\neq t}} {}^{DL}_{c(t')} P^{DL}_{c(t')} \left(\frac{D_{jt'}}{D_{jt}}\right)^{\tau} \alpha z_{jst'} + \sum_{\substack{j'\in B \\ j'\neq j}} \sum_{\substack{s'\in S \\ s'\neq s}} \sum_{\substack{t'\in T \\ t'\neq t}} \Omega^{DL}_{j's'js} \alpha^{DL}_{c(t')} P^{DL}_{c(t')} \left(\frac{D_{j't'}}{D_{j't}}\right)^{\tau} z_{j's't'} \Psi\left(y_{j's'}, y_{js}\right) \Bigg) \\
-L\left(y_{js}\right) \cdot BW^{DL}_{FU} \cdot \left(P^{DL}_{c(t)} + \left(1 - z_{jst}\right)V\right) \Bigg] + \sum_{j\in B}\sum_{s\in S}\sum_{t\in T} v^3_{jst}\left(z_{jst} D_{jt} - r_{js} u_{jst}\right),
\end{aligned}
\tag{LR}
$$

subject to (21), (22), (24–28). Then (LR) can be rewritten as (LR1)

$$
\begin{aligned}
Z_D\left(v^1_{jst}, v^2_{jst}, v^3_{jst}\right) = \min \sum_{t\in T} K^{c(t)} \sum_{j\in B}\sum_{s\in S} w_{js} g^{c(t)}_{js} CB^{c(t)}_{js}\left(VG_{js}, VM_{js}\right) \\
+ \sum_{j\in B}\sum_{s\in S}\sum_{t\in T} \Bigg\{ z_{jst} \Bigg[\sum_{\substack{t'\in T \\ t'\neq t}} \Bigg(v^1_{jst} (E_b/N_{total})^{UL}_{c(t)} d^{UL}_{c(t)} \left(1-\rho^{UL}\right) \alpha^{UL}_{c(t')} P^{UL}_{c(t')} \\
+ v^2_{jst} (E_b/N_{total})^{DL}_{c(t)} d^{DL}_{c(t)} \left(1-\rho^{DL}\right) \alpha^{DL}_{c(t')} P^{DL}_{c(t')} \left(\frac{D_{jt'}}{D_{jt}}\right)^{\tau} \\
+ \sum_{\substack{j'\in B \\ j'\neq j}} \sum_{\substack{s'\in S \\ s'\neq s}} \Psi\left(y_{j's'}, y_{js}\right) \Bigg(v^1_{jst} (E_b/N_{total})^{UL}_{c(t)} d^{UL}_{c(t)} \Omega^{UL}_{j's'js} \alpha^{UL}_{c(t')} P^{UL}_{c(t')} \left(\frac{D_{j't'}}{D_{jt'}}\right)^{\tau} \\
+ v^2_{jst} (E_b/N_{total})^{DL}_{c(t)} d^{DL}_{c(t)} \Omega^{DL}_{j's'js} \alpha^{DL}_{c(t')} P^{DL}_{c(t')} \left(\frac{D_{j't'}}{D_{j't}}\right)^{\tau} \Bigg) \Bigg) \\
+ L\left(y_{js}\right) \left(v^1_{jst} BW^{UL}_{FU} + v^2_{jst} BW^{DL}_{FU}\right) V + v^3_{jst} D_{jt} \Bigg] \\
- L\left(y_{js}\right) \left(v^1_{jst} BW^{UL}_{FU} \left(P^{UL}_{c(t)} + V\right) + v^2_{jst} BW^{DL}_{FU} (P^{DL}_{c(t)} + V)\right) \Bigg\} \\
- \sum_{j\in B}\sum_{s\in S} r_{js} \sum_{t\in T} v^3_{jst} u_{jst}
\end{aligned}
\tag{LR1}
$$

In addition, (LR 1) can be decomposed into two subproblems (SUB 1) and (SUB 2), which we discuss next. We then describe the algorithm used to solve each subproblem.

Subproblem (SUB 1) related to DV r_{js}

$$Z_{SUB\,1} = \min - \sum_{j \in B} \sum_{s \in S} r_{js} \sum_{t \in T} v^3_{jst} u_{jst}, \tag{SUB1}$$

subject to (24)

To find the minimum value of $Z_{SUB\,1}$, we simply assign R_{js} to each r_{js}.

Subproblem (SUB 2) related to DVs $z_{jst}, y_{js}, g^c_{js}, m^c_{js}$:

$$\begin{aligned}
Z_{SUB2} = \min & \sum_{t \in T} K^{c(t)} \sum_{j \in B} \sum_{s \in S} w_{js} g^{c(t)}_{js} CB^{c(t)}_{js} \big(VG_{js}, VM_{js}\big) \\
& + \sum_{j \in B} \sum_{s \in S} \sum_{t \in T} \Bigg\{ z_{jst} \Bigg[\sum_{\substack{t' \in T \\ t' \neq t}} \bigg(v^1_{jst} (E_b/N_{total})^{UL}_{c(t)} d^{UL}_{c(t)} (1 - \rho^{UL}) \alpha^{UL}_{c(t')} P^{UL}_{c(t')} \\
& + v^2_{jst} (E_b/N_{total})^{DL}_{c(t)} d^{DL}_{c(t)} \big(1 - \rho^{DL}\big) \alpha^{DL}_{c(t')} P^{DL}_{c(t')} \left(\frac{D_{jt'}}{D_{jt}} \right)^{\tau} \\
& + \sum_{\substack{j' \in B \\ j' \neq j}} \sum_{\substack{s' \in S \\ s' \neq s}} \Psi\big(y_{j's'}, y_{js}\big) \bigg(v^1_{jst} (E_b/N_{total})^{UL}_{c(t)} d^{UL}_{c(t)} \Omega^{UL}_{j's'js} \alpha^{UL}_{c(t')} P^{UL}_{c(t')} \left(\frac{D_{j't'}}{D_{jt'}} \right)^{\tau} \\
& + v^2_{jst} (E_b/N_{total})^{DL}_{c(t)} d^{DL}_{c(t)} \Omega^{DL}_{j's'js} \alpha^{DL}_{c(t')} P^{DL}_{c(t')} \left(\frac{D_{j't'}}{D_{j't}} \right)^{\tau} \bigg) \bigg) \\
& + L\big(y_{js}\big) \Big(v^1_{jst} BW^{UL}_{FU} + v^2_{jst} BW^{DL}_{FU} \Big) V + v^3_{jst} D_{jt} \Bigg] \\
& - L\big(y_{js}\big) \Big(v^1_{jst} BW^{UL}_{FU} \Big(P^{UL}_{c(t)} + V \Big) + v^2_{jst} BW^{DL}_{FU} \Big(P^{DL}_{c(t)} + V \Big) \Big) \Bigg\},
\end{aligned} \tag{SUB2}$$

subject to (21), (22), (25–28), and

$$LB_{js} \leq m^c_{js} \leq UB_{js} \cdot \forall j \in B, s \in S, c \in C \tag{29}$$

As well as the constraints shown in the problem (IP), we pre-calculate the LB_{js} and UB_{js} of m^c_{js} to improve efficiency when solving the subproblem, for which a constraint is given in (29). The LB_{js} is defined as the total number of channels required by MSs that are only covered by sector$_{js}$, while UB_{js} is the total number of channels required to serve all MSs in sector$_{js}$. To better express the subproblem (SUB 2), we denote $coefz_{jst}$ as the coefficient of z_{jst}, which is a function of DV y_{js}, and calculate it as follows. Each $coefz_{jst}$ searches all combinations of (y_{js}, $y_{j's'}$) with complexity |T| × |B| × |S| × |F_S| × |B| × |S| × |F_S|:

$$\begin{aligned}
coefz_{jst}\big(y_{js}, y_{j's'}\big) = & \sum_{\substack{t' \in T \\ t' \neq t}} \bigg(v^1_{jst} (E_b/N_{total})^{UL}_{c(t)} d^{UL}_{c(t)} (1 - \rho^{UL}) \alpha^{UL}_{c(t')} P^{UL}_{c(t')} \\
& + v^2_{jst} (E_b/N_{total})^{DL}_{c(t)} d^{DL}_{c(t)} \big(1 - \rho^{DL}\big) \alpha^{DL}_{c(t')} P^{DL}_{c(t')} \left(\frac{D_{jt'}}{D_{jt}} \right)^{\tau} \\
& + \sum_{\substack{j' \in B \\ j' \neq j}} \sum_{\substack{s' \in S \\ s' \neq s}} \Psi\big(y_{j's'}, y_{js}\big) \bigg(v^1_{jst} (E_b/N_{total})^{UL}_{c(t)} d^{UL}_{c(t)} \Omega^{UL}_{j's'js} \alpha^{UL}_{c(t')} P^{UL}_{c(t')} \left(\frac{D_{j't'}}{D_{jt'}} \right)^{\tau}
\end{aligned}$$

$$+ v^2_{jst}(E_b/N_{total})^{DL}_{c(t)} d^{DL}_{c(t)} \Omega^{DL}_{j's'js} \alpha^{DL}_{c(t')} P^{DL}_{c(t')} \left(\frac{D_{j't'}}{D_{j't}}\right)^{\tau}\Bigg)\Bigg)$$

$$+ L\left(y_{js}\right)\left(BW^{UL}_{FU}P^{UL}_{c(t)} + BW^{DL}_{FU}P^{DL}_{c(t)}\right) + V\left(v^1_{jst} + v^2_{jst}\right) + v^3_{jst}D_{jt}.$$

We can calculate the minimum value of $coefz_{jst}$ by choosing y_{js} properly. Then (SUB 2) can be concisely rewritten, as shown in (30), and further decomposed into $|T|$ subproblems, each of which is expressed by l_t in (31). To calculate the minimum value of l_t, we must search all of $|T| \times |B| \times |S| \times |F_S| \times |B| \times |S| \times |F_S|$ combinations. For each class-c, since g^c_{js} and m^c_{js} are linear functions of z_{jst}, we only have to find the minimal number of z_{jst} that satisfies constraint $\Phi^c_{js} \le \sum_{t\in T} z_{jsc(t)} \Big/ \sum_{t\in T} u_{jsc(t)}$ in (26). Thus, for each sector$_{js}$, sort all l_t in ascending order and select the number of z_{jst} in class-c that has a guaranteed ratio of Φ^c_{js}.

$$\begin{aligned} Z_{SUB2} = \min \sum_{t\in T} \Bigg\{ \sum_{j\in B}\sum_{s\in S} \left[K^{c(t)} w_{js} g^{c(t)}_{js} CB^{c(t)}_{js}\left(VG_{js}, VM_{js}\right) + z_{jst}coefZ_{jst}\left(y_{js}, y_{j's'}\right)\right] \\ -L\left(y_{js}\right)\left(v^1_{jst}BW^{UL}_{FU}\left(P^{UL}_{c(t)} + V\right) + v^2_{jst}BW^{DL}_{FU}\left(P^{DL}_{c(t)} + V\right)\right)\Bigg\} \end{aligned} \tag{30}$$

$$\begin{aligned} l_{js} = \sum_{j\in B}\sum_{s\in S} \left[K^{c(t)} w_{js} g^{c(t)}_{js} CB^{c(t)}_{js}\left(VG_{js}, VM_{js}\right) + z_{jst}coefz_{jst}\left(y_{js}, y_{j's'}\right)\right] \\ -L(y_{js})\left(v^1_{jst}BW^{UL}_{FU}(P^{UL}_{c(t)} + V) + v^2_{jst}BW^{DL}_{FU}(P^{DL}_{c(t)} + V)\right) \end{aligned} \tag{31}$$

3.4 Getting Primal Feasible Solutions

After solving (SUB 1) and (SUB 2), for any $\left(v^1_{jst}, v^2_{jst}, v^3_{jst}\right) \ge 0$, the objective value of $Z_D\left(v^1_{jst}, v^2_{jst}, v^3_{jst}\right) \ge 0$ is a lower bound (LB) of Z_{IP}. Based on (LR 1), the dual problem $Z_D = \max\left(v^1_{jst}, v^2_{jst}, v^3_{jst}\right)$ is constructed to calculate the tightest LB of (IP). To get primal feasible solutions, we apply the following primal feasible algorithm for the ALB problem.

[*Algorithm ALB*]

Step 1 Sort the FSs in the set F$_S$ in ascending order of length

Step 2 Calculate N$_S$ = $\lceil |B| \times |S| / |F_S| \rceil$, which is the average number of sector$_{js}$, all of which will be allocated the same FS

Step 3 Calculate all distances D$_{jt}$ between MS *t* and the nearest sector$_{js}$ in BS *j* in descending order

Step 4 Assign r_{js} with the maximum radius of power transmission, so that sector$_{js}$ admits MS *t* (assign $z_{jst} = 1$) that are covered in each sector$_{js}$

Step 5 Check the aggregate traffic in all sector$_{js}$, and sort the traffic loads in ascending order

Step 6 Allocate FSs to sector$_{js}$ from the lightest to the heaviest load. Assign the first FS to the first N$_S$ sector$_{js}$, i.e., assign $y_{js} = 1$ to first N$_S$ sector$_{js}$, and the next FS to the next N$_S$ sector$_{js}$, and so on

```
Step 7 Check QoS constraints in all $sector_{js}$
    Step 7.1 Sequentially select all $sector_{js}$ from the heaviest load to lightest load
    Step 7.2 Drop a MS $t$ it is served by $sector_{js}$, and its distance is the farthest from $sector_{js}$, i.e., assign $z_{jst} = 0$ until it satisfies the QoS constraints
    Step 7.3 Repeat Steps 7.1 to 7.2 until the QoS requirements in all $sector_{js}$ are satisfied
Step 8 Deal with all MS $t$ that were not admitted (dropped in Step 7.2), i.e., assign $z_{jst} = 1$, where $u_{jst} = 1$
    Step 8.1 Sequentially select the non-admitted MS $t$
    Step 8.2 Check the aggregate traffic load in all $sector_{js}$, and sort the value of the traffic load
    Step 8.3 Admit the selected MS $t$ to the $sector_{js}$ that covers the MS (i.e., $u_{jst} = 1$) and has a lightly loaded $sector_{js}$ (re-assign $z_{jst} = 1$) if it satisfies QoS in all $sector_{js}$
    Step 8.4 Repeat Steps 8.1 to 8.3 until all non-admitted MS $t$ are dealt with
Step 9 Reduce $r_{js}$ until all MS $t$ are admitted by $sector_{js}$ ($u_{jst} = 1$)
Step10 If $\Phi_{js}^{c} \leq \sum_{t\in T} z_{jst} \Big/ \sum_{t\in T} u_{jst}$, this is an infeasible solution, so ignore this iteration
```

4 Experiment Results

4.1 Parameters

Fig. 1 shows the structure of a 5×5 two-dimensional array with hexagonal cells, given $R_{js} = 5.0$ km. The parameters used in the experiment are those reported in the literature. The required BENR for voice traffic (v) and data traffic (d) are $(E_b/N_{total})_v^{UL} = (E_b/N_{total})_v^{DL} = 7$ dB and $(E_b/N_{total})_d^{UL} = (E_b/N_{total})_d^{DL} = 6$ dB [18], respectively; and the information rates for v and d are $d_v^{UL} = d_v^{DL} = 9.6$ kbps [18–20], $d_d^{UL} = d_d^{DL} = 64$ kbps [20], respectively. The activity factors (AFs) are $\alpha_v^{UL} = \alpha_v^{DL} = \alpha_d^{UL} = \alpha_d^{DL} = 0.5$ [21, 22], and an attenuation factor is given $\tau = 4$. The number of channels required is $m^v = 1$ and $m^d = 7$, defined by the ratio of 64 kbps/9.6 kbps when the unit channel ($m^v = 1$) is used for 9.6 kbps. The orthogonality factors are $\rho^{UL} = 0.5$, $\rho^{DL} = 0.5$ [18]. The power is perfectly controlled by $P_v^{UL} = 10$ dB [19, 23], $P_v^{DL} = 15$ dB, $P_d^{UL} = 15$ dB, and $P_d^{DL} = 20$ dB; and the service rate is $\Phi_{js}^{v} = \Phi_{js}^{d} = 0.1$. Call requests for v and d are generated by the Poisson arrival process with λ_v and λ_d, respectively; and the mean call holding times are given as $1/\mu_v = 180$(s) and $1/\mu_d = 600$(s), respectively [24]. The traffic intensity generated by heavily loaded cells in a uneven distribution is five times greater than that in a uniform distribution. The error gap is calculated to be less than 30% in all cases, and the time consumption in each case is a maximum of 370 (sec).

4.2 Performance Analysis

In this section, we consider three traffic distribution models: uniform (U), linear (L), and hot spot (H) models. The performance analysis on ALB is manipulated by a combination of HFCS and sectorization, denoted as an adaptive scheme (AS). For comparison, we implement a non-adaptive (NA) approach with a common power control scheme [9, 10]. We analyze the weighted CBR (Z_{IP}) as a function of voice (data) traffic intensity with constant data (voice) traffic in order to compare a combination of the load balancing scheme and traffic distribution. The ratio of K^v vs. K^d (0.5 vs. 0.5) is given. First, without sectorization ($|S| = 1$) and given $\lambda_d = 30$, Z_{IP} is a function of voice traffic, as shown in Fig. 10a, while the CBR(Z_{IP}) is a function of data traffic with $\lambda_v = 15$ in Fig. 10b. Generally speaking, the heavier the traffic load, the greater will be the system imbalance. In Fig. 10a, the average system load (call blocking rate) is proportional to the voice traffic. The CBR(Z_{IP}) in Fig. 10b is similar, but the load calculated by the NA scheme increases significantly when the data intensity is 48. To evaluate the performance of the proposed AS approach, we compare it with the NA scheme (AS vs. NA) in terms of the average system load. The AS approach improves on the performance of the NA approach by approximately 6% ~ 8% for light loads and 5% ~ 24% for heavy loads, as shown in Fig. 11b, respectively. The results indicate that AS improves on NA in both cases.

For sectorization ($|S| = 3$) and the other parameters are the same as those in Fig. 10. The average loads are in range 0.085 to 0.66 for the NA scheme, as shown in Fig. 12a; and for the proposed AS approach, the values of CBR(Z_{IP}) are between 0.027 and 0.43. In Fig. 12b, the values are larger than those in Fig. 12a. Under the NA scheme, we get from 0.12 to 1.6, compared to 0.04 to 1.07 for the proposed AS approach. The traffic intensity always affects the load balance, irrespective of whether the cell is sectorized. Comparing the two approaches, we find that the AS approach improves load balance from 35% to 68%, as shown in Fig. 13a; and outperforms the NA scheme by between 34% and 64%, as shown in Fig. 13b. From Fig. 13, we conclude that, irrespective of what traffic intensity is given, the load balance improvement of the AS approach over the NA scheme is almost the same.

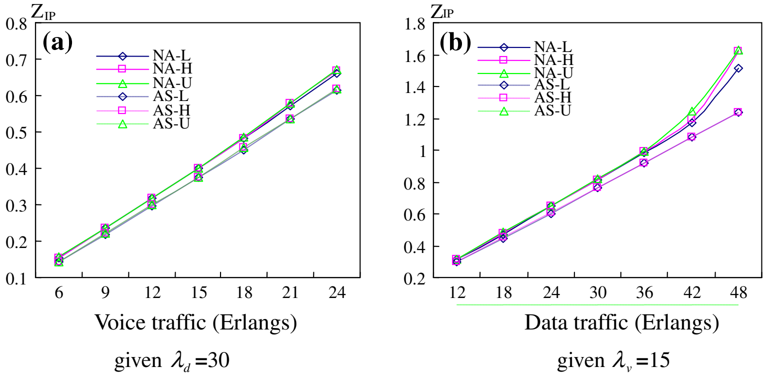


Fig. 10 Loading analysis without sectoriztion. **a** Given $\lambda_d = 30$. **b** Given $\lambda_v = 15$

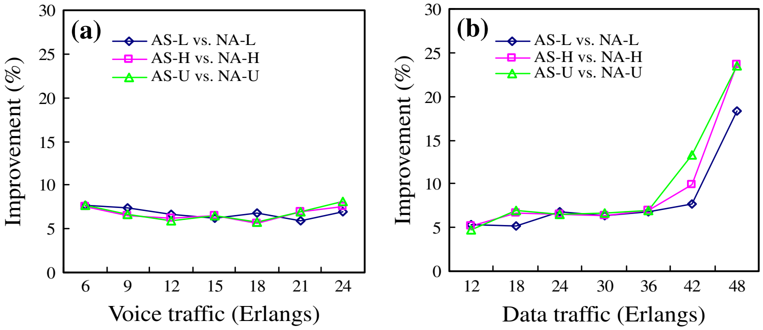


Fig. 11 Average improvement in the load balance without sectorization. **a** Given $\lambda_d = 30$. **b** Given $\lambda_v = 15$

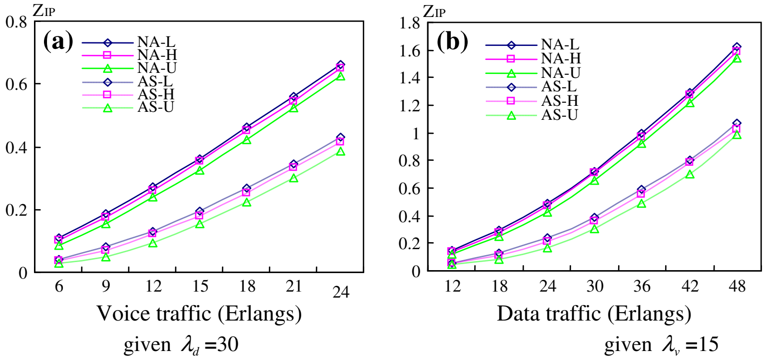


Fig. 12 Loading analysis with sectorization. **a** Given $\lambda_d = 30$. **b** Given $\lambda_v = 15$

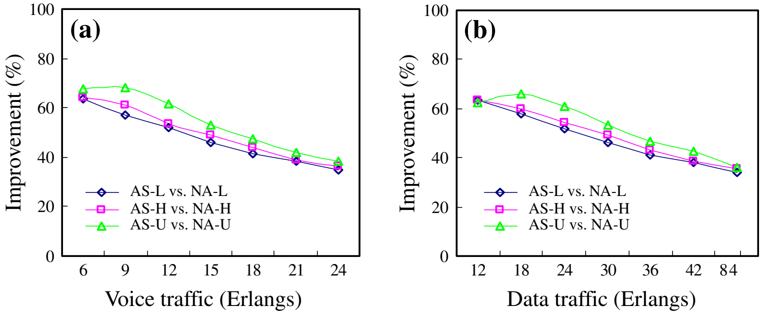


Fig. 13 Average improvement in the load balance with sectorization. **a** Given $\lambda_d = 30$. **b** Given $\lambda_v = 15$

This is because, by its nature, sectorization spreads the traffic load among sectors. In summary, under the AS approach, the more traffic that is loaded, the more the performance will improve. The load balance improves markedly by using the proposed AS approach.

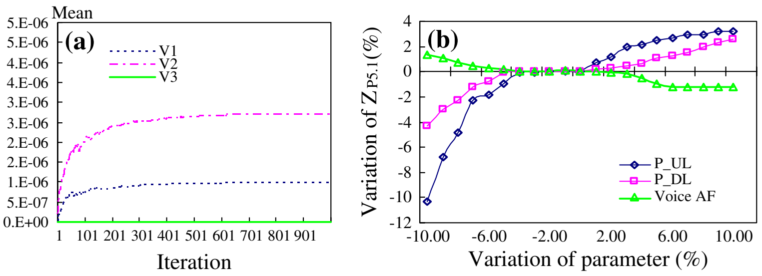


Fig. 14 Sensitivity analysis of the load balancing problem. **a** For Lagrangean multipliers. **b** For various load balancing parameters

4.3 Sensitivity Analysis

In the ALB model, we relax two constraints related to the SIR requirements. The analysis is based on 1,000 iterations, as shown in Fig. 14a, and the multiplier values V1 related to constraint (19) and V2 related to constraint (20) are converged to the corresponding constant values 9.8E-7 and 2.7E-6, respectively. This shows that the QoS of call connections is the most important issue in the load balancing problem. The multiplier value V3 related to constraint (23) is much smaller the other two, so we omit it from the discussion. In addition, the associated parameters, which represent the power of both links and the voice AF, are adjusted in the range of −10% to + 10% with a 1% step. The effects of the parameters on load balancing are presented in Fig. 14b. The strength of the UL power affects the load balance more significantly than the DL strength. It ranges from −10.30% to 2.45%, while the range is −4.30% to 2.55% in the DL. The load balance is also influenced by the voice AF, but the effect is moderate from +1.38 to −1.18%.

Most studies analyze the system capacity based on a uniform distribution; however, because of the demands for multimedia applications in next-generation systems, uneven traffic distribution will become increasingly common. A number of approaches have been proposed to deal with uneven traffic loads, e.g., changing the size of power controlled cells and adaptive sectorization. To some extent, the approaches mitigate the interference caused by uneven cell loads. Because the imperfect code orthogonality and the current frequency reuse factor of 1 is not always optimal, flexible bandwidth allocation by the HFCS is an alternative to balance the various distributions.

5 Conclusion

5.1 Contributions

To maximize the entire system capacity under ever-increasing uneven network traffic distributions, we propose a load balancing mechanism based on a combined sectorization and HFCS. By allocating resources efficiently for power and bandwidth assignments, we fulfill the requirements for load balancing. Furthermore, bandwidth arrangement adapts the scenarios of different loads. The load balancing is more significant in omni-cell than sectorized cell structure; that is achievement if the load balancing is more significant in non-uniformly sectorized structures than in uniformly sectorized structures. Irrespective of which distribution is considered, the proposed AS scheme reduces the blocking rate by up to 24% without sectorization, and by as much as 68% with sectorization. Overall, the proposed AS approach outperforms the NA scheme by as much as 68%.

5.2 Engineering Guidelines

For more practical application, further experiments need to be conducted, e.g., more generic cell planning rather than only using hexagonal cell structures, and a greater diversity of uneven distributions. The AS model, which combines sectorization and HFCS approach is a novel ALB scheme. In general, the average system load is proportional to the traffic distributed between sectors. The more traffic that is loaded, the greater the system imbalance will be. Nevertheless, if the network is properly sectorized, the imbalance can be reduced. The proposed AS approach outperforms the NA method under heavy load conditions. Significantly, under AS, the more traffic that is loaded, the more the scheme's performance will improve. The load balance improves significantly when the AS approach is used.

Acknowledgments This paper was awarded the Dragon Thesis Award, Gold Medal, by the Acer Foundation in 2005.

References

1. Wu, J.-S., Chung, J.-K., Sze, M.-T.: Analysis of uplink and downlink capacities for two-tier cellular system. IEE Proc. Commun. **144**(6), 405–411 (1997). doi:10.1049/ip-com:19971648
2. Liberti, J.C., Rappaport, T.S.: Analytical results for capacity improvements in CDMA. IEEE Trans. Vehicular Technol. **43**(3), 680–690 (1994). doi:10.1109/25.312781
3. Saraydar, C.U., Yener, A.: Adaptive cell sectorization for CDMA systems. IEEE J. Sel. Areas Comm. **19**(6), 1041–1051 (2001). doi:10.1109/49.926360
4. Lee, C.Y., Kang, H.G., Park, T.: A dynamic sectorization of microcells for balanced traffic in CDMA: genetic algorithms approach. IEEE Trans. Vehicular Technol. **51**(1), 63–72 (2002). doi:10.1109/25.992068
5. Hanly, S.V.: An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity. IEEE J. Sel. Areas Comm. **13**(7), 1332–1340 (1995). doi:10.1109/49.414650
6. Takeo, K., Sato, S.: Evaluation of a CDMA cell design algorithm considering non-uniformity of traffic and base station locations. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **E81-A**(7), 1367–1377 (1998)
7. Houtao, Z., Buot, T., Nagaike, R., Harmen, S.: Load balancing in WCDMA systems by adjusting pilot power. Proc. IEEE WPMC **3**, 936–940 (2002)
8. Tam, W.-M., Lau, F.C.M.: Analysis of power control and its imperfections in CDMA cellular systems. IEEE Trans. Vehicular Technol. **48**(5), 1706–1717 (1999). doi:10.1109/25.790552
9. Chen, X.H., Lee, K.L.: A novel adaptive traffic load shedding scheme for CDMA celular mobile systems. Proc. IEEE ICCS **2**, 566–570 (1994)
10. Chen, X.H.: Adaptive traffic-load shedding and its capacity gain in CDMA cellular mobile systems. IEE Proc. Commun. **142**(3), 186–192 (1995). doi:10.1049/ip-com:19951913

11. Wang, Y.-T., Sheu, J.-P.: A dynamic channel-borrowing approach with fuzzy logic control in distributed cellular networks. Simul. Model. Pract. Theory **12**(3–4), 287–303 (2004). doi:10.1016/j.simpat.2003.10.002
12. Ning, G., Zhu, G., Peng, L., Lu, X.: Research on hybrid dynamic load balancing algorithm in heterogeneous hierarchical wireless networks. J. Commun. **28**(1), 75–81+86 (2007)
13. Sang, A., Wang, X., Madihian, M., Gitlin, R.D.: Coordinated load balancing, handoff/cell-site selection, and scheduling in multi-cell packet data systems. Wirel. Netw. **14**(1), 103–120 (2008). doi:10.1007/s11276-006-8533-7
14. Eng, T., Milstein, L.B.: Comparison of hybrid FDMA/CDMA systems in frequency selective Rayleigh fading. IEEE J. Sel. Areas Comm. **12**(5), 938–951 (1994). doi:10.1109/49.298068
15. Zhuge, L., Li, V.O.K.: Reverse-link capacity of multiband overlaid DS-CDMA systems. Mobile Netw. Appl. **7**(2), 101–113 (2002). doi:10.1023/A:1013770704158
16. Zhuge, L., Li, V.O.K.: Overlaying CDMA systems with interference differentials. Mobile Netw. Appl. **8**(3), 269–278 (2003). doi:10.1023/A:1023345815475
17. Kaufman, J.S.: Blocking in a shared resource environment. IEEE Trans. Commun. **29**(10), 1474–1481 (1981). doi:10.1109/TCOM.1981.1094894
18. Chu, K.C., Lin, F.Y.S.: Survivability and performance optimization of mobile wireless communication networks in the event of base station failure. Comput. Electr. Eng. **32**(1–3), 50–64 (2006). doi:10.1016/j.compeleceng.2006.01.015
19. Chu, K.C., Hung, L.P., Lin, F.Y.S.: Adaptive channel reservation for call admission control to support prioritized soft handoff calls in a cellular CDMA system. Ann. Telecommun. **64**(11–12), 777–791 (2009). doi:10.1007/s12243-009-0126-x
20. Kim, D., Jeong, D.G.: Capacity unbalance between uplink and downlink in spectrally overlaid narrow-band and wide-band CDMA mobile systems. IEEE Trans. Vehicular Technol. **49**(4), 1086–1093 (2000). doi:10.1109/25.875215
21. Jeon, W.S., Jeong, D.G.: Call admission control for mobile multimedia communications with traffic asymmetry between uplink and downlink. IEEE Trans. Vehicular Technol. **50**(1), 59–66 (2001). doi:10.1109/25.917873
22. Jeon, W.S., Jeong, D.G.: Call admission control for CDMA mobile communications systems supporting multimedia services. IEEE Trans. Wirel. Comm. **1**(4), 649–659 (2002). doi:10.1109/TWC.2002.804189
23. Chu, K.C., Lin, F.Y.S., Wang, C.-S.: An admission control-based benefit optimization model for mobile communications: the effect of a decision time budget. J. Network Systems Manage., forthcoming. doi:10.1007/s10922-009-9152-2
24. Kim, K., Han, Y.: A call admission control scheme for multi-rate traffic based on total received power. IEICE Trans. Commun. **E84-B**(3), 457–463 (2001)

Author Biographies

Kuo-Chung Chu received his PhD degree in Information Management from the National Taiwan University in 2005. In 1992, he joined the Computer Center, Academia Sinica, Taiwan, where he was responsible for network systems management; followed by the Faculty of Information Management, Jinwen University of Science and Technology, Taipei, and as a department chair from 2006 to 2007. Since 2007, he has been with the Department of Information Management, National Taipei College of Nursing (NTCN), Taiwan. He is currently an Associate Professor in that department, and as a Director of Computer Center at NTCN. He received the Dragon Thesis Award, Gold Medal, by the Acer Foundation in 2005, and the Excellent Practical Dissertation Award, by the Industrial Development Bureau/MOEA and Chinese Society of Information Management of Taiwan in 2006, and the Excellent Paper Award, by the Operations Research Society of Taiwan in 2006. His papers have appeared in Annals of Telecommunications, Computers and Electrical Engineering, International Journal of Information and Management Sciences, Journal of Network and Systems Management, Telecommunication Systems, etc. His research interests include data mining, decision modeling, optimization approach, simulation, and their applications on healthcare and network management.

Frank Yeong-Sung Lin received his BS degree from the Electrical Engineering Department, National Taiwan University in 1983; and his PhD degree in Electrical Engineering from the University of Southern

California in 1991. After graduating from the USC, he joined Telcordia Technologies (formerly Bell Communications Research, abbreviated as Bellcore) in New Jersey, USA, where he was responsible for developing network planning and capacity management algorithms for a wide range of advanced networks. In 1994, Prof. Lin joined the Faculty of Electronic Engineering, National Taiwan University of Science and Technology. Since 1996, he has been with the Department of Information Management, National Taiwan University. His research interests include network optimization, network planning, performance evaluation, high-speed networks, wireless communications systems, distributed algorithms, and information security.

Chun-Sheng Wang received an MBA and a PhD degree in Information Management from National Taiwan University, Taiwan, ROC, in 1996 and 2007, respectively. He is currently an assistant professor of Information Management at the Jinwen University of Science and Technology, Taiwan, ROC. His papers have appeared in Journal of Systems and Software, Information Sciences, Expert Systems with Applications, etc. His current research interests include data mining, information systems, and network management.