Downlink Optimization and Performance of Relay-Assisted Cellular Networks in Multicell Environments

Wern-Ho Sheen, *Member, IEEE*, Shiang-Jiun Lin, and Chia-Chi Huang

Abstract—Deploying fixed relay stations (RSs) in traditional mobile cellular networks is widely recognized as a promising technology to improve cell coverage, user throughput, and system capacity, to save the transmit power of a mobile station (MS) in the uplink, and to provide a low-cost deployment for coverage extension. One crucial step toward developing such a relay-assisted cellular network is to fully evaluate its performance from both theoretical and practical viewpoints. In the literature, however, the system has only been evaluated for very limited system configurations: with a fixed number of RSs and locations and/or a fixed frequency-reuse pattern. This paper aims to investigate the downlink performance limits of a general relay-assisted network with optimized system parameters in a multicell environment. A genetic algorithm (GA) is proposed for the joint optimization of system parameters, including the number of RSs and their locations, frequency-reuse pattern, path selection, and resource allocation to maximize the system spectral efficiency (SE). Two types of quality of end-user experience (QoE), i.e., 1) fixed-bandwidth allocation (FBA) and 2) fixed-throughput allocation (FTA), are investigated along with two path-selection methods, i.e., 1) SE based and 2) signal-to-interference-plus-noise ratio (SINR) based. Numerical results show that significant improvement on system performance can be achieved with the optimized system parameters.

Index Terms—Downlink performance, genetic algorithm (GA)-based optimization, relay-assisted cellular system.

I. INTRODUCTION

DEPLOYING fixed relay stations (RSs) in traditional cellular networks, which are also known as relay-assisted cellular networks, has widely been recognized as a promising technology to improve cell coverage, user throughput, and system capacity, to provide more uniform data rates to users who are scattered over a cell, and to save the transmit power

W.-H. Sheen is with the Department of Information and Communication Engineering, Chaoyang University of Technology, Taichung 41349, Taiwan (e-mail: whsheen@cyut.nctu.edu.tw).

S.-J. Lin is with the Department of Communication Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan, and also with the Industrial Technology Research Institute, Hsinchu 31040, Taiwan (e-mail: hjlin@cm.nctu.edu.tw).

C.-C. Huang is with the Department of Communication Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan (e-mail: huangcc@cm.nctu.edu.tw).

of a mobile station (MS) in the uplink [1]–[16]. In addition, the relay-assisted cellular architecture allows a flexible and fast network rollout, which is particularly important at the initial stage of system deployment, where an extensive deployment of base stations (BSs) may not be economically viable [1], [3]–[5]. Very recently, the first commercial relay-assisted network has been standardized by the IEEE 80216j Task Group, aiming to improve the performance of IEEE 802.16e wireless mobile broadband networks [6].

One crucial step in developing a relay-assisted cellular network is to fully evaluate its performance from both theoretical and practical viewpoints. It has been a topic of extensive research in both academia and industry [2], [7]–[15]. In [2], a relay-assisted network was investigated in a multicell environment with six RSs in a cell, where a frequency-reuse scheme over the relaying links was proposed to improve the system spectral efficiency (SE). In [7], the uplink performance of a relay-assisted code-division multiple-access network was analyzed in a single-cell environment in which six RSs are deployed at fixed locations to improve system performance. Pole capacity and coverage were evaluated for different frequency allocations. In [8], the issues of RS positioning and spectrum partitioning were investigated with RSs located on the lines connecting BS and the six vertices of a hexagonal cell. Again, they considered the case of six RSs in a cell. The RSs' positions are optimized (along the line) to maximize the user throughput at cell boundary. In [9], the performance of a relay-assisted orthogonal frequency-division multiple-access network was evaluated for the specific setup of three RSs in a cell with and without intracell resource reuse; numerical results showed that the relay-assisted network significantly outperforms the conventional cellular network with respect to system capacity and coverage. Very recently, the IEEE 802.16j multihop relay network was evaluated in [10]–[15]; in particular, the uplink capacity was simulated with two RSs in a sector in [10], and the downlink capacity was simulated in [11] with one RS or three RSs in a sector for different modes of RS operations, including the transparent and nontransparent RSs with centralized or distributed scheduling. In [12], the deployment of RSs for coverage extension was investigated, and in [13] and [14], the issue of placement of BSs and RSs was considered with [13] focusing on the coverage extension application and [14] on the downlink performance with different number of RSs and cell radii. Finally, in [15], the optimal placement of RS and the time allocation were studied for the system employing one RS in a cell with uneven traffic distribution.

● BS   ○ Fixed RS

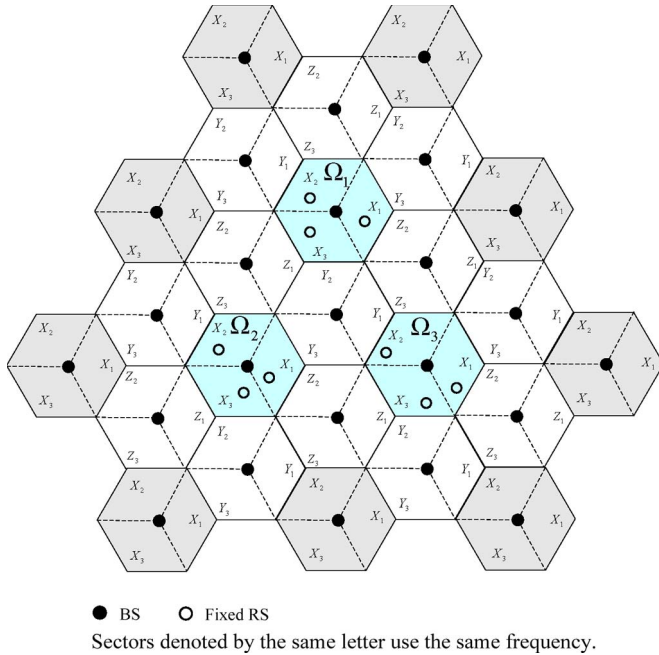Sectors denoted by the same letter use the same frequency.

Fig. 1.   Multicell architecture (3, 3, 3).

So far, as previously discussed, the performance evaluation and optimization of relay-assisted cellular networks have been limited to very specific system configurations: with a fixed number and location of RSs and/or using fixed reuse patterns. In this paper, our aim is to investigate the downlink optimization and performance limits of a general relay-assisted network in a multicell environment from an information-theoretic point of view. Two types of quality of end-user experience (QoE), i.e., 1) fixed-bandwidth allocation (FBA) and 2) fixed-throughput allocation (FTA), are investigated along with two path selection methods, i.e., SE based and signal-to-interference-plus-noise ratio (SINR) based. A genetic algorithm (GA) is proposed for the joint optimization of the system parameters, including RS's positions, reuse pattern, path selection, and resource allocation among different links to maximize the system SE. The theoretical performance serves as a benchmark. For more practical relay-assisted cellular networks, the effects of modulation and coding schemes and signaling overhead that enables the RSs' operation need to be taken into considerations.

The rest of this paper is organized as follows: Section II describes the system models, including multicell architecture, RS operation modes, propagation models, power setting criteria, path-selection methods, and reuse patterns. Section III proposes the joint optimization based on GA for different system configurations. Section IV gives the numerical results, and this paper is concluded in Section V.

II. SYSTEM MODELS

We investigate a multicell network that consists of BSs, fixed RSs, and MSs. Fig. 1 shows such an example, where RSs are deployed in sectors of a sectorized network to improve the system performance. The cell architecture is described by a three-tuple $(K_{\text{cluster}}, K_{\text{sector}}, K_{\text{band}})$, where K_{cluster}, K_{sector}, and K_{band} are the size of the cluster, the number of sectors in

a cell, and the number of frequency bands used in the sectors of a cell, respectively. For example, the cell architecture in Fig. 1 is described by (3, 3, 3). An area of K cochannel cells will be regarded as the design area in which RSs are deployed for performance enhancement. The BSs of these cells are designated as B_i, $i = 1, \ldots, K$.

Fig. 2(a) shows a more detailed layout of a cell, for example, ith cell, with no sectorization. The BS is located at the center of the cell, D is the cell radius, N_i RSs are deployed to improve the cell performance, $\vec{r}_{i,j}$ is the position vector of the jth RS (which is denoted by $R_{i,j}$), and $\vec{m}$ is the MS's position vector. The MSs are assumed to be uniformly distributed over the cell region Ω_i. The number of RSs can be different from cell to cell in its most general case.

The RSs operate in the in-band decode-and-forward mode, which means that the RSs decode the received signal first and then forward it to MS using the same radio technology as the BS in the same frequency band. An MS can communicate with the BS either through the direct path or the two-hop path (via an RS) that constitutes the BS–RS and RS–MS links. Two-hop relaying is considered as the most common scenario in practical applications because of the excessive delay incurred in more than two-hop relaying. Fig. 2(b) shows an example of the radio resource allocation, where orthogonal radio resources are allocated to the direct and two-hop paths, respectively. The radio resource for the two-hop path is further divided into two parts: one for the BS–RS link and the other for the RS–MS link. With this type of resource allocation, RSs are not required to transmit and receive at the same time (half-duplex relaying).

A. Propagation Models

The path and shadowing losses are separately treated from small-scale fading. For the path loss, the line-of-sight (LOS) and nonline-of-sight (NLOS) models for the suburban macro-cell environment in [16] are adopted, that is

$$L_{\text{path}}^{\text{LOS}}(d) = 23.8 \log_{10}(d) + 41.9 \text{ dB} \tag{1}$$

$$L_{\text{path}}^{\text{NLOS}}(d) = 40.2 \log_{10}(d) + 27.7 \text{ dB} \tag{2}$$

where d is the separation (in meters) between transmitter and receiver. The LOS model is used for the BS–RS link because RSs are often mounted over the rooftops, whereas the NLOS model is for the BS–MS and RS–MS links. In addition, the NLOS model is used for the calculation of multiple access interference (MAI).

For shadowing loss, a simplified model given in (3) is mainly adopted for verifying the effectiveness of the proposed method in a shadowed environment. Some other shadowing models can be used as well, e.g.,

$$L_{\text{shadow}}(\vec{m}) = \begin{cases} \delta \text{ dB}, & \text{if } \vec{m} \text{ is in a shadowed area} \\ 0 \text{ dB}, & \text{otherwise.} \end{cases} \tag{3}$$

In particular, the shadowed environment in Fig. 3, which consists of four obstacles, will be used in Section IV for numerical results. In this setup, an MS is said to be in a shadowed area if the LOS between BS (RS) and MS is blocked by an obstacle. No shadowing loss is imposed on the BS–RS link.
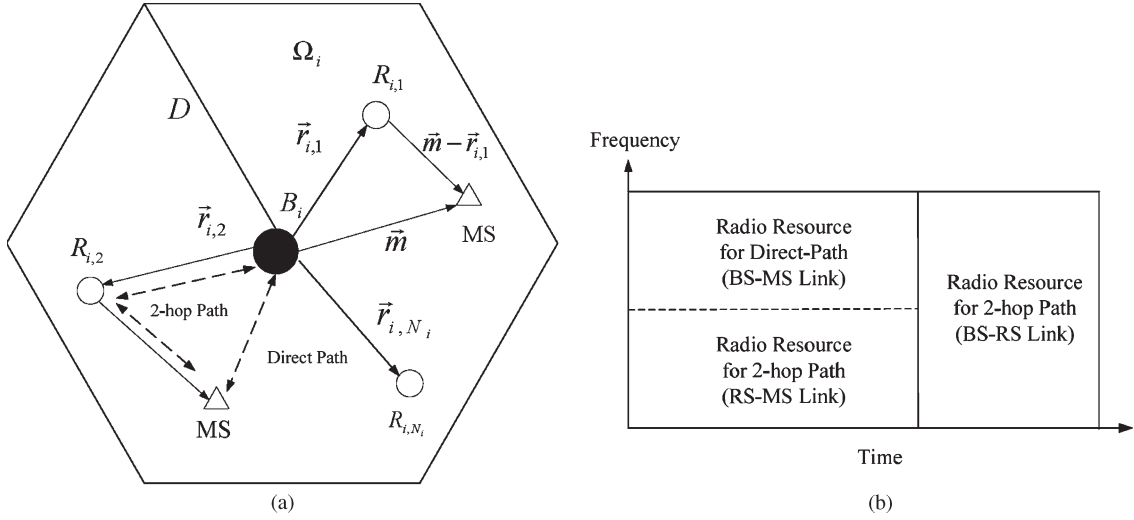
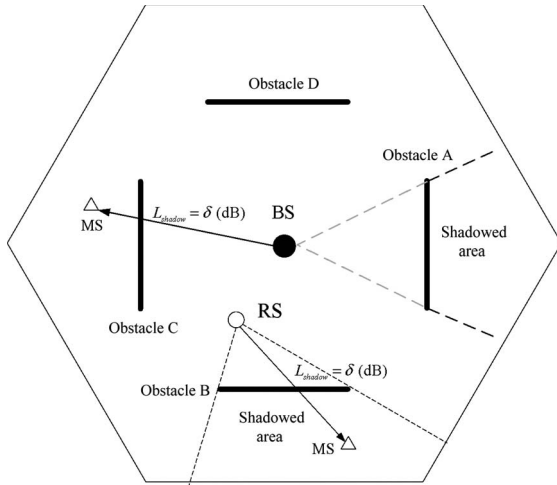Fig. 2. (a) Detailed layout of a cell. (b) Radio-resource allocation.

Fig. 3. Simplified shadowing model.

B. Antenna Configurations

RS and MS are equipped with one omnidirectional antenna, respectively, whereas both omnidirectional and sectored antenna configurations are investigated for BSs. For the sectored antenna, the antenna pattern proposed in [17] is adopted, which is

$$A(\theta) = -\min\left[12\left(\frac{\theta}{\theta_{3\,\mathrm{dB}}}\right)^2, A_m\right] \quad \mathrm{dB} \qquad (4)$$

where $-180° \leq \theta \leq 180°$ is the angle between the direction of interest and the bearing direction of the antenna, $\theta_{3\,\mathrm{dB}} = 70°$ is the 3-dB beamwidth, and $A_m = 20$ is the maximum attenuation.

C. Power Setting of BS and RS

A prespecified SE for users at the cell boundary is used to set up the transmitted power of a BS and its subordinate RSs. Let

$S = f(\mathrm{SINR})$ (in bits per second per hertz) denote the link SE as a function of average SINR, which is defined by

$$\mathrm{SINR} = \frac{p_t \cdot g_t \cdot L_{\mathrm{all}}^{-1}(d) \cdot g_r \cdot \mathrm{E}\left[|h|^2\right]}{N_0 + I_0} \qquad (5)$$

where p_t is the transmit power spectral density (PSD), g_t and g_r are the transmit and receive antenna gains, respectively, $L_{\mathrm{all}}(d) \doteq L_{\mathrm{path}}(d) \cdot L_{\mathrm{shadow}}$ is the composite effect of the path and shadowing losses in linear scale, h is a complex channel gain due to small-scale fading, $\mathrm{E}[\cdot]$ denotes the expectation operation, and N_0 and I_0 are the PSDs of additive white Gaussian noise (AWGN) and MAI, respectively.

For AWGN channels

$$f(\mathrm{SINR}) = \log_2(1 + \mathrm{SINR}) \qquad (6)$$

where h is a constant in (5). For fading channels, on the other hand, $f(\mathrm{SINR})$ may not assume a close form as in (6) and needs to numerically be evaluated, depending on the considered fading characteristics and whether the channel-state information is available at the transmitter [18]. In any case, for a specific SE S, the required SINR can be found by $\mathrm{SINR} = f^{-1}(S)$.

Letting S_{edge} be the targeted SE for an MS at the cell boundary, the transmit PSD of the BS is then set as

$$p_{B_i} = \frac{\mathrm{SINR}_{\mathrm{edge}} \cdot (N_0 + I_0)}{g_t \cdot L_{\mathrm{all}}^{-1}(D) \cdot g_r \cdot \mathrm{E}\left[\left|h_{B_i \to M}\right|^2\right]}$$

in watts per hertz (7)

where $\mathrm{SINR}_{\mathrm{edge}} = f^{-1}(S_{\mathrm{edge}})$. Likewise, the transmit PSD of $R_{i,j}$ is set as

$$p_{R_{i,j}} = \frac{\mathrm{SINR}_{\mathrm{edge}} \cdot (N_0 + I_0)}{g_t \cdot L_{\mathrm{all}}^{-1}\left(D - \|\vec{r}_{i,j}\|\right) \cdot g_r \cdot \mathrm{E}\left[\left|h_{R_{i,j} \to M}\right|^2\right]}$$

in watts per hertz (8)

where $\|\vec{r}_{i,j}\|$ denotes the Euclidean distance of the position vector $\vec{r}_{i,j}$. Note that with the power settings in (7) and (8), the coverage of RSs is entirely within that of the BS. In this

setup, RSs are mainly used to improve system capacity, user throughput, and removal of coverage holes due to shadowing rather than for coverage extension, where RSs are deployed at the cell boundary to cover areas that are originally not covered by the BS.

D. Path Selection

Path selection is a procedure to determine whether a direct path or a two-hop path is to be chosen by a BS to communicate with an MS. Two types of path selection will be investigated: SINR based and SE based. For the SINR-based path selection, the direct path is selected if $\text{SINR}_{B_i \to M}(\vec{m}) \geq \max_{R_{i,j}}\{\text{SINR}_{R_{i,j} \to M}(\vec{m})\}$; otherwise, the two-hop path through $R_{i,k}$ is selected, where $\text{SINR}_{B_i \to M}(\vec{m})$ and $\text{SINR}_{R_{i,j} \to M}(\vec{m})$ are the received SINRs over the B_i-MS and $R_{i,j}$-MS links, respectively, and $R_{i,k} = \arg\{\max_{R_{i,j}}\{\text{SINR}_{R_{i,j} \to M}(\vec{m})\}\}$.

Let $S_{B_i \to M}(\vec{m})$, $S_{B_i \to R_{i,j}}(\vec{m})$, and $S_{R_{i,j} \to M}(\vec{m})$ be the SEs of B_i-MS, B_i-$R_{i,j}$, and $R_{i,j}$-MS links, respectively. For the SE-based selection, the direct path is selected if $S_{B_i \to M}(\vec{m}) \geq \max_{R_{i,j}}\{S_{B_i \to R_{i,j} \to M}(\vec{m})\}$; otherwise, the two-hop path through $R_{i,k}$ is selected, where

$$S_{B_i \to R_{i,j} \to M}(\vec{m}) = \frac{1}{\frac{1}{S_{B_i \to R_{i,j}}(\vec{m})} + \frac{1}{S_{R_{i,j} \to M}(\vec{m})}}$$
$$= \frac{S_{B_i \to R_{i,j}}(\vec{m}) \cdot S_{R_{i,j} \to M}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} \quad (9)$$

is the effective SE of the $B_i - R_{i,j}$-MS link [19], and $R_{i,k} = \arg\{\max_{R_{i,j}}\{S_{B_i \to R_{i,j} \to M}(\vec{m})\}\}$.

E. Frequency Reuse Over RS-MS Links

In addition to frequency reuse in cochannel cells, the frequency band for the RS–MS links in a cell can be reused as well, thanks to the spatial separation between RSs. To exploit this advantage, RSs in a cell are divided into L reuse groups, where $L \leq N_i$, $i = 1, \ldots, K$, and each group shares the same radio resource. In addition, the reuse pattern can be specified by the set $\text{G}_i \doteq \{\text{G}_{i,l}\}_{l=1}^L \forall i$, where $\text{G}_{i,l} = \{R_{i,k_l}\}$ is the set of RSs in the lth group of cell i. Obviously, $N_i = \sum_{l=1}^L |\text{G}_{i,l}|$, where $|\text{G}_{i,l}|$ is the cardinality of the set $\text{G}_{i,l}$. The number of reuse groups is the same for all cochannel cells in this paper, and some reuse groups in a particular cell may be empty, that is, it contains no RSs in its most general case.

III. OPTIMIZATION OF SYSTEM PARAMETERS

In this section, the RSs' position and the frequency-reuse pattern are jointly optimized over the design area $\Omega \doteq \bigcup_{i=1}^K \Omega_i$ to maximize the system SE under different QoE criteria and path-selection methods.

A. Objective Function

Let T_Ω and W_Ω be the aggregate throughput and the total allocated bandwidth over the design area Ω, respectively. The system SE S_Ω is then defined by

$$S_\Omega \doteq \frac{T_\Omega}{W_\Omega}. \quad (10)$$

Our objective is to search for the optimal set of RSs' positions $\Upsilon \doteq \{\vec{r}_{i,j}\}$ and the reuse pattern $\text{G} \doteq \{\text{G}_i\}$ so that the SE S_Ω is maximized.

Two types of QoE criteria will be investigated along with two path selection methods (given in Section II-D). One is the FBA, where a fixed amount of bandwidth is allocated to MSs in the cell, and the other is the FTA, where a fixed throughput (data rate) is supported no matter where the MS is located (uniform data rate coverage). FBA and FTA represent two extremes of the QoE criteria, and the real system may lie between these two.

1) FBA: In FBA, a fixed bandwidth per unit area w_t (in hertz per unit area) is allocated to an MS. Therefore, $w_t = w_{B_i \to M}(\vec{m}) = w_{B_i \to R_{i,j}}(\vec{m}) + w_{R_{i,j} \to M}(\vec{m})$ for an MS located at $\vec{m}$, where $w_{B_i \to M}(\vec{m})$ is the bandwidth allocated to the $B_i \to$ MS link, $w_{B_i \to R_{i,j}}(\vec{m})$ is the bandwidth allocated to the $B_i \to R_{i,j}$ link, and $w_{R_{i,j} \to M}(\vec{m})$ is the bandwidth allocated to the $R_{i,j} \to$ MS link, respectively. Furthermore, the total allocated bandwidth over Ω is calculated by

$$W_\Omega = W_{B \to M} + W_{B \to R} + W_{R \to M} \quad (11)$$

where

$$W_{B \to M} = \max_i \{W_{B_i \to M}\} \quad (12)$$

$$W_{B \to R} = \max_i \left\{ \sum_{j=1}^{N_i} W_{B_i \to R_{i,j}} \right\} \quad (13)$$

$$W_{R \to M} = \sum_{l=1}^L \max_i \{W_{\text{G}_{i,l} \to M}\}, \qquad i = 1, \ldots, K \quad (14)$$

are the aggregated bandwidths allocated to the $B \to$ MS, $B \to R$, and $R \to$ MS links, respectively.

In (12)–(14)

$$W_{B_i \to M} = \int_{\vec{m} \in \Omega_{B_i}} w_{B_i \to M}(\vec{m}) dA \quad (15)$$

$$W_{B_i \to R_{i,j}} = \int_{\vec{m} \in \Omega_{R_{i,j}}} w_{B_i \to R_{i,j}}(\vec{m}) dA \quad (16)$$

$$W_{\text{G}_{i,l} \to M} = \max_{R_{i,j} \in \text{G}_{i,l}} \int_{\vec{m} \in \Omega_{R_{i,j}}} w_{R_{i,j} \to M}(\vec{m}) dA \quad (17)$$

where Ω_{B_i} and $\Omega_{R_{i,j}}$ are the coverage areas of B_i and $R_{i,j}$, respectively. Note that given Υ and the path selection method, the coverage areas of Ω_{B_i} and $\Omega_{R_{i,j}}$ are determined under the power settings in (7) and (8). In addition, in (11), we have used the fact that the cochannel regions are allocated with the same bandwidth, and thus, the maximum bandwidth required among the cochannel regions is used in the calculation of the total bandwidth.

Recall that $S_{B_i \to M}(\vec{m})$, $S_{B_i \to R_{i,j}}(\vec{m})$, and $S_{R_{i,j} \to M}(\vec{m})$ are the SEs of the $B_i \to$ MS, $B_i \to R_{i,j}$, and $R_{i,j} \to$ MS links,

respectively. Using these notations, the throughputs per unit area supported for those links are given by

$$t_{B_i \to M}(\vec{m}) = w_{B_i \to M}(\vec{m}) \cdot S_{B_i \to M}(\vec{m}) \qquad (18)$$

$$t_{B_i \to R_{i,j}}(\vec{m}) = w_{B_i \to R_{i,j}}(\vec{m}) \cdot S_{B_i \to R_{i,j}}(\vec{m}) \qquad (19)$$

$$t_{R_{i,j} \to M}(\vec{m}) = w_{R_{i,j} \to M}(\vec{m}) \cdot S_{R_{i,j} \to M}(\vec{m}). \qquad (20)$$

In addition, the effective throughput $t_{B_i \to R_{i,j} \to M}(\vec{m})$ for the two-hop path (via $R_{i,j}$) is obtained by $t_{B_i \to R_{i,j} \to M}(\vec{m}) = \min\{t_{B_i \to R_{i,j}}(\vec{m}), t_{R_{i,j} \to M}(\vec{m})\}$. It is shown in Lemma 1 of the Appendix that $t_{B_i \to R_{i,j}}(\vec{m}) = t_{R_{i,j} \to M}(\vec{m})$ provides the highest spectrum efficiency. Thus

$$w_{B_i \to R_{i,j}}(\vec{m}) = \frac{S_{R_{i,j} \to M}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} \cdot w_t \qquad (21)$$

$$w_{R_{i,j} \to M}(\vec{m}) = \frac{S_{B_i \to R_{i,j}}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} \cdot w_t \qquad (22)$$

$$t_{B_i \to R_{i,j} \to M}(\vec{m}) = \frac{S_{B_i \to R_{i,j}}(\vec{m}) \cdot S_{R_{i,j} \to M}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} \cdot w_t. \qquad (23)$$

Using (18)–(23), the aggregate throughput (in bits per second) can be calculated as

$$T_\Omega = \sum_{i=1}^{K} \left(T_{B_i} + \sum_{l=1}^{L} T_{G_{i,l}} \right) \qquad (24)$$

where

$$T_{B_i} = \int_{\vec{m} \in \Omega_{B_i}} t_{B_i \to M}(\vec{m}) dA \qquad (25)$$

$$T_{G_{i,l}} = \sum_{R_{i,j} \in G_{i,l}} \int_{\vec{m} \in \Omega_{R_{i,j}}} t_{B_i \to R_{i,j} \to M}(\vec{m}) dA. \qquad (26)$$

2) FTA: In FTA, a targeted throughput per unit area t_t (in bits per second per unit area) is supported for an MS, no matter where it is located. To achieve this, the bandwidth allocation for the direct path is

$$w_{B_i \to M}(\vec{m}) = \frac{t_t}{S_{B_i \to M}(\vec{m})}. \qquad (27)$$

For the two-hop path via $R_{i,j}$, from Lemma 2 in the Appendix, the best bandwidth allocation is to make $t_{B_i \to R_{i,j}}(\vec{m}) = t_{R_{i,j} \to M}(\vec{m}) = t_t$. As a result

$$w_{B_i \to R_{i,j}}(\vec{m}) = \frac{t_t}{S_{B_i \to R_{i,j}}(\vec{m})} \qquad (28)$$

$$w_{R_{i,j} \to M}(\vec{m}) = \frac{t_t}{S_{R_{i,j} \to M}(\vec{m})}. \qquad (29)$$

By substituting (27)–(29) into (15)–(17), respectively, the aggregate bandwidth can be calculated as in (11). Furthermore, in this case, we have

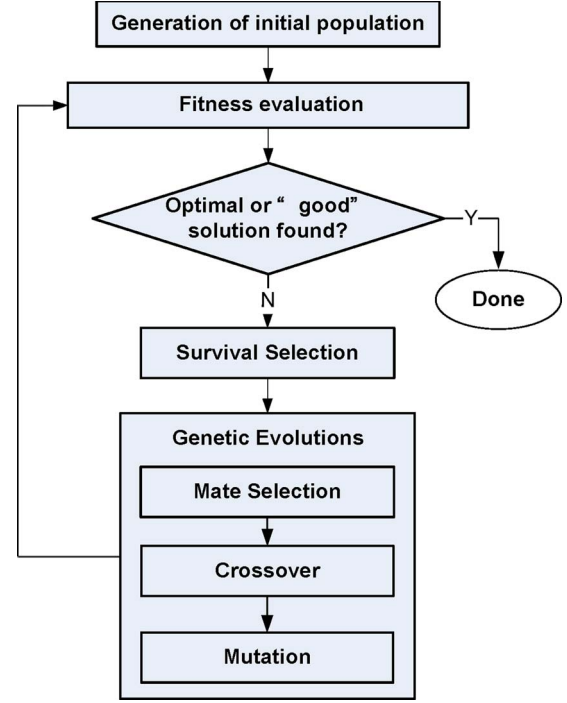$$T_\Omega = \sum_{i=1}^{K} \int_{\vec{m} \in \Omega_i} t_t dA. \qquad (30)$$

Fig. 4. General operation principle of GAs.

Note that $W_{B \to M}$, $W_{B \to R}$, and $W_{R \to M}$ in (12)–(14) are derived under the assumption of a fully loaded system, that is, there is one user per unit area. A practical radio-resource allocation, as illustrated in Fig. 2(b), can subsequently be done, depending on how many users are to be supported in a cell. Let $W_{\rm req}$ be the bandwidth required to support a total of $N_{\rm user}$ users in a cell. For FBA, $W_{\rm req} = N_{\rm user} \cdot W_t$, where W_t (in hertz) is the bandwidth allocated an MS, whereas for FTA, $W_{\rm req} = N_{\rm user} \cdot T_t / S_\Omega$, where T_t (in bits per second) is the target throughput for an MS. Then, the bandwidth for the $B \to$ MS, $B \to R$, and $R \to$ MS links can, respectively, be allocated as

$$\frac{W_{B \to M}}{W_{B \to M} + W_{B \to R} + W_{R \to M}} \cdot W_{\rm req}$$

$$\frac{W_{B \to R}}{W_{B \to M} + W_{B \to R} + W_{R \to M}} \cdot W_{\rm req}$$

$$\frac{W_{R \to M}}{W_{B \to M} + W_{B \to R} + W_{R \to M}} \cdot W_{\rm req}.$$

B. Optimization Algorithm

It is easy to see from (10) that S_Ω is a highly nonlinear function of Υ and G, given a QoE criterion and a path-selection method, and generally, an analytic close-form solution is not available. In this section, a GA-based optimization algorithm is devised to solve the optimization problem.

GA is an evolutionary algorithm invented by Holland in the 1970s, apparently inspired by Darwin's theory of natural evolution, and has widely been recognized as a powerful technique to find exact or approximate solutions to a wide range of optimization and search problems. The general operation of GAs is illustrated in Fig. 4. Assume that the parameters of the targeted objective function can be encoded as genes and that

the set of genes consists of a chromosome that is a candidate solution to the objective function. Initially, GA is started with a random generation of a set of chromosomes called population. Each chromosome will be evaluated against the objective function. Then, a survivor selection criterion is applied to select better-fitness chromosomes for evolution if the best chromosome so far is considered not satisfying. The evolution operations include mate selection, crossover, and mutation. Mate selection picks chromosome mates (parents) from the survivors. Crossover (recombination) is then carried out over the selected mates to reproduce new offspring who inherits partial genes from its parents. Mutation is an operation performed to produce a few offspring where some genes may be altered. The evolution process repeats from one generation to another until some condition (for example, number of iterations or improvement of the best solution) is met.

GA has successfully been applied to a wide range of optimization problems that involve a large number of variables [21]–[24]. In our problem, the GA-based approach is particularly useful when the total number of RSs is large, where optimization based on an exhaustive search is quite inefficient.

1) Proposed GA: In this section, a more detailed description on the proposed GA is provided.

1) Chromosome representation: The Cartesian coordinate and the reuse group of $R_{i,j}$ $(\vec{r}_{i,j}, G_{i,l})$ are encoded as a gene, and the set of RSs' positions and the corresponding reuse pattern (Υ, G) are a chromosome. In our formulation, the cell region is discretized into grids with $\vec{r}_{i,j}$ located at the grid intersection points.

2) Initial population: In the beginning, N_{pop} chromosomes are randomly generated and used as the population size involved in evolution at each generation. Each randomly generated coordinate is rounded into the nearest grid-intersection point.

3) Fitness evaluation: Each chromosome is evaluated against the objective function in (10) under different QoE criteria, path-selection methods, and reuse patterns.

4) Survivor selection: Chromosomes with better fitness survive and will be involved in the evolution in the next generation, whereas those with less fitness are discarded. The selection rate to survive is denoted as β so that only the best $N_{\mathrm{sur}} = \beta \cdot N_{\mathrm{pop}}$ chromosomes survive after evaluation, and the rest of the chromosomes are discarded to make room for the new offspring.

5) Mate selection: Two mates (chromosomes) are selected from the mating pool of N_{sur} chromosomes to produce two new offspring. We apply the most commonly used mate-selection scheme in GA applications: Roulette wheel selection [22], where the survival chromosome is chosen to be one of the parents with the probability

$$P(\Upsilon_k, G_k) = \frac{S_\Omega(\Upsilon_k, G_k)}{\sum_{\Upsilon_i} S_\Omega(\Upsilon_i, G_i)} \qquad (31)$$

where $S_\Omega(\Upsilon_k, G_k)$ is the fitness value for the particular chromosome Υ_k. Roulette wheel selection emulates the survival of the fittest mechanism in nature, where the fitter chromosome (to have good genes to survive) will

have a higher chance to be selected to create offspring. The offspring that inherits good genes from parents may have a higher probability of survival in the subsequent generation.

6) Crossover: After two distinct chromosomes (mates) are selected, a crossover is performed to produce two new offspring. The uniform crossover is adopted in this paper, where for each gene in offspring, it is randomly decided whether to interchange information between the two mates [22], [23]. In addition, in a real-valued encoding scheme, a small zero-mean random perturbation is suggested to add to each gene of the offspring [24] to prevent the evolutions being dominated by a few genes. In our method, the Gaussian variable with variance equal to a grid length is used as the perturbation. Sufficient offspring are produced until the total number of survivors and offspring is equal to N_{pop}. The new population is then involved in the next step evolution.

7) Mutation: In addition to the crossover, mutation is another way to force GA to explore other areas of the solution space to avoid overly fast converging into a local optimum. A mutation probability P_{mut} is set to determine whether a gene is mutated or not. For the case of no intracell frequency reuse, the coordinate of the RS (the gene) is randomly regenerated for each mutation. For the case of intracell frequency reuse, on the other hand, the mutation is done with half a chance to regenerate the RS's coordinate and with the other half to swap the coordinate with one randomly chosen from other reuse groups in the cell. This design maintains the genetic diversity and usually leads a faster convergence to the optimum result.

In the proposed GA-based optimization, the computational complexity is dominated by the complexity in evaluating the objective function in (10), which has to be evaluated N_{pop} times in each iteration. Furthermore, the computational complexity of (10) is of $O((N+1) \cdot N_{\mathrm{grid}} \cdot K)$, where N_{grid} is the number of partitioned grids in a cell because the radio link information (spectrum efficiency, consumed bandwidth, received SINR, etc.) has to be calculated between RS (BS) and MS, which is located at one of the grids in the cell. Therefore, the total complexity of the proposed method is $O(N_{\mathrm{ite}} \cdot N_{\mathrm{pop}} \cdot (N+1) \cdot N_{\mathrm{grid}} \cdot K)$, where N_{ite} is the number of iterations in GA.

IV. NUMERICAL RESULTS

In our numerical results, the cell radius is set to be 1400 m, the cell is divided into grids with each side equal to 20 m, and all stations (BS, RSs, and MS) are located at the grid-intersection points. Four system configurations will be considered, including FBA with SE-based path selection (FBA-SE), FBA with SINR-based path selection (FBA-SINR), FTA with SE-based path selection (FTA-SE), and FTA with SINR-based path selection (FTA-SINR). All of the results are obtained for AWGN channels with $w_t = 1$ Hz per unit area for FBA and $t_t = 1$ b/s per unit area for FTA. In addition, $S_{\mathrm{edge}} = 0.5$ (b/s/Hz) for all results unless otherwise specified, and $N_i = N$ $\forall i$. In GA, clearly, the control parameters, such as initial population, survivor-selection scheme, mutation probability,
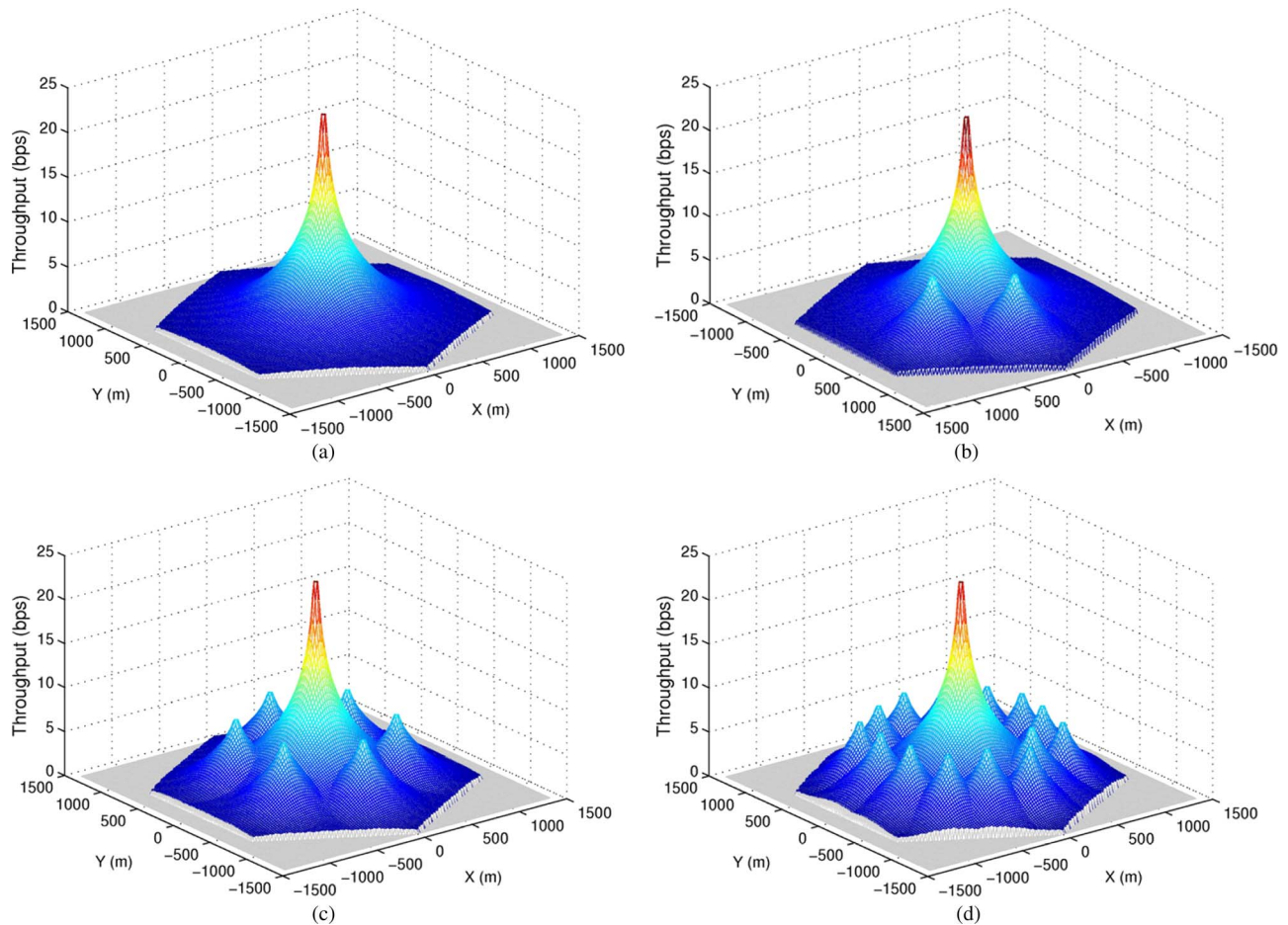
Fig. 5. Optimal RSs' positions and throughput distribution for FBA-SE. (a) $N = 0$. (b) $N = 2$. (c) $N = 6$. (d) $N = 14$.

crossover mechanism, etc., affect the convergence of GA. To the best of our knowledge, the conditions for GAs to converge have been proved only for the binary encoding type of GAs with infinite iteration number by using Markov chain models [25], [26]. For the case of finite iteration number and/or integer or real encoding, the convergence of the algorithm is still an open problem. In this paper, instead of using an analytical approach, extensive simulations are employed to look at the convergence issue. As shown in Section IV-A, the convergence of the proposed GA is quite insensitive to the control parameters $(N_{\text{pop}}, \beta, P_{\text{mut}})$ if they are selected to within a proper range of values. In all our simulations, $N_{\text{pop}} = 200$, $\beta = 0.5$, and $P_{\text{mut}} = 0.05$. In addition, the number of iterations of GA depends on the number of RSs involved; more generations are needed for a larger number of RSs.

A. Single Cell

Fig. 5 illustrates how RSs are placed in an optimal way and how they can improve the SE and user throughput for FBA-SE. The BS is equipped with an omnidirectional antenna, and no frequency reuse and shadowing are considered. Fig. 5(a) shows the case of no RS $(N = 0)$, where the maximum throughput is 23.37 b/s (per unit area) at the locations close to the BS, and the minimum throughput is 0.5 b/s at the cell vertices (as planned at the power setting of BS). Clearly, the achievable

throughput rapidly decreases from the cell center to the edge due to path loss. Fig. 5(b) shows the case of two RSs, where the RSs are placed (after optimization) on the lines connecting the cell center and two vertices of the hexagonal cell at distance of 680 m from BS. As can be seen, user throughput is significantly improved at close proximity of RSs. Note that the optimal placement of RSs is not unique in this case; as long as RSs are placed on the lines connecting the cell center and any two of the vertices at the same distance from the BS, the system performance is the same. Fig. 5(c) illustrates the case of six RSs. Again, RSs are placed on the lines connecting the cell center and vertices, and all of them locate on the ring with distance of about 680 m from BS. As the number of RS increases, however, the optimal RS locations may split into more than one ring, as one might expect. Fig. 5(d) is such an example with 14 RSs.

Table I(a) summarizes the important system parameters for FBA-SE, including the service area ratio of BS and RSs, the bandwidth allocation ratio for the BS–MS link, the BS–RS and RS–MS links, and the system SE. As shown, the system SE is significantly improved with the deployment of RSs, which is around 4.23% for each addition of RS when $N \leq 6$. The improvement becomes smaller, however, as the number of RSs is larger than 6, where RSs begin to compete with each other for serving MS rather than to compete with BS. Fig. 6 shows the cumulative distribution function (cdf) of the user throughput for

TABLE I
IMPORTANT PARAMETERS FOR FOUR SYSTEM SETUPS

|	N	0	2	4	6	8	10	12	14
(a) FBA-SE	$\Omega_B{:}\Omega_R$ (%)	100:0	77:23	54:46	31:69	23:77	22:78	21:79	22:78
	$W_{B\to M}{:}\,W_{B\to R}{:}W_{R\to M}$ (%)	100:0:0	77:5:18	54:9:37	31:14:55	23:17:60	22:18:60	21:19:60	22:20:58
	S_Ω (bps/Hz)	2.7382	2.9698	3.2014	3.4329	3.6398	3.7986	3.9142	4.0024
(b) FBA-SINR	$\Omega_B{:}\Omega_R$ (%)	100:0	75:25	52:48	27:73	19:81	18:82	18:82	20:80
	$W_{B\to M}{:}\,W_{B\to R}{:}W_{R\to M}$ (%)	100:0:0	75:5:20	52:10:38	27:15:58	19:18:63	18:19:63	18:20:62	20:20:60
	S_Ω (bps/Hz)	2.7382	2.9605	3.1821	3.4043	3.6148	3.7773	3.8960	3.9848
(c) FTA-SE	$\Omega_B{:}\Omega_R$ (%)	100:0	82:18	65:35	47:53	36:64	31:69	31:69	32:68
	$W_{B\to M}{:}\,W_{B\to R}{:}W_{R\to M}$ (%)	100:0:0	84:2:14	66:5:29	43:8:49	30:11:59	22:13:65	18:14:68	20:15:65
	$\overline{w}$ (Hz)	0.6423	0.5906	0.5390	0.4874	0.4500	0.4135	0.3826	0.3647
	S_Ω (bps/Hz)	1.5570	1.6932	1.8552	2.0516	2.2221	2.4181	2.6134	2.7421
(d) FTA-SINR	$\Omega_B{:}\Omega_R$ (%)	100:0	82:18	63:37	44:56	34:66	28:72	28:72	29:71
	$W_{B\to M}{:}\,W_{B\to R}{:}W_{R\to M}$ (%)	100:0:0	83:3:14	63:5:32	39:9:52	27:11:62	17:14:69	15:15:70	17:15:68
	$\overline{w}$ (Hz)	0.6423	0.5913	0.5403	0.4895	0.4518	0.4152	0.3835	0.3662
	S_Ω (bps/Hz)	1.5570	1.6912	1.8508	2.0429	2.2134	2.4087	2.6073	2.7311
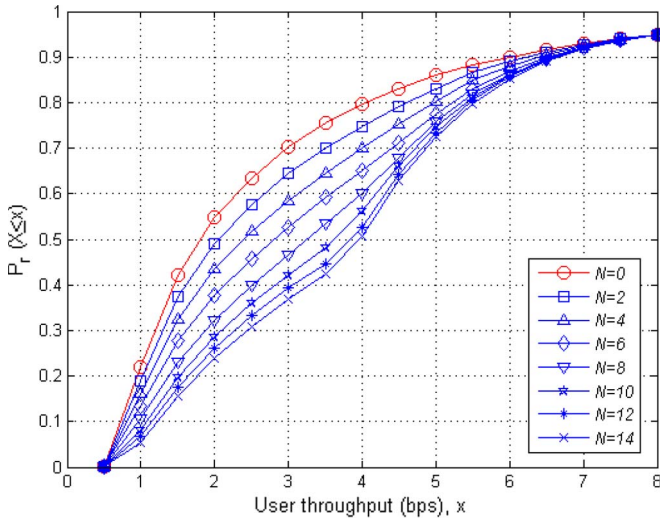
Fig. 6.   CDF of user throughput for FBA-SE.

Fig. 7.   Convergence behavior of the proposed GA for FBA-SE ($N = 6$) under different control parameters ($N_{\text{pop}}, \beta, P_{\text{mut}}$).

different N. Clearly, the percentage of user with low throughput decreases as N increases. In other words, the achievable data rate is distributed more uniformly over the cell. In addition, it is observed that the improvement becomes smaller when $N \geq 10$.

Fig. 7 shows simulations to illustrate the convergence behavior of the proposed GA for FBA-SE with $N = 6$. As can be seen, the convergence of the proposed GA algorithm is quite insensitive to the control parameters ($N_{\text{pop}}, \beta, P_{\text{mut}}$) if they are selected to within a proper range of values. In addition, we investigate the convergence behavior for the fixed set of control parameters ($N_{\text{pop}} = 200, \beta = 0.5, P_{\text{mut}} = 0.05$) under different initial populations (seeds). For example, for FBA-SE ($N = 6$), our result shows that with a total number of 50 runs, the empirical mean and standard deviation of the reached values are 3.432926 and 3.9e-5, respectively. This says that the proposed GA algorithm converges very well in this case. Similar convergence behaviors are observed for different system setups.

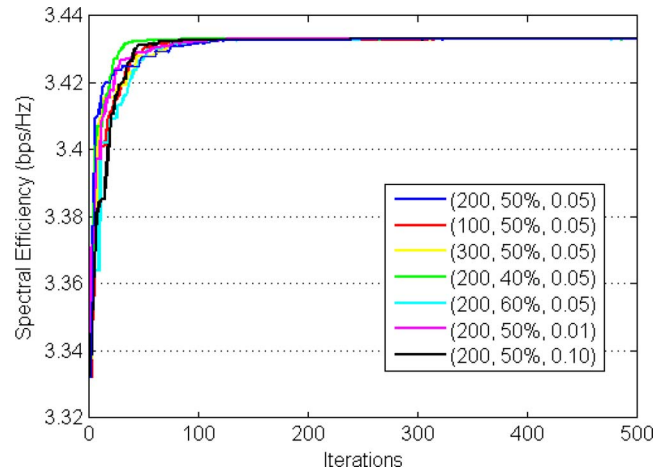As for FBA-SINR, the optimal RSs are placed in a similar way to FBA-SE, except that the RSs are generally a bit farther from BS than that of FBA-SE. Table I(b) shows a summary of important system parameters. Somewhat surprisingly, FBA-SINR performs very close to that of FBA-SE with respect to system SE after optimization with GA. The reason is that for a large proportion of two-hop users, the BS–RS link is much better than the RS–MS links due to the LOS condition in the BS–RS link, and when $S_{B_i \to R_{i,j}}(\vec{m})$ is much greater than $S_{R_{i,j} \to M}(\vec{m})$, the SE-based path selection degenerates to the SINR-based path selection, as can be seen in (9). Note that the SINR-based path selection is much simpler to implement than the SE-based path selection.

Fig. 8 shows the complementary cdf of bandwidth consumption for FTA-SE. As can be seen, the percentage of high bandwidth consumption is reduced as the number of RSs increases. Table I(c) summarizes the important parameters for FTA-SE, and those for FTA-SINR are summarized in Table I(d), where $\bar{w}$ is the average bandwidth consumption. Again, FTA-SINR performs close to FTA-SE.

Fig. 9 compares the system SE of the four system configurations. As can be seen, the system SE of FTA is much less than that of FBA in all cases. To explain this, consider a system with m users who have the link SE of $s_1, s_2, \ldots, s_m$, $\forall s_i \geq 0$, respectively. For FBA, the system SE is given by the
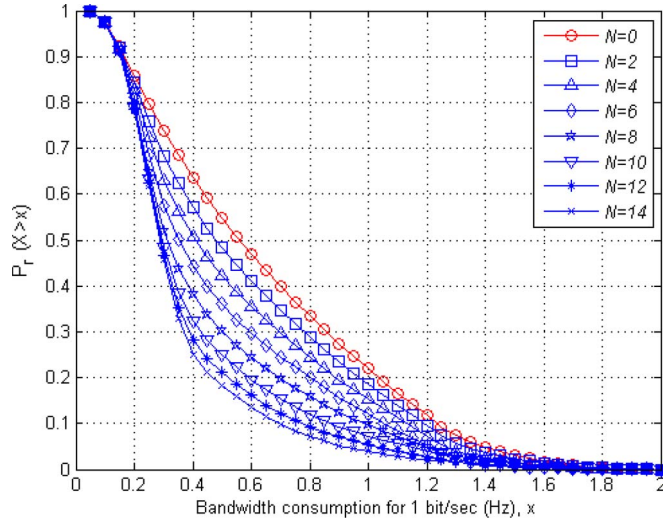
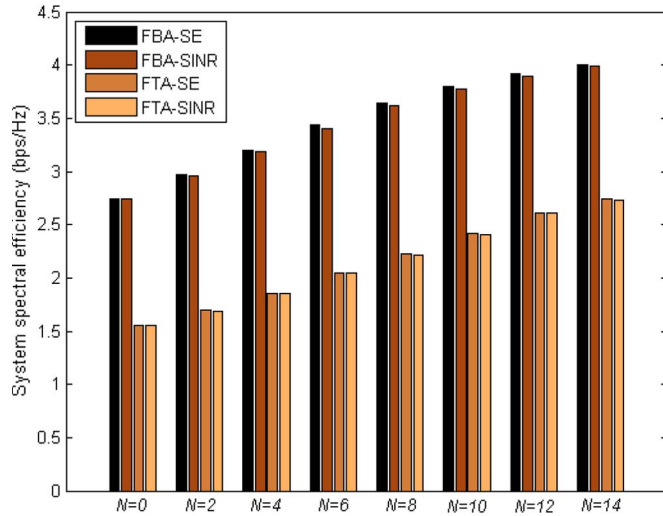Fig. 8.   Complementary cdf of bandwidth consumption for FTA-SE.

Fig. 9.   System SE of four system configurations.

arithmetic mean $(1/m)\sum_{i=1}^{m}s_i$, whereas for FTA, it is given by the harmonic mean $1/(1/m)\sum_{i=1}^{m}1/s_i$. Clearly

$$\frac{1}{m}\sum_{i=1}^{m}s_i \geq \frac{1}{\frac{1}{m}\sum_{i=1}^{m}1/s_i}. \tag{32}$$

The equality holds if and only if $s_1 = s_2 = \cdots = s_m$; a condition does not hold in our cases. Note that the harmonic mean is dominated by the smallest s_i. In other words, the system capacity may significantly be lowered if a fixed throughput is to be maintained for those users in a bad channel condition, that is, uniform data rate coverage comes at a large expense on system capacity.

1) Frequency Reuse Over RS-MS Links: To better utilize the radio resource, frequency reuse over the RS–MS links is explored here. The case of six RSs is taken as an example with the following reuse patterns, $G = \{G_1, G_2, G_3, G_4, G_5, G_6\}$, $G = \{G_1, G_2, G_3\}$, $G = \{G_1, G_2\}$, and $G = \{G_1\}$, where each reuse group has equal number of RSs. Fig. 10 depicts the throughput distribution of FBA-SE. Circles with the same letter

indicate the service area of RSs in the same reuse groups. Again, RSs are placed on the line connecting the cell center to six vertexes of the cell, with the RSs in the same reuse group being pulled away as far as possible. In the fully reuse case, as shown in Fig. 10(d), severe interference leads to a shrinking of the service area of RSs. With respect to system SE, the reuse cases always outperform the no reuse cases, as shown in Fig. 11, where the reuse pattern $G = \{G_1, G_2\}$ gives the highest SE. In addition, it is very interesting to see that SINR-based path selection has a higher SE than SE-based path selection for both FBA and FTA in the frequency-reuse cases. This can be attributed to the fact that RSs always have a larger service area with the SINR-based path selection, which leads to a larger bandwidth saving in the frequency-reuse case. Table II summarizes the important system parameters for four different system configurations. Note that, in this example, the reuse pattern $G = \{G_1, G_2, G_3\}$ gives the largest service area of RS, whereas $G = \{G_1, G_2\}$ gives the highest system SE, that is, there is a tradeoff between MAI and frequency reuse.

2) Shadowing Effects: In this section, an example is given to verify the effectiveness of the proposed method in a shadowed environment. First, we show how the shadowing effect degrades the system performance under the same power setup criteria in Section IV-A. The simplified shadow model in Fig. 3 is adopted with four obstacles (of length 800 m) centered at $(0, \pm750)$ and $(\pm750, 0)$. Two sets of shadowing loss $\delta = 10$ dB and $\delta = 20$ dB are considered [see (3)]. In Fig. 12(a) and (c), the achievable throughput in the shadowed area is relatively low, as one might expect. Then, four RSs are deployed into the cell. FBA-SE is considered in this case. The optimization algorithm suggests that RS should be placed at $(0, \pm760)$ and $(\pm760, 0)$, which are right behind the four obstacles. The throughput enhancement is shown in Fig. 12(b) and (d) for $\delta = 10$ dB and $\delta = 20$ dB, respectively. In Fig. 12(b), in addition to serving the severely shadowed area, RSs also provide service to small fractions of the area across the obstacle, where the throughput provided by RSs is higher than that by BS. In Fig. 12(d), we can see that this area is reduced as the shadowing loss increases. As can be seen, in a relay-assisted cellular system, the shadowing effects can largely be removed by deploying RSs.

3) Sectorization: In this section, we investigate how RSs are placed and how they can improve the system performance when the BS is equipped with a sectored antenna. A three-sector antenna is employed with the antenna pattern given in (4). FTA-SE is taken as an example. In the sectorized cell, the power of the BS is set to achieve the targeted SE $S_{\text{edge}} = 0.5$ b/s/Hz for users at the worst locations, which are the cell vertices on the sector boundary. Notice that with this power setting, the users at the cell vertices in the antenna-bearing direction can enjoy higher SE. Fig. 13(a) shows the distribution of bandwidth consumption for achieving $t_t = 1$ b/s (per unit area) with no RS in the cell. As can be seen, users at the cell vertices on the sector boundary need 2 Hz to support the targeted throughput, whereas users at the cell vertices in the antenna bearing direction only need 0.5 Hz. The system SE S_Ω in this case is 2.92 b/s/Hz.

Fig. 13(b)–(d) shows the results for using three RSs. No frequency reuse over the RS–MS links is considered. Fig. 13(b) shows the optimal RS positions and the distribution of
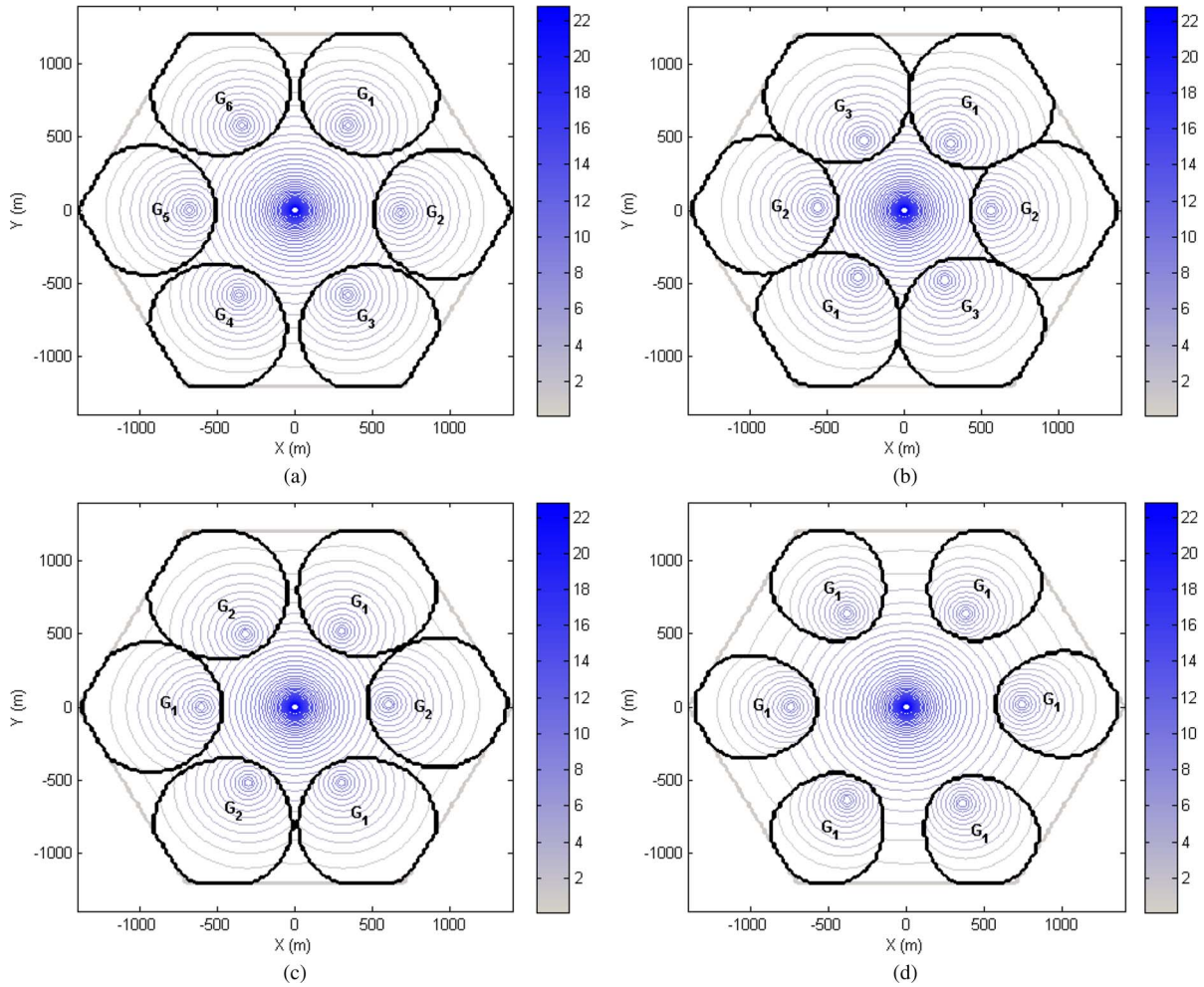
Fig. 10. Throughput distribution (in bits per second per unit area) of different frequency-reuse patterns for FBA-SE ($N = 6$). (a) G = $\{G_1, G_2, G_3, G_4, G_5, G_6\}$. (b) G = $\{G_1, G_2, G_3\}$. (c) G = $\{G_1, G_2\}$. (d) G = $\{G_1\}$.
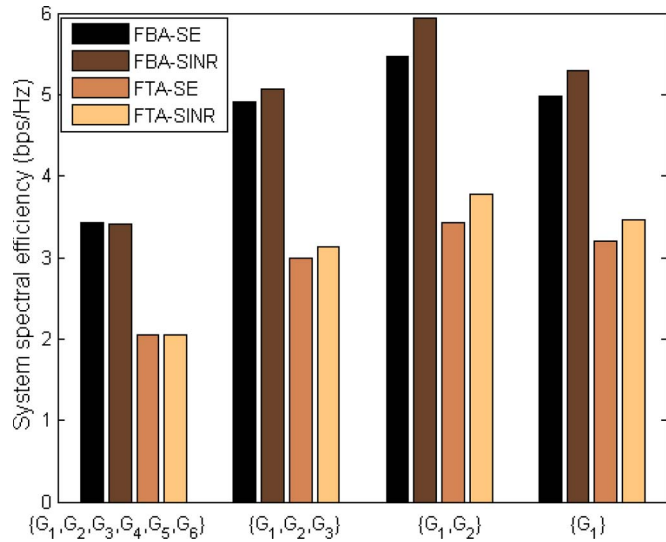

Fig. 11. System SE of different frequency-reuse patterns ($N = 6$).

bandwidth consumption for $S_{\text{edge}} = 0.5$ b/s/Hz for RSs. As expected, RSs tend to assist the relatively low link SE areas. That is, RSs are placed on the sector boundaries to reduce the overall bandwidth consumption. From our numerical results, RSs serve about 13% of the total area, and S_Ω is 3.32 b/s/Hz,

which is a 14% improvement as compared with the case in Fig. 13(a). Fig. 13(c) and (d) shows the results for $S_{\text{edge}} = 1.0$ b/s/Hz and $S_{\text{edge}} = 1.5$ b/s/Hz for RSs, respectively. Again, RSs are placed on the sector boundaries to enhance the system SE. In these cases, RSs provide 22% and 30% of the service area, and S_Ω are increased to 3.61 and 3.83 b/s/Hz, respectively.

B. Multicell

In this section, we use an example to illustrate how the joint optimization is done over multiple cells. We consider the cell structure of (3, 3, 3) (see Fig. 1), where a total of nine RSs are deployed in the cells Ω_1, Ω_2, and Ω_3 with three RSs for each cell. For BSs, the power is set to achieve $S_{\text{edge}} = 0.5$ b/s/Hz for the users located at the cell vertices on the sector boundary, and for RSs, the setting is to achieve $S_{\text{edge}} = 1.0$ b/s/Hz. The cochannel interference from the first-tier cochannel cells is taken into consideration. Fig. 14(a)–(c) shows the optimal placement of RSs and the distribution of bandwidth consumption in Ω_1, Ω_2, and Ω_3, respectively, for FTA-SE. As expected, RSs are placed on the sector boundaries to best improve the SE, and the cells have a similar layout, where RSs in the same reuse group are placed as far apart as possible to reduce the cochannel interference. In addition, our numerical results show

TABLE II
IMPORTANT PARAMETERS FOR FREQUENCY-REUSE PATTERNS ($N = 6$)

|	G	$\{G_1, G_2, G_3, G_4, G_5, G_6\}$	$\{G_1, G_2, G_3\}$	$\{G_1, G_2\}$	$\{G_1\}$
(a) FBA-SE	$\Omega_B{:}\Omega_R$ (%)	31:69	21:79	28:72	49:51
	$w_{B\to M}{:}\, w_{B\to R}{:}w_{R\to M}$ (%)	31:14:55	31:22:47	45:23:32	74:16:10
	S_Ω (bps/Hz)	3.4329	4.9120	5.4713	4.9848
(b) FBA-SINR	$\Omega_B{:}\Omega_R$ (%)	27:73	14:86	16:84	43:57
	$w_{B\to M}{:}\, w_{B\to R}{:}w_{R\to M}$(%)	27:15:58	21:26:53	29:29:42	69:18:13
	S_Ω (bps/Hz)	3.4043	5.0586	5.9428	5.2868
(c) FTA-SE	$\Omega_B{:}\Omega_R$ (%)	47:53	25:75	30:70	50:50
	$w_{B\to M}{:}\, w_{B\to R}{:}w_{R\to M}$ (%)	43:8:49	31:16:53	44:17:39	75:12:13
	S_Ω (bps/Hz)	2.0516	2.9570	3.4322	3.2042
(d) FTA-SINR	$\Omega_B{:}\Omega_R$ (%)	44:56	16:84	18:82	43:57
	$w_{B\to M}{:}\, w_{B\to R}{:}w_{R\to M}$(%)	39:9:52	19:19:62	28:22:50	69:15:16
	S_Ω (bps/Hz)	2.0429	3.1383	3.7729	3.4575
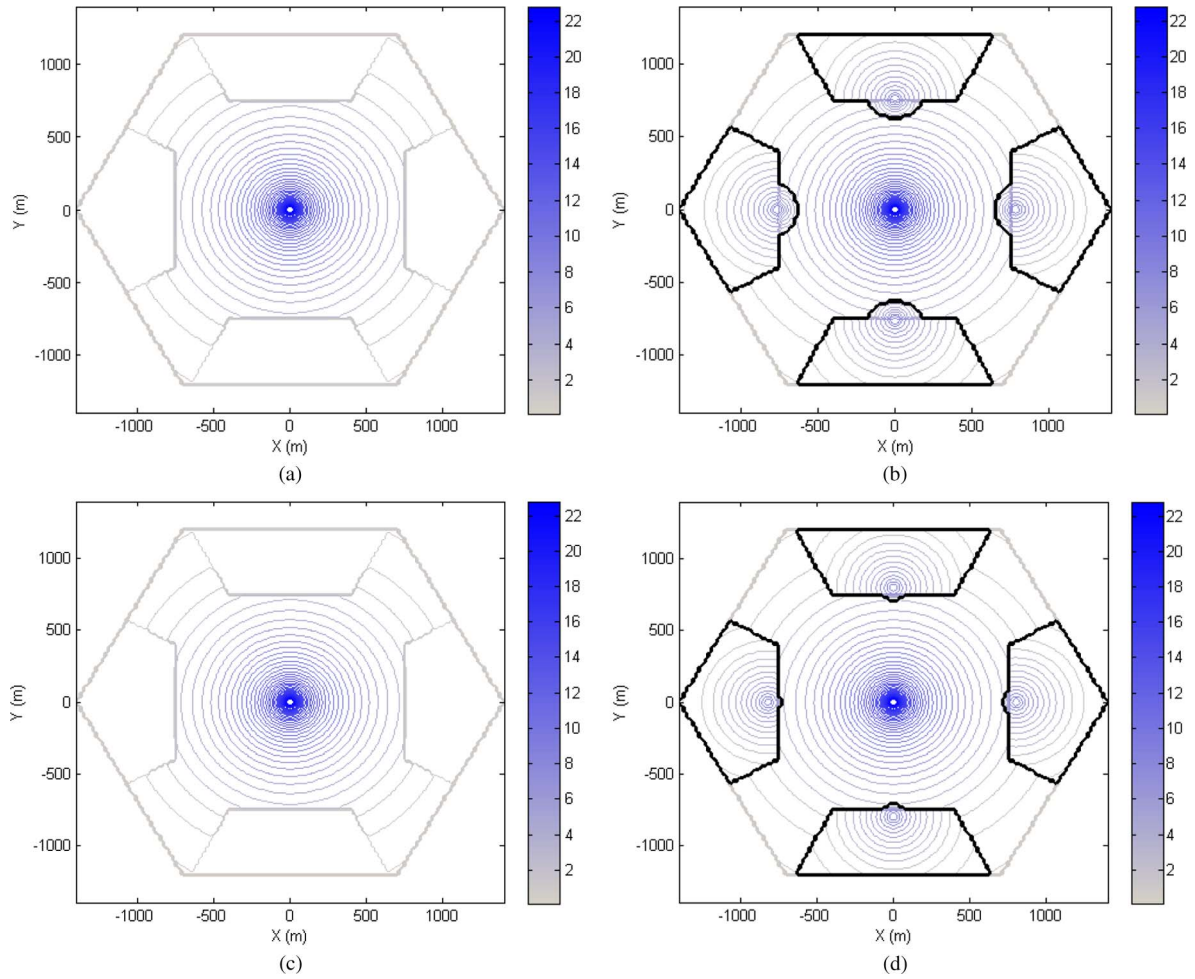
Fig. 12. Throughput distribution (in bits per second per unit area) of FBA-SE under shadowing effects. (a) $N = 0$, $\delta = 10$ dB. (b) $N = 4$, $\delta = 10$ dB. (c) N = 0, $\delta = 20$ dB. (d) N = 4, $\delta = 20$ dB.

that $S_\Omega = 2.94$ b/s/Hz in the case of no RS, and it is improved to $S_\Omega = 3.60$ b/s/Hz with three RSs deployed, where RSs serve about 20% of the total area of each cell.

V. CONCLUSION

In this paper, the optimization and theoretical performance of relay-assisted cellular networks have been investigated in a multicell environment. A GA is proposed for the joint multicell optimization of system parameters, including RS's positions, path selection, frequency-reuse pattern, and resource allocation, to maximize system SE. Four different system configurations are considered, which includes FBA-SE (FBA with SE-based path selection), FBA-SINR (FBA with SINR-based path selection), FTA-SE (FTA with SE-based path selection), and FTA-SINR (FTA with SINR-based path selection). Numerical results
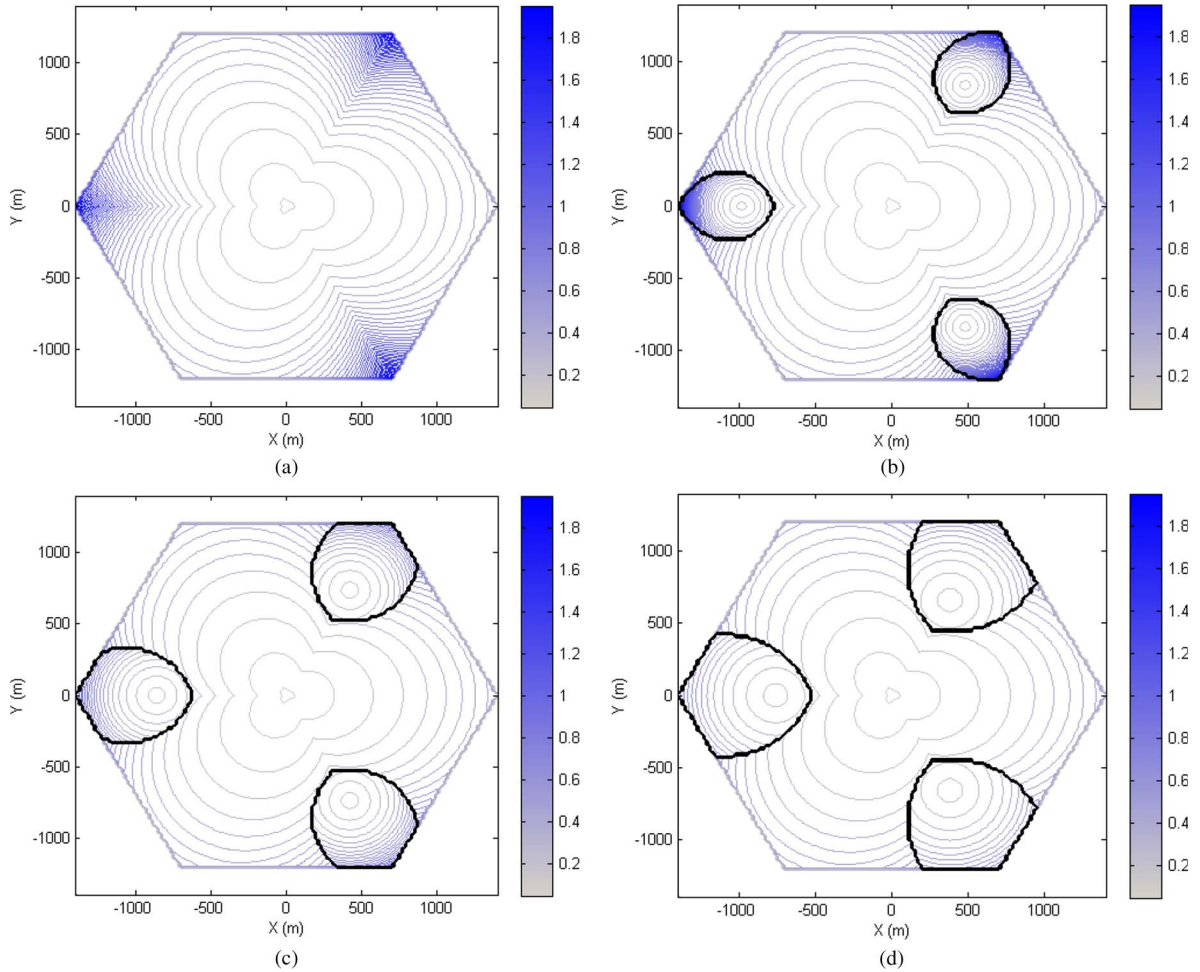
Fig. 13. Optimal RSs' positions and bandwidth consumption (in hertz per unit area) for FTA-SE in a three-sector cell. (a) $N = 0$. (b) $N = 3$, $S_{\text{edge}} = 0.5$ b/s/Hz for RSs. (c) $N = 3$, $S_{\text{edge}} = 1.0$ b/s/Hz for RSs. (d) $N = 3$, $S_{\text{edge}} = 1.5$ b/s/Hz for RSs.

show the following: 1) RSs provide significant improvement with respect to system SE and user throughput over traditional cellular networks. 2) The uniformity of user data rate comes at the expense of a large loss in system SE when FTA is employed. 3) Somewhat surprisingly, the low-complexity SINR-based path selection performs nearly as well as the SE-based path selection for the no-reuse case, while it is slightly better in the frequency-reuse case.

APPENDIX

In this Appendix, we prove the following two lemmas that state the conditions for the best bandwidth allocation in a two-hop path.

Lemma 1: For the FBA, the highest SE of a two-hop path is achieved with the following bandwidth allocation:

$$w_{B_i \to R_{i,j}}(\vec{m}) = \frac{S_{R_{i,j} \to M}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} \cdot w_t \quad (A.1)$$

$$w_{R_{i,j} \to M}(\vec{m}) = \frac{S_{B_i \to R_{i,j}}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} \cdot w_t \quad (A.2)$$

where $S_{B_i \to R_{i,j}}(\vec{m})$ and $S_{R_{i,j} \to M}(\vec{m})$ are the link SEs of the $B_i \to R_{i,j}$ link and the $R_{i,j} \to M$ link, respectively.

Proof: Recall that, for the FBA, $w_t = w_{B_i \to R_{i,j}}(\vec{m}) + w_{R_{i,j} \to M}(\vec{m})$ for a two-hop path. Let

$$w_{B_i \to R_{i,j}}(\vec{m}) = \left(\frac{S_{R_{i,j} \to M}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} + \alpha \right) \cdot w_t \quad (A.3)$$

be the bandwidth allocated to the $B_i \to R_{i,j}$ link, where $0 \le \alpha \le 1$. Then, the bandwidth for the $R_{i,j} \to M$ link is

$$w_{R_{i,j} \to M}(\vec{m}) = \left(\frac{S_{B_i \to R_{i,j}}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} - \alpha \right) \cdot w_t. \quad (A.4)$$

Thus, the throughput for the $B_i \to R_{i,j}$ link and the $R_{i,j} \to M$ link is, respectively, given by

$$t_{B_i \to R_{i,j}}(\vec{m}) = \left(\frac{S_{R_{i,j} \to M}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} + \alpha \right)$$
$$\cdot w_t \cdot S_{B_i \to R_{i,j}}(\vec{m}) \quad (A.5)$$

$$t_{R_{i,j} \to M}(\vec{m}) = \left(\frac{S_{B_i \to R_{i,j}}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} - \alpha \right)$$
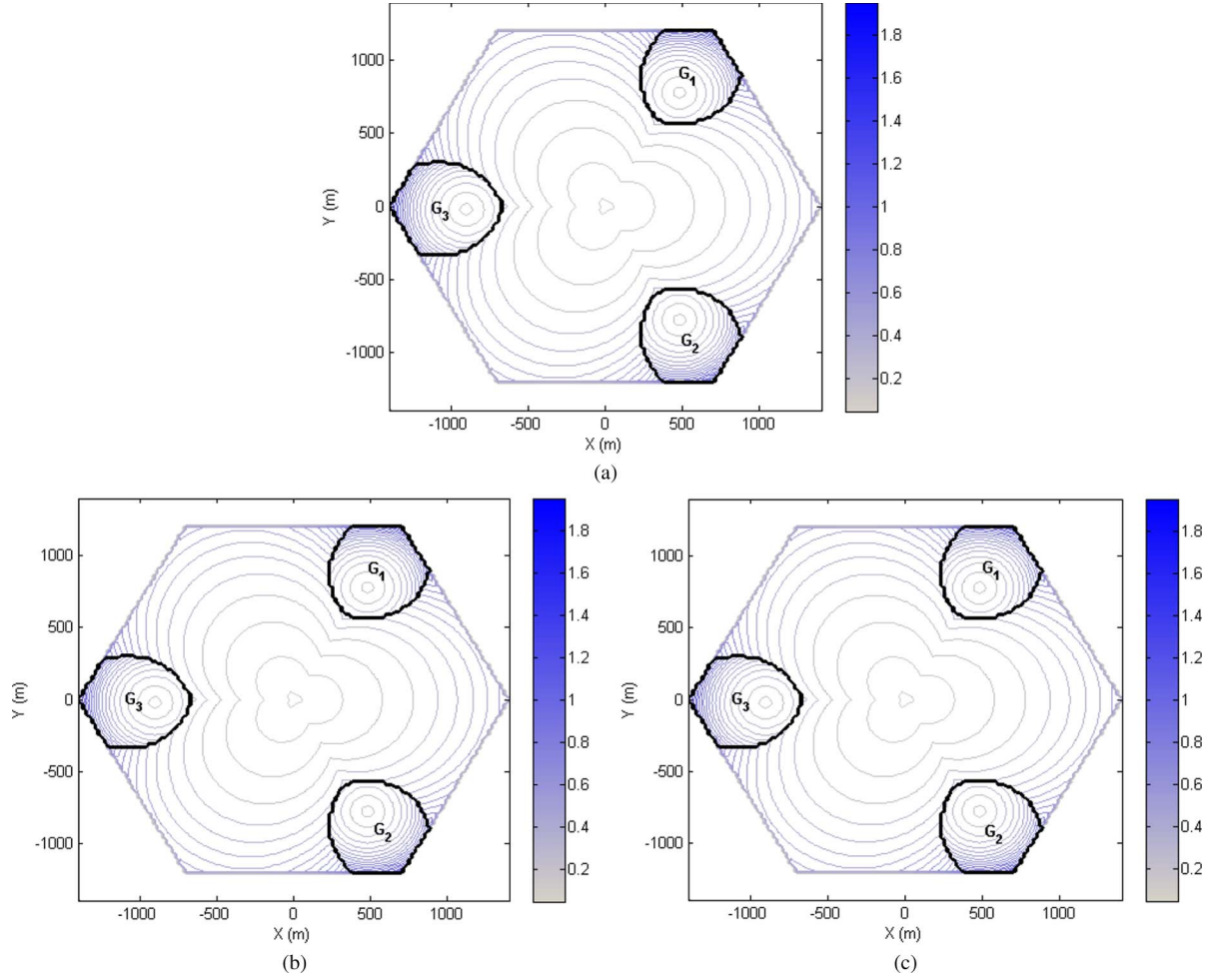$$\cdot w_t \cdot S_{R_{i,j} \to M}(\vec{m}). \quad (A.6)$$

Fig. 14. RSs' positions and bandwidth consumption (in hertz per unit area) with joint optimization of B_1, B_2, and B_3 in Fig. 1 for FTA-SE. (a) Ω_1. (b) Ω_2. (c) Ω_3.

Since $t_{B_i \to R_{i,j} \to M}(\vec{m}) = \min\{t_{B_i \to R_{i,j}}(\vec{m}), t_{R_{i,j} \to M}(\vec{m})\}$

$$t_{B_i \to R_{i,j} \to M}(\vec{m}) \leq \left(\frac{S_{B_i \to R_{i,j}}(\vec{m}) \cdot S_{R_{i,j} \to M}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} \right) \cdot w_t. \tag{A.7}$$

The right side of (A.7) is obtained by setting $\alpha = 0$ in (A.1) and (A.2), i.e.,

$$w_{B_i \to R_{i,j}}(\vec{m}) = \frac{S_{R_{i,j} \to M}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} \cdot w_t \tag{A.8}$$

$$w_{R_{i,j} \to M}(\vec{m}) = \frac{S_{B_i \to R_{i,j}}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} \cdot w_t. \tag{A.9}$$

With $\alpha = 0$, we have the maximum throughput of the two-hop path as

$$t_{B_i \to R_{i,j} \to M}(\vec{m}) = t_{B_i \to R_{i,j}}(\vec{m}) = t_{R_{i,j} \to M}(\vec{m})$$
$$= \frac{S_{B_i \to R_{i,j}}(\vec{m}) \cdot S_{R_{i,j} \to M}(\vec{m})}{S_{B_i \to R_{i,j}}(\vec{m}) + S_{R_{i,j} \to M}(\vec{m})} \cdot w_t. \tag{A.10}$$

■

Lemma 2: For the FTA, the highest SE of a two-hop path to achieve the target throughput t_t is obtained with the following bandwidth allocation:

$$w_{B_i \to R_{i,j}}(\vec{m}) = \frac{t_t}{S_{B_i \to R_{i,j}}(\vec{m})} \tag{A.11}$$

$$w_{R_{i,j} \to M}(\vec{m}) = \frac{t_t}{S_{R_{i,j} \to M}(\vec{m})}. \tag{A.12}$$

Proof: Since $t_{B_i \to R_{i,j} \to M}(\vec{m}) = \min\{t_{B_i \to R_{i,j}}(\vec{m}),$ $t_{R_{i,j} \to M}(\vec{m})\}$, one needs to have

$$t_{B_i \to R_{i,j}}(\vec{m}) = S_{B_i \to R_{i,j}}(\vec{m}) \cdot w_{B_i \to R_{i,j}}(\vec{m}) \geq t_t \tag{A.13}$$

$$t_{R_{i,j} \to M}(\vec{m}) = S_{R_{i,j} \to M}(\vec{m}) \cdot w_{R_{i,j} \to M}(\vec{m}) \geq t_t. \tag{A.14}$$

The proof immediately follows from (A.13) and (A.14). ∎

REFERENCES

[1] R. Pabst, B. H. Walke, D. C. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. D. Falconer, and G. P. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag.*, vol. 42, no. 9, pp. 80–89, Sep. 2004.

[2] H. Hu, H. Yanikomeroglu, D. D. Falconer, and S. Periyalwar, "Range extension without capacity penalty in cellular networks with digital

fixed relays," in *Proc. IEEE Global Telecommun. Conf.*, Nov. 2004, pp. 3053–3057.

[3] O. Oyman, J. N. Laneman, and S. Sandhu, "Multihop relaying for broadband wireless mesh networks: From theory to practice," *IEEE Commun. Mag.*, vol. 45, no. 11, pp. 116–122, Nov. 2007.

[4] V. Genc, S. Murphy, Y. Yu, and J. Murphy, "IEEE 802.16j relay-based wireless access networks: An overview," *Wireless Commun.*, vol. 15, no. 5, pp. 56–63, Oct. 2008.

[5] S. W. Peters and R. W. Heath, "The future of WiMAX: Multihop relaying with IEEE 802.16j," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 104–111, Jan. 2009.

[6] Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems: Multihop Relay Specification, IEEE Std. P802.16j/D9, Feb. 2009.

[7] D. R. Basgeet and Y. C. Chow, "Uplink performance analysis for a relay based cellular system," in *Proc. IEEE VTC—Spring*, May 2006, pp. 132–136.

[8] P. Li, M. Rong, Y. Xue, D. Yu, L. Wang, and H. Shi, "Spectrum partitioning and relay positioning for cellular system enhanced with two-hop fixed relay nodes," *IEICE Trans. Commun.*, vol. E90-B, no. 11, pp. 3181–3188, Nov. 2007.

[9] E. Weiss, S. Max, O. Klein, G. Hiertz, and B. Walke, "Relay-based vs. conventional wireless networks: Capacity and spectrum efficiency," in *Proc. IEEE PIMRC*, Sep. 2007, pp. 1–5.

[10] E. Visotsky, J. Bae, R. Peterson, R. Berry, and M. L. Honig, "On the uplink capacity of an 802.16j system," in *Proc. IEEE WCNC*, Mar./Apr. 2008, pp. 2657–2662.

[11] H. Zeng and C. Zhu, "System-level modeling and performance evaluation of multi-hop 802.16j systems," in *Proc. IEEE IWCMC*, Aug. 2008, pp. 354–359.

[12] S.-J. Kim, S.-Y. Kim, B.-B. Lee, S.-W. Ryu, H.-W. Lee, and C.-H. Cho, "Multi-hop relay based coverage extension in the IEEE 802.16j based mobile WiMAX systems," in *Proc. IEEE Netw. Comput. Adv. Inf. Manag.*, Sep. 2008, pp. 516–522.

[13] C. Hoymann, M. Dittrich, and S. Goebbels, "Dimensioning cellular multi-hop WiMAX networks," in *Proc. IEEE Mobile WiMAX Symp.*, Mar. 2007, pp. 150–157.

[14] J. Vidal, N. Marina, and A. Høst-Madsen, "Dimensioning cellular networks with fixed relay stations," in *Proc. IEEE ICT*, Jun. 2008, pp. 1–5.

[15] B. Lin, P.-H. Ho, L.-L. Xie, and X. Shen, "Optimal relay station placement in IEEE 802.16j networks," in *Proc. IEEE IWCMC*, Aug. 2007, pp. 25–30.

[16] A. Adinoyi *et al.*, Description of Identified New Relay Based Radio Network Deployment Concepts and First Assessment by Comparison Against Benchmarks of Well Known Deployment Concepts Using Enhanced Radio Interface Technology, IST-2003-507581 WINNER D3.1, Nov. 2004.

[17] Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations, 3GPP TR 25.996 V6.1.0, Sep. 2003.

[18] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[19] O. Oyman and S. Sandhu, "Non-ergodic power–bandwidth tradeoff in linear multi-hop networks," in *Proc. IEEE ISIT*, Jul. 2006, pp. 1514–1518.

[20] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: Univ. Michigan Press, 1975.

[21] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley, 1989.

[22] R. L. Haupt and S. E. Haupt, *Practical Genetic Algorithms*. Hoboken, NJ: Wiley, 2004.

[23] K. F. Man, K. S. Tang, and S. Kwong, "Genetic algorithms: Concepts and applications," *IEEE Trans. Ind. Electron.*, vol. 43, no. 5, pp. 519–534, Oct. 1996.

[24] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization*. Hoboken, NJ: Wiley, 2001.

[25] A. Nix and M. D. Vose, "Modeling genetic algorithms with Markov chains," *Ann. Math. Artif. Intell.*, vol. 5, no. 1, pp. 79–88, Mar. 1992.

[26] O. Francois, "An evolutionary strategy for global minimization and its Markov chain analysis," *IEEE Trans. Evol. Comput.*, vol. 2, no. 3, pp. 77–90, Sep. 1998.

Wern-Ho Sheen (M'91) received the B.S. degree from the National Taiwan University of Science and Technology, Taipei, Taiwan, in 1982, the M.S. degree from the National Chiao Tung University, Hsinchu, Taiwan, in 1984, and the Ph.D. degree from the Georgia Institute of Technology, Atlanta, in 1991.

From 1993 to 2001, he was with the National Chung Cheng University, Chiayi, Taiwan, where he was a Professor with the Department of Electrical Engineering and the Managing Director of the Center for Telecommunication Research. From 2001 to 2009, he was a Professor with the Department of Communications Engineering, National Chiao Tung University. He has extensively been consulting for industry and research institutes in Taiwan. He is currently a Professor with the Department of Information and Communication Engineering, Chaoyang University of Technology, Taichung, Taiwan. His research interests include the general areas of communication theory, cellular mobile and personal radio systems, adaptive signal processing for wireless communications, spread spectrum communications, and very-large-scale integrated design for wireless communication systems.
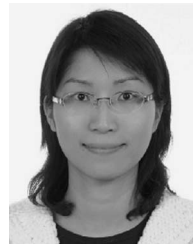
Shiang-Jiun Lin was born in Taiwan. She received the B.S. degree in computer science and information engineering and the M.S. and Ph.D. degrees in communication engineering from the National Chiao Tung University, Hsinchu, Taiwan, in 2000, 2002, and 2010, respectively.

In 2008, she was a Visiting Scholar with the Department of Electrical Engineering, University of Washington, Seattle. She is currently a Research Engineer with the Industrial Technology Research Institute, Hsinchu, Taiwan. She is also with the Department of Communication Engineering, National Chiao Tung University. Her research interests include wireless communication networks, relaying communication technology, and optimizations for wireless communication systems.

Chia-Chi Huang was born in Taiwan. He received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1977 and the M.S. and Ph.D. degrees in electrical engineering from the University of California at Berkeley, in 1980 and 1984, respectively.

From 1984 to 1988, he was a Radio Frequency and Communication System Engineer with the Corporate Research and Development Center, the General Electric Company, Schenectady, NY, where he worked on mobile radio communication system design. From 1989 to 1992, he was a Research Staff Member with the IBM T. J. Watson Research Center, Yorktown Heights, NY, where he worked on indoor radio communication system design. He has been a faculty member since 1992 and is currently a Professor with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu, Taiwan.