Optimal Resource Planning with Interference Coordination for Relay-Based Cellular Networks

Taejoon Kim[1], Kwanghoon An[1], and Heejung Yu[2]
[1] School of Information and Communications Engineering,
Chungbuk National University,
Chungju-si, 28644 – South Korea
[e-mail: ktjcc@chungbuk.ac.kr, wdwf990@chungbuk.ac.kr]
[2]Department of Information and Communications Engineering,
Yeungnam University,
Gyeongsan-si, 38541– South Korea
[e-mail: heejung@yu.ac.kr]
*Corresponding author: Heejung Yu

Received February 23, 2016; revised May 12, 2017; revised July 3, 2017; accepted July 24, 2017; published November 30, 2017

Abstract

Multihop relay-based cellular networks are attracting much interest because of their throughput enhancement, coverage extension, and low infrastructure cost. In these networks, relay stations (RSs) between a base station (BS) and mobile stations (MSs) drastically increase the overall spectral efficiency, with improved channel quality for MSs located at the cell edge or in shadow areas, and enhanced throughput of MSs in hot spots. These relay-based networks require an advanced radio resource management scheme because the optimal amount of radio resource for a BS-to-RS link should be allocated according to the MS channel quality and distribution, considering the interference among RSs and neighbor BSs. In this paper, we propose optimal resource planning algorithms that maximize the overall utility of relay-based networks under a proportional fair scheduling policy. In the first phase, we determine an optimal scheduling policy for distributing BS-to-RS link resources to RSs. In the second phase, we determine the optimal amount of the BS-to-RS link resources using the results of the first phase. The proposed algorithms efficiently calculate the optimal amount of resource without exhaustive searches, and their accuracy is verified by comparison with simulation results, in which the algorithms show a perfect match with simulations.

Keywords: Multihop relay network, radio resource allocation, optimization, proportional fair scheduling

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry Education (NRF-2017R1D1A1B03032911), by the Human Resources Development of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government Ministry of Trade, industry & Energy (20144030200450).

https://doi.org/10.3837/tiis.2017.11.005 **ISSN : 1976-7277**

1. Introduction

THE rapid increase of mobile data traffic in cellular networks calls for low-cost and high-performance infrastructure technologies. A multihop relay network (MRN) provides multihop wireless connectivity by deploying relay stations (RSs) to assist the communication between a base station (BS) and mobile stations (MSs). This technology is a cost-effective method of upgrading the conventional cellular networks because RSs are equipped with fewer functionalities than BSs, and the expensive wireline backhaul links among BSs can be substituted with wireless backhaul links between BSs and RSs [1]-[2]. Moreover, the mobility of RSs, supported by the well-established MRN standard of IEEE 802.16j [3]-[4], enables a traffic demand-based dynamic RS deployment strategy. Therefore, the relaying techique is considered to be the best option for extending the coverage and improving the channel quality of MSs located at cell edges, shadow areas, or hot spots. Accordingly, the heterogeneous network (HetNet) with RSs has emerged as a core component of the fifth generation (5G) mobile communication system [5]-[7]. An examplary MRN architecture, in which MSs can be associated with a BS or an RS is depicted in **Fig. 1**. In general, MSs located near a BS are associated with the BS, and MSs located at the cell edge or in a shadow area are associated with RSs. RSs enhance the quality of end-to-end communication by forwarding the received data packets to MSs.

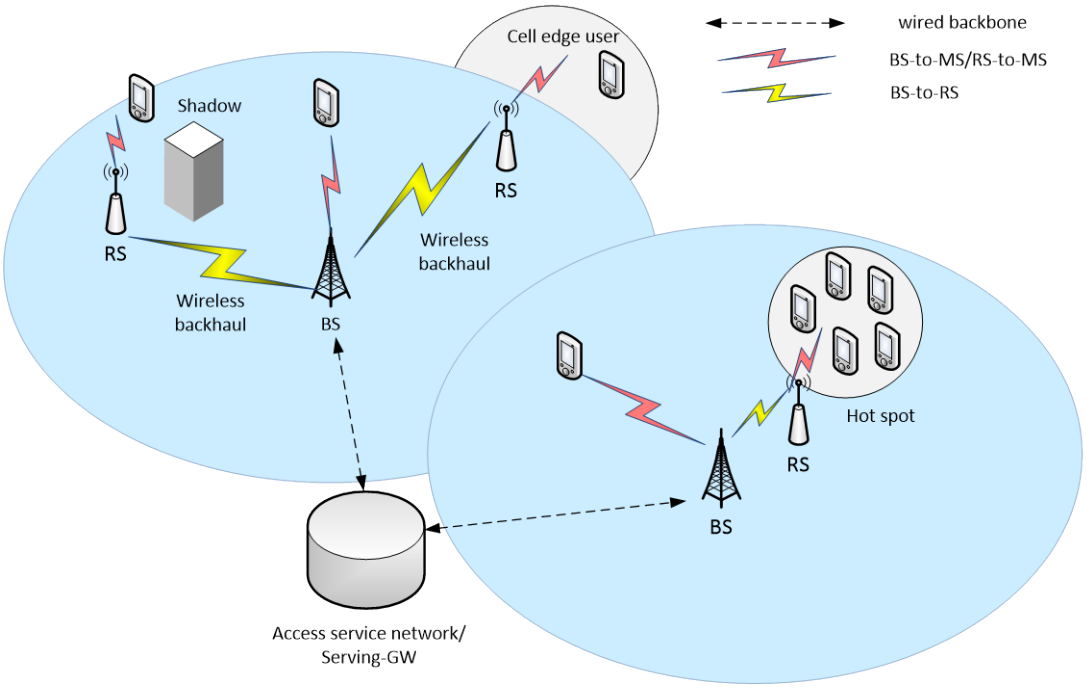

Fig. 1. Examplary MRN architecture

In spite of these advantages, there are some issues that still need to be addressed to operate MRNs efficiently. First, additional radio resources should be allocated to the wireless links

between a BS and RSs, and a quantitative analysis is required to determine the optimal amount of radio resources that should be allocated thusly. Moreover, because multiple RSs can be connected to a BS, an optimal scheduling policy should be determined for distributing the resources to each BS-to-RS link. In general, BSs and RSs transmit packets to the attached MSs in the same frame duration, and the radio resources for this frame duration can be either partitioned or reused. In the resource partition scheme, the BS and RSs of a cell use different frequency bands, and the interference among them is greatly mitigated. In contrast, in the resource resuse scheme, the BS and RSs use the same frequency band. Accordingly, more radio resources can be allocated to MSs in the reuse scheme than in the partition scheme; however, in the reuse scheme, MSs experience interference from BSs and RSs. Therefore, a sophisticated radio resource management scheme is required in MRNs.

1.1 Related Works

In [8], an optimization problem is formulated for maximizing system capacity while guaranteeing minimum resources for each MS, and an efficient heuristic algorithm is proposed as well. In [9], a resource allocation problem is formulated for maximizing cell throughput using binary integer linear programming (BILP), and the Hungarian method [10] is used in the proposed heuristic algorithm. In [11], a resource scheduling algorithm is proposed for extracting multiuser diversity gain using channel information in a relay-based cellular network. In [12], Choi *et al*. modify the optimization problem of proportional fair (PF) [13] scheduling for orthogonal frequency division multiple access (OFDMA) systems [14] to improve cell-edge user throughput while maximizing total system throughput. In [15], a resource allocation algorithm is proposed for MRNs under a PF scheduling policy; the proposed algorithm determines the amount of resources for the BS-to-RS links by assigning resource slots according to the PF scheduling metric. In [16], a scheduling scheme is proposed under a max-min fair constraint; however, PF is not considered. In [15] and [17], the intercell and the intracell interferences are greatly reduced through an orthogonal frequency band allocation scheme. In [18], an optimal resource allocation is considered in relation to relay networks of greater complexity than dual-hop, and a distributed resource allocation algorithm is proposed. In [19], a combined relay selection and link scheduling method is considered, and the problem is decoupled into two subproblems of frame segmentation and relay selection. In [20], Li *et al.* consider the problem of intracell resource allocation in relay networks, accounting for coordinated multipoint (CoMP) transmission to improve the channel quality of MSs. In [21], Park *et al.* propose a resource allocation algorithm that extracts multiuser diversity gain over a relay-enhanced network. In [24], co-channel interference under multi-cell and multi-antenna environments is investigated and the closed-form downlink capacity is derived. In [25], the performance of random cellular networks is evaluated through novel spatial spectrum and energy efficiency models.

1.2 Contributions

In this paper, we focus on an optimal resource planning problem to determine the optimal amount of resources to be allocated for BS-to-RS links. The most of existing research work for MRNs does not cope with the BS-to-RS links resource allocation and the interference mitigation simultaneously. The main contribution of our work is the design of a radio resource management scheme which considers the interference management, the optimal resource allocation, and the utility maximization through PF scheduling altogether. A resource partition scheme is adopted, as well as an orthogonal frequency band allocation scheme for interference coordination. Moreover, we derive the throughput of each MS for MRNs under the PF scheduling and formulate an optimization for problem of maximizing the total utility of MRNs.

We obtain optimal resource management algorithms through mathematical derivation. We verify the accuracy of the proposed algorithms by comparing the results of the algorithms with simulation results, and the results of the algorithms are shown to perfectly match those of the simulations. In additional, the algorithms are shown to be computationally efficient because no iterative search is required.

2. System Model

We consider a dual-hop downlink (DL) MRN with one BS and intermediate RSs. The frame structure, RS deployment, and orthogonal frequency allocation are depicted in **Fig. 2**. The cell area of this network is divided into three sectors with two RSs deployed in each sector, as shown in **Fig. 2**(a), which also illustrates the relationship between the one BS-RS system under investigation and two other neighbor cells. According to the IEEE 802.16j standard, the MRN frame is divided into a DL subframe and an uplink (UL) subframe. In a time domain, a DL subframe is made up of an access zone and a relay zone, as shown in **Fig. 2**(b). In the access zone, the BS and RSs send packets to their subordinate MSs. In contrast, packets to be relayed through the RSs are transmitted from the BS to the RSs through the relay zone. After receiving the packets, the RSs forward them to their subordinate MSs through the access zone of the next frame. The BS and RSs should use the symbol-aligned common zone boundary between the access zone and the relay zone shown as a vertical line with arrows in **Fig. 2**(b). An interference coordination scheme based on orthogonal frequency allocation [15], [17] is used to mitigate intercell and intracell interferences. More specifically, in the case of the access zone frequency allocation, the total frequency band is partitioned into three subbands, and these three subbands are allocated to the three sector segments of the BS, respectively. In addition, the RSs opposing each other across the cell center reuse the same subband as shown in **Fig. 2**(a).

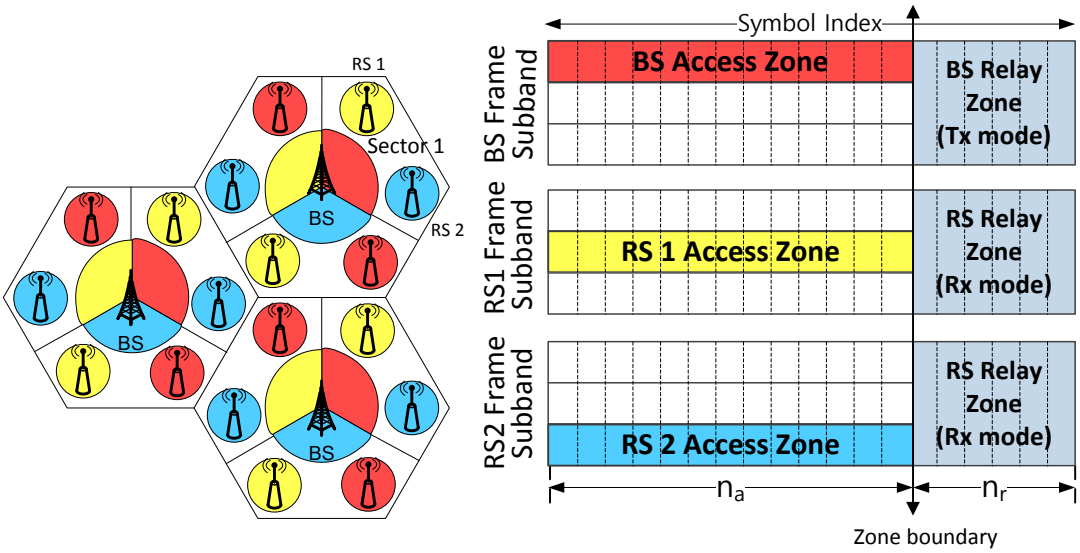

Fig. 2. Interference coordinated subband allocation and DL subframe structure [15].

In the case of the relay zone subband allocation, all of the available subbands are allocated to the BS-to-RS links of the relay zone because the interference among the BS-to-RS links can be effectively managed by locating the RSs at the appropriate positions. We assume that the BS has infinitely backlogged packets to transmit to each MS. However, the RSs may suffer from buffer underruns when the BS-to-RS links become bottlenecked, that is, when the achievable throughput of the BS-to-RS links is lower than that of the RS-to-MS links. In addition, we assume that each MS experiences an independent Rayleigh block-fading channel. Due to mathematical tractability, we adopt the simplified PF scheduling metric [22], which is the current signal-to-noise ratio (SNR) divided by the average SNR, instead of the conventional PF metric [23]. Finally, we do not apply PF scheduling to the BS-to-RS links because the channel fading of these links is very slow.

3. Optimal Resource Planning for MRNs

3.1 Optimization Problem Formulation

Focusing on the DL traffic flows of one sector of a cell, in which two RSs (RS 1 and RS 2) are deployed: let K_b, K_{r1}, and K_{r2} represent the numbers of MSs associated with the BS, RS 1, and RS 2, respectively. We denote M_i^b, M_j^{r1}, and M_k^{r2} as the *i*-th, *j*-th, and *k*-th MS, respectively, associated with the BS, RS 1, and RS 2, with M^b, M^{r1}, and M^{r2} representing the sets of MSs associated with the BS, RS 1, and RS 2, respectively, e.g., $M^b = \{M_1^b, ..., M_{K_b}^b\}$. The scheduling metric for M_i^b is defined as follows:

$$\Gamma_i^b(t) = \frac{\gamma_i^b(t)}{\overline{\gamma}_i^b},\ i = 1,\ ..., K_b, \tag{1}$$

where $\gamma_i^b(t)$ is the SNR of M_i^b at frame *t* and $\overline{\gamma}_i^b$ is the average SNR of M_i^b. The scheduling metrics for M_j^{r1} and M_k^{r2} are defined in the same manner. In the Rayleigh block-fading model, the probability density function (pdf) is given by $f_{\overline{\gamma}}(\gamma) = 1/\overline{\gamma}\exp(-\gamma/\overline{\gamma})$, where γ is the instantaneous SNR, and $\overline{\gamma}$ is the average SNR. Therefore, by substituting $x = \gamma/\overline{\gamma}$, the pdf of the scheduling metric for any MS, regardless of the station with which it is associated, is given by $f_s(x) = \exp(-x)$. In every scheduling epoch, the BS selects the optimal MS with the highest scheduling metric value from M^b. Likewise, RS 1 and RS 2 select the optimal MS from M^{r1} and M^{r2}, respectively. Accordingly, when M_i^b with the highest metric value x is selected, its cumulative distribution function (cdf) is given by

$$P\{\max(\Gamma_1^b, ..., \Gamma_{K_b}^b) = \Gamma_i^b \le x\} = \int_0^x (1-e^{-s})^{K_b-1} e^{-s} ds. \tag{2}$$

We can obtain the pdf of x for M_i^b by differentiating (2), which is given by

$$f_b(x) = (1-e^{-x})^{K_b-1} e^{-x},\ x \ge 0. \tag{3}$$

Likewise, the pdf of x for M_j^{r1}, M_k^{r2} can be obtained as $f_{r1}(x)=(1-e^{-x})^{K_{r1}-1}e^{-x}, x\geq 0$ and $f_{r2}(x)=(1-e^{-x})^{K_{r2}-1}e^{-x}, x\geq 0$, respectively. From the Shannon's capacity theorem, the spectral efficiency, i.e., the achievable rate per hertz, for the selected M_i^b is given by

$$C_i^b=\int_0^\infty \log_2(1+\overline{\gamma}_i^b x)\, f_b(x)dx. \tag{4}$$

Therefore, the total access zone spectral efficiency of M^b is $C^B=\sum_{i=1}^{K_b} C_i^b$. Likewise, C^{R1} and C^{R2} are the total access zone spectral efficiencies of M^{r1} and M^{r2}, respectively, which can be obtained in the same manner. The spectral efficiencies for the BS-to-RS 1 and BS-to-RS 2 links are given by $C^{br1}=\log_2(1+\gamma_{R1})$ and $C^{br2}=\log_2(1+\gamma_{R2})$, respectively, where γ_{R1} and γ_{R2} are the SNRs of the BS-to-RS 1 and BS-to-RS 2 links. We denote n_a, n_r, and $n_t=(n_a+n_r)$ as the number of OFDM symbols for the access zone, relay zone, and DL subframe, respectively. As shown in **Fig. 2**(b), the x-axis of the DL subframe is the OFDM symbol index. In the case of the relay zone, as shown in **Fig. 2**(b), all three subbands are allocated to this zone, and the BS distributes this relay zone resource to RS 1 and RS 2. We denote $\alpha, (0\leq\alpha\leq 1)$, as the *scheduling ratio* of RS 1's relay zone resource to the total relay zone resource. The BS scheduler builds DL subframes according to the spectral efficiencies and the amount of radio resources allocated to each data burst, following the procedure shown in the block diagram depicted in **Fig. 3**.

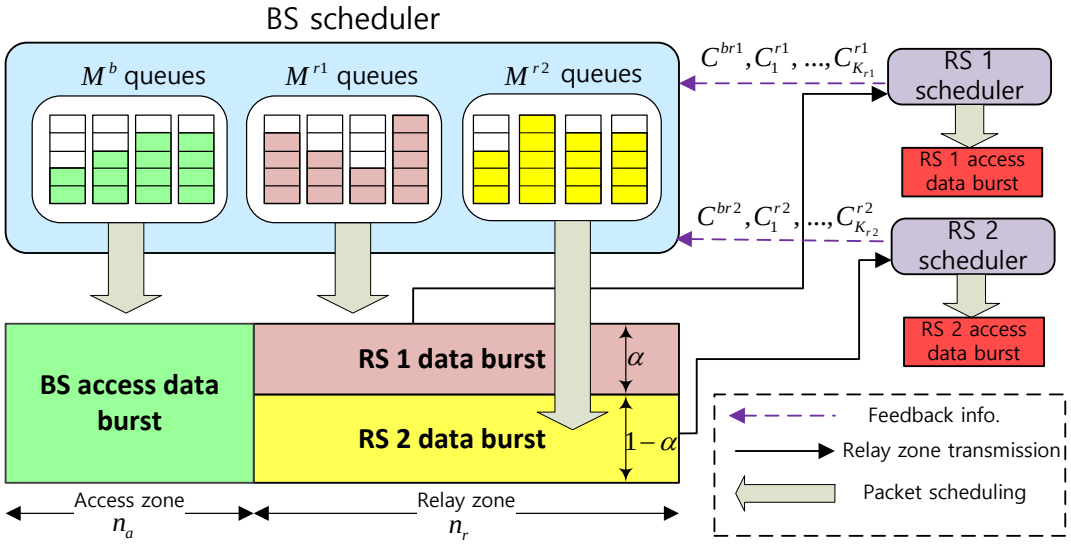

Fig. 3. Block diagram of BS scheduler procedure

Let g be the number of subcarriers per OFDM symbol for a single subband; then the throughput of M_i^b through the access zone of the BS is expressed as:

$$T_i^b = \frac{n_a g C_i^b}{T_f}, i = 1, ..., K_b, \tag{5}$$

where T_f is the frame duration in seconds. As shown in **Fig. 3**, the amounts of data that can be loaded on the data bursts of RS 1 and RS 2 are determined by C^{br1}, C^{br2}, and the amount of resource assigned. Hence, the throughputs of the BS-to-RS links are given as follows:

$$T^{br1} = \frac{3 n_r g \alpha C^{br1}}{T_f}, \quad \text{BS-to-RS 1,} \tag{6}$$

$$T^{br2} = \frac{3 n_r g (1-\alpha) C^{br2}}{T_f}, \text{ BS-to-RS 2,} \tag{7}$$

where we should note that all the three subbands are used in the relay zone. Moreover, the BS scheduler should determine how much data from each MS's queue should be loaded on the data bursts in each scheduling epoch. The RSs adopt the PF scheduling policy, and this policy guarantees equal resource allocation to each subordinate MS, regardless of its average SNR. Therefore, the amounts of data for M_j^{r1} and M_k^{r2} that are loaded on each data burst should be proportional to their spectral efficiencies C_j^{r1} and C_k^{r2}, respectively; their throughputs in BS-to-RS links are accordingly expressed as $T^{br1} \cdot C_j^{r1} / C^{R1}$ and $T^{br2} \cdot C_k^{r2} / C^{R2}$, respectively. The maximum achievable throughputs of M_j^{r1} and M_k^{r2} over the RS-to-MS links are $n_a g C_j^{r1} / T_f$ and $n_a g C_k^{r2} / T_f$, respectively. The end-to-end throughputs of BS-to-MS via RSs are the minimum throughputs of BS-to-RS and RS-to-MS links, and they are given as follows:

$$\begin{aligned} T_j^{r1} &= \min\left(T^{br1} \cdot \frac{C_j^{r1}}{C^{R1}}, \frac{n_a g C_j^{r1}}{T_f}\right) \\ &= \frac{g \cdot C_j^{r1}}{T_f} \min\left(3 n_r \alpha \frac{C^{br1}}{C^{R1}}, n_a\right), \ j = 1, ..., K_{r1}, \end{aligned} \tag{8}$$

$$\begin{aligned} T_k^{r2} &= \min\left(T^{br2} \cdot \frac{C_k^{r2}}{C^{R2}}, \frac{n_a g C_k^{r2}}{T_f}\right) \\ &= \frac{g \cdot C_k^{r2}}{T_f} \min\left(3 n_r (1-\alpha) \frac{C^{br2}}{C^{R2}}, n_a\right), \ k = 1, ..., K_{r2}. \end{aligned} \tag{9}$$

Considering all the MSs in the network, the total log utility is calculated as:

$$U_T(\alpha, n_a) = \sum_{i=1}^{K_b} \log(T_i^b) + \sum_{j=1}^{K_{r1}} \log(T_j^{r1}) + \sum_{k=1}^{K_{r2}} \log(T_k^{r2}). \tag{10}$$

Therefore, the optimal resource management problem can be formulated as the following optimization problem:

$$\begin{aligned} &\max_{n_a,\alpha} \; U_T(\alpha, n_a) \\ &\text{subject to} \quad n_a + n_r = n_t, \\ &\qquad\qquad 0 \le \alpha \le 1. \end{aligned} \tag{11}$$

3.2 Optimal Inter-Relay Resource Scheduling Policy

When a certain amount of resource, given by $3n_r g$ where $n_r = n_t - n_a$, is allocated to the relay zone, we find the optimal $\alpha^*(n_a)$ that maximizes U_T. We define an RS as a bottleneck RS if the BS-to-RS link's throughput is lower than the maximum achievable throughput of the RS-to-MS link, i.e., $T^{br1} \le n_a g C^{R1} / T_f$ or $T^{br2} \le n_a g C^{R2} / T_f$. If we increase the scheduling ratio α from 0 to 1, the RS 1 starts as a bottleneck and becomes a non-bottleneck when α is greater than a threshold α_1. Likewise, the RS 2 starts as a non-bottleneck and becomes a bottleneck when α is greater than a threshold α_2. These thresholds can be obtained from (8) and (9) as follows:

$$\alpha_1 = \frac{n_a C^{R1}}{3n_r C^{br1}}, \quad \alpha_2 = 1 - \frac{n_a C^{R2}}{3n_r C^{br2}}. \tag{12}$$

Note that as long as the BS scheduler builds the RSs' data bursts according to the PF scheduling policy as described above, all the MSs belonging to M^{r1} and M^{r2} have the same α_1 and α_2, respectively. Sample graphs of throughput vs α with varying n_a are depicted in **Fig. 4**. If $\alpha_1 \le \alpha_2$, an arbitrary $\forall \alpha \in [\alpha_1, \alpha_2]$ achieves the maximum throughputs $n_a g C_j^{r1} / T_f$ and $n_a g C_k^{r2} / T_f$ for RS 1 and RS 2, as shown in **Fig. 4**(a). Otherwise, when $\alpha_1 \ge \alpha_2$, at least one of the two RSs becomes a bottleneck, as shown in **Fig. 4**(c). When n_a is a high value the both α_1 and α_2 lie out of [0,1], the both RS 1 and RS 2 become bottleneck as shown in **Fig. 4**(d) .

We denote U_R as the sum of utilities for M^{r1} and M^{r2}. If both RS 1 and RS 2 become bottlenecks, U_R is calculated as:

$$\begin{aligned} U_R &= \sum_{j=1}^{K_{r1}} \log(T_j^{r1}) + \sum_{k=1}^{K_{r2}} \log(T_k^{r2}) \\ &= \log\Big[\prod_{j=1}^{K_{r1}} \frac{3n_r g C^{br1} C_j^{r1}}{C^{R1} \cdot T_f} \prod_{k=1}^{K_{r2}} \frac{3n_r g C^{br2} C_j^{r2}}{C^{R2} \cdot T_f}\Big] + \log(\alpha^{K_{r1}} (1-\alpha)^{K_{r2}}), \end{aligned} \tag{13}$$

where only the second term is a function of α. The differentiation of the function $h_1(\alpha) = \alpha^{K_{r1}}(1-\alpha)^{K_{r2}}$ shows that it achieves its maximum at $\alpha_0 = K_{r1} / (K_{r1} + K_{r2})$. Therefore, if $\alpha_0 \in [\alpha_2, \alpha_1]$, U_R achieves its maximum at $\alpha = \alpha_0$. Otherwise, when $\alpha_0 > \alpha_1$,

increasing α more than α_1 is a waste of resource because T_j^{r1} is bounded by $n_a g C_j^{r1} / T_f$. Therefore, in this case, U_R achieves its maximum at $\alpha = \alpha_1$. Likewise, if $\alpha_0 < \alpha_2$, U_R achieves its maximum at $\alpha = \alpha_2$. From these relationship, the optimal ratio $\alpha^*(n_a)$ is described as:

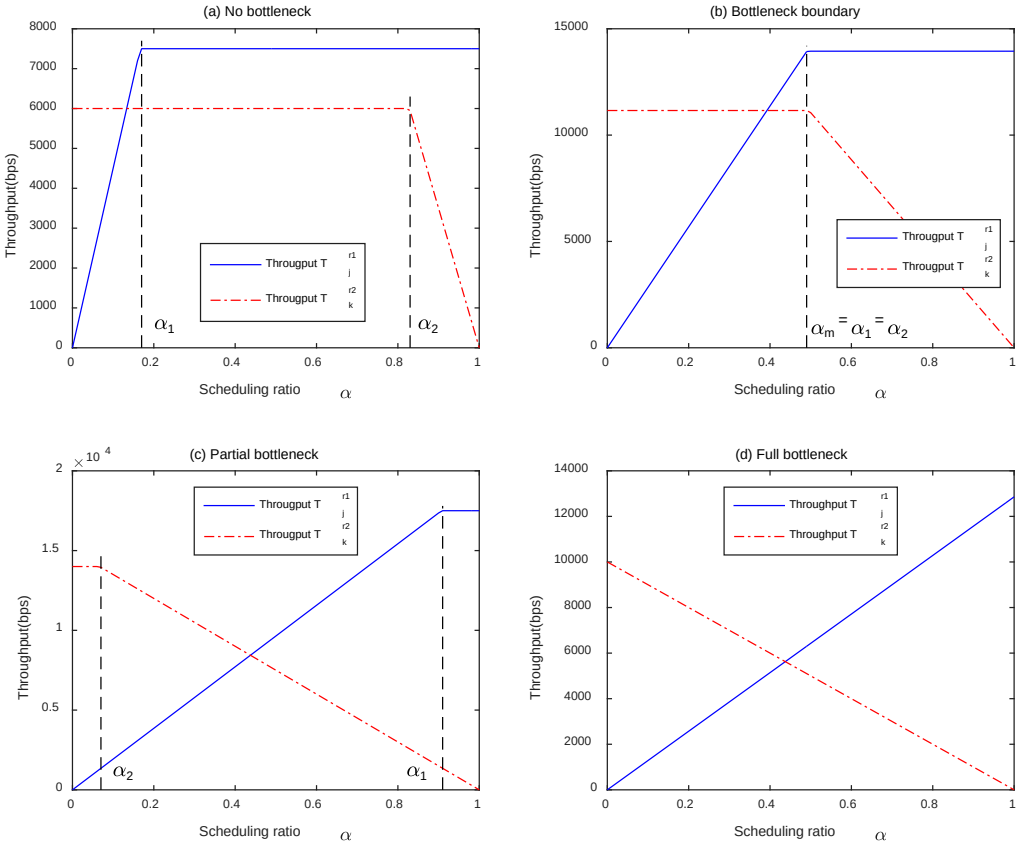

Fig. 4. RS bottleneck situations with increasing n_a

$$\alpha^*(n_a) = \max(\alpha_2, \min(\alpha_0, \alpha_1)). \tag{14}$$

Even when $\alpha_1 > 1$ or $\alpha_2 < 0$, (14) is a valid expression because $0 < \alpha_0 < 1$ is satisfied. Moreover, when $\alpha_1 \le \alpha_2$, (14) encompasses the abovementioned condition of $\alpha^*(n_a) \in [\alpha_1, \alpha_2]$. Therefore, given n_a, or n_r, (14) represents the optimal relay zone scheduling policy for distributing the relay zone resource to RS 1 and RS 2.

3.3 Optimal Resource Allocation

The optimal amount of resource for the access and relay zones is determined by adopting the PF scheduling policy for the access zone and the optimal scheduling policy given by (14) for the relay zone. At this time, we treat n_a as a continuous variable with $n_a \in [0, n_t]$, and the

range $[0, n_t]$ can be partitioned into three regions: a *no bottleneck region* with $[0, B_1]$, a *partial bottleneck region* with $(B_1, B_2]$, and a *full bottleneck region* with $(B_2, n_t]$. When n_a is in the *no bottleneck region*, neither RS 1 nor RS 2 is a bottleneck. However, when n_a lies in the *partial bottleneck region*, only one RS becomes a bottleneck. In the *full bottleneck region*, both RSs become bottlenecks. In the *no bottleneck region*, U_T is expressed as follows:

$$\begin{aligned} U_T &= \sum_{i=1}^{K_b} \log(\frac{n_a g C_i^b}{T_f}) + \sum_{j=1}^{K_{r1}} \log(\frac{n_a g C_j^{r1}}{T_f}) + \sum_{k=1}^{K_{r2}} \log(\frac{n_a g C_k^{r2}}{T_f}) \\ &= \log(n_a^{K_b+K_{r1}+K_{r2}}) + \log[\prod_i \frac{g C_i^b}{T_f} \prod_j \frac{g C_j^{r1}}{T_f} \prod_k \frac{g C_k^{r2}}{T_f}], \end{aligned} \tag{15}$$

which is maximized at $n_a = B_1$ because U_T monotonically increases with n_a. When $n_a = B_1$, the condition $\alpha_1 = \alpha_2$ is satisfied as shown in **Fig. 4**(b). According to (12) and the given condition $\alpha_1 = \alpha_2$, we can obtain B_1 as follows:

$$\begin{aligned} \alpha_1 &= \frac{B_1 C^{R1}}{3(n_t - B_1) C^{br1}} = \alpha_2 = 1 - \frac{B_1 C^{R2}}{3(n_t - B_1) C^{br2}} \\ B_1 &= \frac{3n_t}{\frac{C^{R1}}{C^{br1}} + \frac{C^{R2}}{C^{br2}} + 3} \end{aligned} . \tag{16}$$

In the *partial bottleneck region*, it should be determined which RS becomes a bottleneck. Under an *RS 1-dominant* condition, it is not RS 1 but RS 2 that becomes a bottleneck, and vice versa. When $n_a = B_1$, we denote α_m as $\alpha_m = \alpha_1 = \alpha_2 = \frac{B_1 C^{R1}}{3(n_t - B_1) C^{br1}}$. As indicated by (14), if $\alpha_0 \geq \alpha_m$, an *RS 1-dominant* condition will prevail. Otherwise, when $\alpha_0 < \alpha_m$, an *RS 2-dominant* condition will prevail. These conditions can be expressed as follows:

$$\begin{aligned} &\alpha_0 \cdot (\frac{C^{R1}}{C^{br1}} + \frac{C^{R2}}{C^{br2}}) \geq \frac{C^{R1}}{C^{br1}}, \quad RS\ 1-dominant \\ &\alpha_0 \cdot (\frac{C^{R1}}{C^{br1}} + \frac{C^{R2}}{C^{br2}}) < \frac{C^{R1}}{C^{br1}}, \quad RS\ 2-dominant \end{aligned} \tag{17}$$

Under the *RS 1-dominant* condition, $\alpha_0 \geq \alpha_1$ is satisfied. Similarly, under the *RS 2-dominant* condition, $\alpha_0 \leq \alpha_2$ is satisfied. From these inequalities, B_2 is obtained as follows:

$$B_2 = \begin{cases} n_t\left(1+\dfrac{C^{R1}}{3\alpha_0 C^{br1}}\right)^{-1}, & RS1-dominant \\ n_t\left(1+\dfrac{C^{R2}}{3(1-\alpha_0) C^{br2}}\right)^{-1}, & RS2-dominant \end{cases} \tag{18}$$

Under the *RS 1-dominant* condition, U_T is given by

$$\begin{aligned} U_T &= \log\left[n_a^{K_b+K_{r1}}\left(n_t-(1+\frac{C^{R1}}{3C^{br1}})n_a\right)^{K_{r2}}\right] \\ &\quad +\log\left[\left(\frac{3C^{br2}}{C^{R2}}\right)^{K_{r2}}\prod_i \frac{gC_i^b}{T_f}\prod_j \frac{gC_j^{r1}}{T_f}\prod_k \frac{gC_k^{r2}}{T_f}\right], \end{aligned} \tag{19}$$

where only the first log[·] term is a function of n_a. The bracketed portion of this term is denoted as $h_2(\cdot)$, which is given by

$$h_2(x) = x^A(n_t - Dx)^B,\ B_1 < x \le B_2, \tag{20}$$

where $A = K_b + K_{r1}$, $B = K_{r2}$, and $D = 1 + C^{R1}/(3C^{br1})$. It should be noted that $h_2(n_t/D) = 0$, and the differentiation of $h_2(x)$ shows that $h_2(x)$ may achieve its maximum at $x = \dfrac{A}{A+B}\dfrac{n_t}{D}$ or $x = n_t$. However, $B_2 < n_t/D$ is always satisfied as shown by

$$B_2 = \frac{n_t}{1+\dfrac{C^{R1}}{3\alpha_0 C^{br1}}} < \frac{n_t}{1+\dfrac{C^{R1}}{3C^{br1}}} = \frac{n_t}{D}. \tag{21}$$

Therefore, $h_2(x)$ maintains the unimodal property over the *partial bottleneck region*. From these relationships, under the *RS 1-dominant* condition, the candidate maximizer for U_T is $x = \dfrac{A}{A+B}\dfrac{n_t}{D}$, and the same logic can be applied to the *RS 2-dominant* condition. Considering both conditions, the candidate maximizer n_a^1 for U_T over the *partial bottleneck region* is obtained as follows:

$$n_a^1 = \begin{cases} \dfrac{K_b+K_{r1}}{K_b+K_{r1}+K_{r2}}\dfrac{n_t}{1+\dfrac{C^{R1}}{3C^{br1}}}, & RS1-dominant \\ \dfrac{K_b+K_{r2}}{K_b+K_{r1}+K_{r2}}\dfrac{n_t}{1+\dfrac{C^{R2}}{3C^{br2}}}, & RS2-dominant \end{cases} \tag{22}$$

In the *full bottleneck region*, both RS 1 and RS 2 become a bottleneck, and $\alpha^*(n_a)=\alpha_0$ is satisfied. Therefore, U_T is given by

$$U_T=\log\left[n_a^{K_b}n_r^{K_{r1}+K_{r2}}\right]+\log\left[\prod_i\frac{gC_i^b}{T_f}\prod_j\frac{3g\alpha_0C_j^{r1}}{C^{R1}\cdot T_f}\prod_k\frac{3g(1-\alpha_0)C_k^{r2}}{C^{R2}\cdot T_f}\right], \tag{23}$$

where U_T exhibits a unimodal property and can be maximized at n_a^2, which is described as follows:

$$n_a^2=\frac{K_bn_t}{K_b+K_{r1}+K_{r2}}. \tag{24}$$

Considering the range of each region, we denote the maximizers for U_T over the *partial bottleneck region* and the *full bottleneck region* as n_1 and n_2, respectively, which can be obtained as follows:

$$\begin{aligned}n_1&=\max(B_1,\min(n_a^1,B_2))\\ n_2&=\max(B_2,n_a^2)\end{aligned} \tag{25}$$

Finally, we find the maximizer over $[0,n_t]$, i.e., n^*, by comparing n_1 and n_2, which is given by

$$n^*=\left\langle\arg\max_{n_1,n_2}(U_T(\alpha^*(n_1),n_1),U_T(\alpha^*(n_2),n_2))\right\rangle, \tag{26}$$

where $\langle\cdot\rangle$ is a rounding operator.

4. Simulation and Analysis Results

In this section, we compare the proposed analytic resource management scheme provided by (14) and (26) with simulation-based exhaustive search schemes. The comparison results demonstrate that the proposed scheme is very accurate at allocating the optimal resource amounts, and very efficient, in that it requires fairly little computational complexity without iterative computation. In the analyses and simulations, we adopt the system parameters specified by IEEE 802.16j, which are summarized in **Table 1**.

Table 1. System parameters for simulation and analysis

Parameter	Value
Channel bandwidth	10 MHz
FFT size	1024
Number of symbols	48
Frame duration	5 ms
Number of MSs associated with BS	25 ~ 40
Number of MSs associated with RS	5 ~ 20
Number of subcarriers per a symbol	720
SNR of BS-to-RS link	15 ~ 25 dB
SNR of BS-to-MS links	5 ~ 20 dB
SNR of RS-to-MS links	10 ~ 25 dB

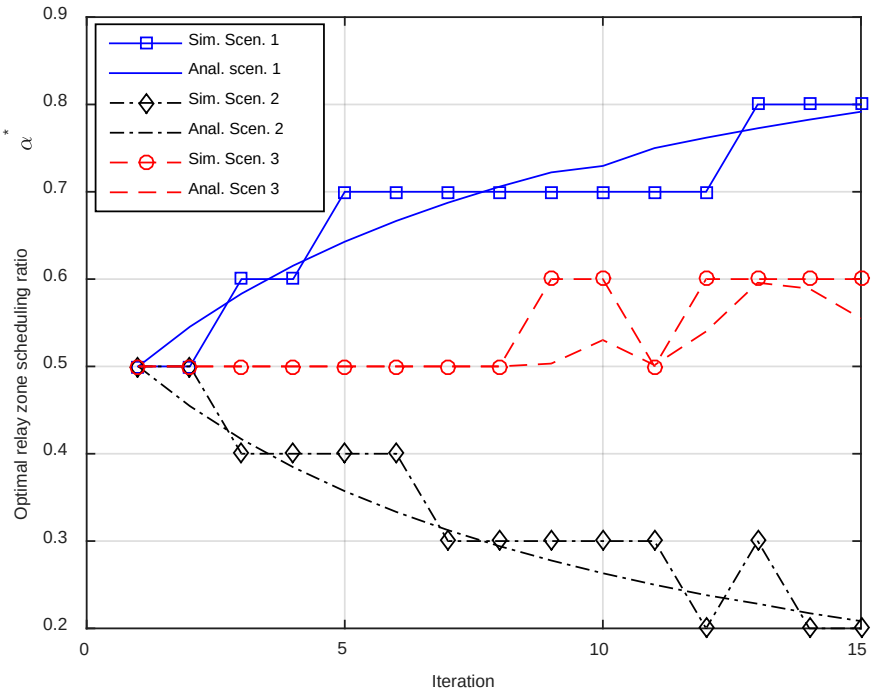

Fig. 5. The optimal scheduling ratio α^* with varying parameters

Subsequently, we obtain the optimal amount of resource for the access zone and the relay zone through the proposed optimal resource allocation method given by (26), and compare the results with those obtained through simulation-based exhaustive search. The computational burden for these simulations is severe, because the (n_a, α) pair should be optimized concurrently. Because the total number of 48 symbols per frame is assumed, the exhaustive search should test all the n_a values ranging from 0 to 48 in each iteration. Moreover, in each n_a test, the optimal α should also be searched through the aforementioned exhaustive search scheme. However, the proposed scheme efficiently finds the optimal amount of resource simply by comparing only two candidate values without iterative searches. In the first scenario, we increase K_b from 25 to 39 in each iteration, fixing other parameters at $K_{r1} = 5$, $K_{r2} = 5$, and $\overline{\gamma}^{br1} = \overline{\gamma}^{br2} = 15$ dB. In addition, $\overline{\gamma}_i^b$, $\overline{\gamma}_j^{r1}$, and $\overline{\gamma}_k^{r2}$ are assigned randomly to the same ranges as above. The second and third scenarios are similar to the second and third scenarios described above in searching for the optimal α, except that n_a is no longer fixed. The simulation time per iteration is 100 s, which amounts to $100 \times 200 = 20000$ frames.

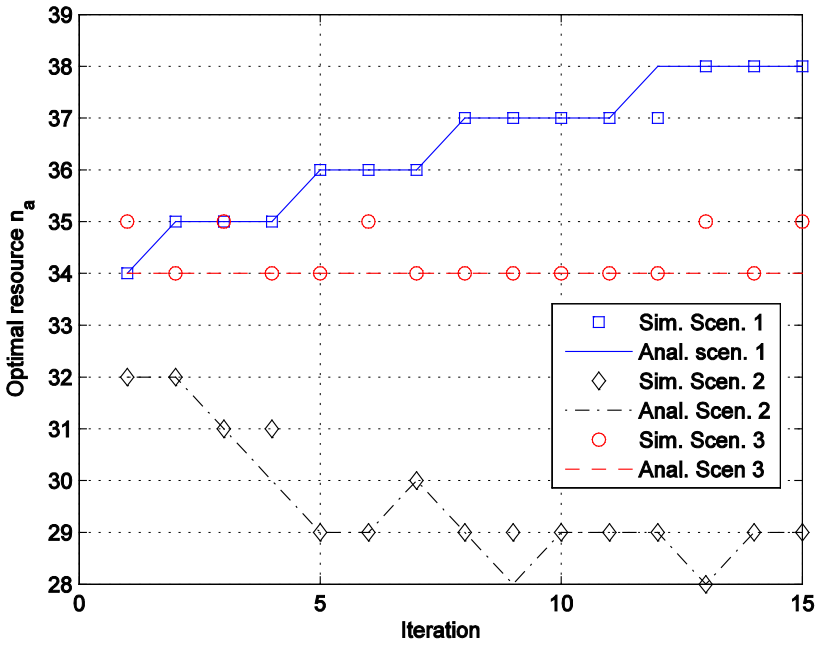

Fig. 6. The optimal resource n_a with varying parameters

Fig. 6 presents the results of both the simulations and the analyses, showing that they closely match. In the first scenario, the optimal access zone resource n^* increases gradually as K_b increases, because the BS requires a relatively large amount of resource to accommodate many MSs. In contrast, in the second scenario, n^* decreases as K_{r1} increases, but soon becomes bounded and does not decrease further. This result can be explained because although the BS scheduler allocates more resource to the BS-to-RS link as K_{r1} increases, if too much resource is allocated to that link, the RSs suffer from bandwidth shortage in their access zones, hindering their ability to serve the increased number of MSs. In the third scenario, n^* rarely changes as $\overline{\gamma}^{br2}$ increases, indicating that the average SNRs of the BS-to-RS links have a limited effect on n^*, compared with the number of MSs associated with the BS and RSs, because the PF scheduling guarantees equal resource allocation per MS. Note that even though n^* rarely changes in this scenario, as shown in **Fig. 5**, α^* gradually increases to allocate more resource to the BS-to-RS 1 link, which has comparatively poorer channel quality than the BS-to-RS 2 link. Although 20000 frames per iteration is adopted, some glitches in the simulation results occurred. On the other hand, the analysis results require fairly low computational complexity and achieve very high accuracy.

Finally, we consider sector throughput and total user utility, changing n_a from 1 to 47, and, at each n_a, the throughput and utility are obtained through simulation. In the first scenario, the MS distribution $K_b = 20$, $K_{r1} = K_{r2} = 15$, and $\overline{\gamma}^{br1} = \overline{\gamma}^{br2} = 5$ dB is assumed. In the second scenario, the MS distribution is $K_b = 25$, $K_{r1} = K_{r2} = 15$, and $\overline{\gamma}^{br1} = \overline{\gamma}^{br2} = 15$ dB. In the third scenario, the MS distribution is $K_b = 40$, $K_{r1} = K_{r2} = 5$, and $\overline{\gamma}^{br1} = \overline{\gamma}^{br2} = 25$ dB. The

throughputs and utilities for these scenarios are depicted in **Figs. 7 - 8**, respectively, and the optimal n_a s obtained from the proposed schemes are marked as with pentagons.

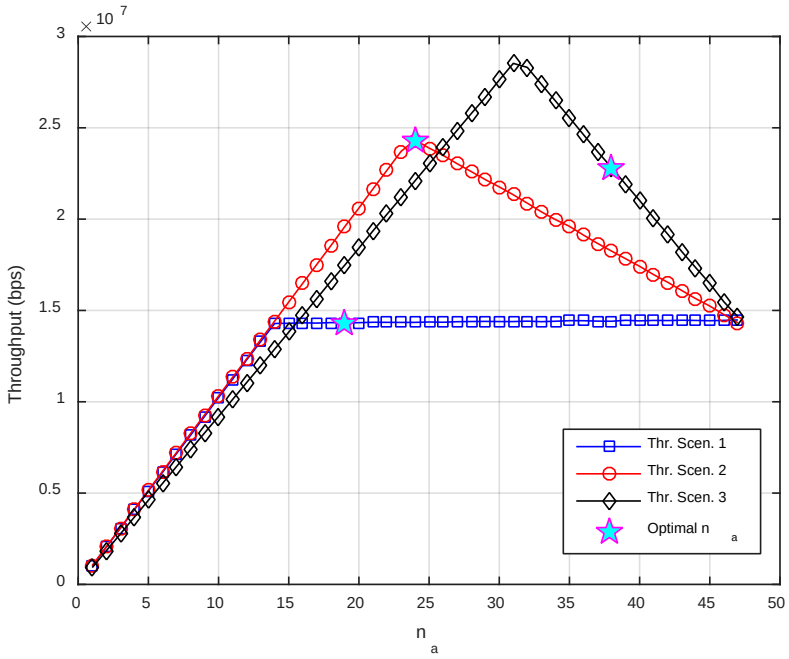

Fig. 7. Sector throughput with different MS distributions and varying n_a

In the first scenario, as shown in **Fig. 7,** sector throughput increases until $n_a = 14$; then, it is saturated. Because the SNRs of the BS-to-RS 1 and 2 links are poor for the first scenario, the amount of traffic from RS 1 and 2 to their attahced MSs is limited, even with high n_a s. In the second scenario, the throughput increases until $n_a = 24$; then, it decreases. When n_a increases past 24, the amount of traffic from RS 1 and 2 to their MSs is decreased, because the amount of resource allocated to BS-to-RS 1 and 2 links decreased. A similar thoughput pattern is observed in the third scenario. In this case, the n_a achieving the maximum throughput is higher than the other scenarios, because the SNRs of BS-to-RS links for this scenario is good. As shown in this figure, the SNRs of BS-to-RS links plays an important role in maximizing throughput. Because the optimal n_a maximizes the log utility function, guaranteeing strict fairness among MSs, it does not match the throughput maximizing n_a. In the third scenario, the optimal n_a should be higher than 32 (the value maximizing the throughput), because many MSs are attached to the BS, requiring a lot of access zone resource.

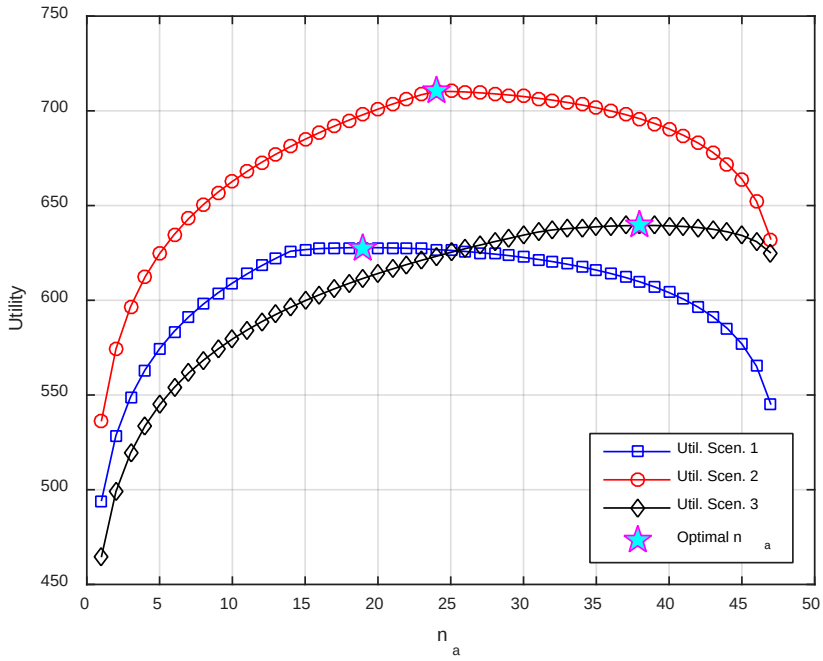

Fig. 8. Utility with different MS distribution and varying n_a

In **Fig. 8**, the utilities for the above-mentioned scenarios are depicted. When n_a is increased, the utilities increase because the throughput of the MSs increases. However, if n_a is further increased, the utilities decrease because the throughput of the MSs attached to the RS 1 and 2 decrease sharply due to the lack of relay zone resource. Accordingly, for all three scenarios, the utility graphs are bell-shaped, and the optimal n_a s achieve the maximum utilities.

5. Conclusion

The main contribution of our work is the development of optimal resource management algorithms for MRNs under a PF scheduling and interference coordination schemes. The proposed algorithms enable MRNs to perform optimal PF scheduling through utility maximization. The proposed schemes determine the optimal scheduling ratio for distributing the relay zone resource to RSs and calculate the optimal amount of resource to be allocated to the access zone and relay zone. Moreover, the proposed schemes are very accurate at determining the optimal parameters, as verified through comparisons with simulations. In addition, the proposed schemes are very efficient, requiring low computational complexity without any iterative search.

References

[1] Ö. Oyman, J. N. Laneman and S. Sandhu, “Multihop relaying forbroadband wireless mesh networks: from theory to practice," *IEEE Commun. Mag*., vol. 45, no. 11, pp. 116-122, November, 2007. Article (CrossRef Link)

[2] R. Pabst, et al., "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Commun. Mag*., vol. 42, no. 9, pp. 80-89, September, 2004. Article (CrossRef Link)
[3] IEEE Std. 802.16j-2009, *IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Broadband Wireless Access Systems Amendment 1: Multiple Relay Specification*, June, 2009. Article (CrossRef Link)
[4] V. Genc, S. Murphy, Y. Yu and J. Murphy, "IEEE 802.16j relay-based wireless access networks: an overview," *IEEE Trans. Wireless Commun*., vol. 15, no. 5, pp. 56-63, October, 2008. Article (CrossRef Link)
[5] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: challenges and research advances," *IEEE Networks*, vol. 28, no. 6, pp. 6-11, Nov. 2014. Article (CrossRef Link)
[6] X. Ge, S. Tu, G. Mao, C.-X. Wang, and T. Han, "5G ultra-dense cellular networks," *IEEE Wireless Commun.*, vol. 23, no. 1, pp. 72-79, March, 2016. Article (CrossRef Link)
[7] X. Ge, J. Chen, C.-X. Wang, J. Thompson, and J. Zhang, "5G green cellular networks considering power allocation schemes," *Science China Information Sciences*, vol. 59, no. 2, pp. 1-14, February, 2015. Article (CrossRef Link)
[8] C. Bae and D.-H. Cho, "Fairness-aware adaptive resource allocation scheme in multihop OFDMA systems," *IEEE Commun. Lett*., vol. 11, no. 2, pp. 134-136, February, 2007. Article (CrossRef Link)
[9] M. Salem, et al., "Fairness-aware radio resource management in downlink OFDMA cellular relay networks," *IEEE Trans. Wireless Commun*., vol. 9, no. 5, pp. 1628-1639, May, 2010. Article (CrossRef Link)
[10] C. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, New Jersey, 1982.
[11] Ö. Oyman, "Opportunistic scheduling and spectrum reuse in relay-based cellular networks," *IEEE Trans. Wireless Commun*., vol. 9, no. 3, pp. 1074-1085, March, 2010. Article (CrossRef Link)
[12] B.-G. Choi, I. Doh and M. Y. Chung, "Radio resource management scheme for relieving interference to MUEs in relay-based cellular networks," *IEEE Trans. Veh. Technol*., vol. 64, no. 7, pp. 3018-3029, July, 2015.
[13] F. P. Kelly, A. K. Maulloo and D. K. H. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *J. Oper. Res. Soc*., vol. 49, no. 3, pp. 237-252, March, 1998. Article (CrossRef Link)
[14] H. Kim and Y. Han, "A proportional fair scheduling for multicarrier transmission systems," *IEEE Commun. Lett*., vol. 9, no. 3, pp. 210-212, March, 2005. Article (CrossRef Link)
[15] Y. Zhao, X. Fang, R. Huang and Y. Fang, "Joint interference coordination and load balancing for OFDMA multihop cellular networks," *IEEE Trans. Mob. Comput*., vol. 13, no. 1, pp. 89-101, January, 2014. Article (CrossRef Link)
[16] Y. Kim and M. L. Sichitiu, "Optimal max-min fair resource allocation in multihop relay-enhanced WiMAX networks," *IEEE Trans. Veh. Technol*., vol. 60, no. 8, pp. 3907-3918, October, 2011. Article (CrossRef Link)
[17] M. Liang, F. Liu, Z. Chen, Y. F. Wang and D. C. Yang, "A novel frequency reuse scheme for OFDMA based relay enhanced cellular networks," in *Proc. IEEE 69th Vehicular Technology Conf*., Barcelona, pp. 1-5, April, 2009. Article (CrossRef Link)
[18] A. Behnad, X. Gao and X. Wang, "Distributed resource allocation for multihop decode-and-forward relay systems," *IEEE Trans. Veh. Technol*., vol. 64, no. 10, pp. 4821-4826, October, 2015. Article (CrossRef Link)
[19] Z. Yang, Q. Zhang and Z. Niu, "Throughput improvement by joint relay selection and link scheduling in relay-assisted cellular networks," *IEEE Trans. Veh. Technol*., vol. 61, no. 6, pp. 2824-2835, July, 2012. Article (CrossRef Link)
[20] Q. Li, R. Q. Hu, Y. Qian and G. Wu, "Intracell cooperation and resource allocation in a heterogeneous network with relays," *IEEE Trans. Veh. Technol*., vol. 62, no. 4, pp. 1770-1784, May, 2013. Article (CrossRef Link)
[21] C.-W. Park, H.-J. Lee and J.-T. Lim, "Fair semi-distributed resource allocation scheme over relay-enhanced OFDMA networks," *IEEE Commun. Lett*., vol. 16, no. 8, pp. 1188-1191, August,

2012. Article (CrossRef Link)
[22] J.-G. Choi and S. Bahk, "Cell-throughput analysis of the proportional fair scheduler in the single-cell environment," *IEEE Trans. Veh. Technol*., vol. 56, no. 2, pp. 766-778, March, 2007. Article (CrossRef Link)
[23] A. Jalali, R. Padovani and R. Pankaj, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proc. of IEEE Veh. Technol. Conf*., Tokyo, pp. 1854-1858, May, 2000. Article (CrossRef Link)
[24] X. Ge, K. Huang, C.-X. Wang, X. Hong, and X. Yang, "Capacity analysis of a multi-cell multi-antenna cooperative cellular network with co-channel interference," *IEEE Trans. Wireless Commun*., vol. 10, no. 10, pp. 3298-3309, October, 2011. Article (CrossRef Link)
[25] X. Ge, B. Yang, J. Ye, G. Mao, C.-X. Wang, and T. Han, "Spatial spectrum and energy efficiency of random cellular networks," *IEEE Trans. Commun*., vol. 63, no. 3, pp. 1019-1030, January, 2015. Article (CrossRef Link)

Taejoon Kim received his BS in Electronics Engineering from Yonsei University, Seoul, Republic of Korea, in 2003, and his PhD in Electrical Engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea, in 2011. He is currently an assistant professor with the School of Information and Communication Engineering, Chungbuk National University, Republic of Korea. From 2003 to 2005, he was a researcher with LG Electronics, Seoul, Republic of Korea. From 2011 to 2013, he was a senior researcher with ETRI, Daejeon, Republic of Korea. His research areas include the analysis and optimization of wireless networks and communication theory.

Kwanghoon An received his BS in Information and Communication Engineering from Chungbuk National University, Chungbuk, Republic of Korea, in 2015. He is currently working toward the MS degree in Information and Communication Engineering from Chungbuk National University. His research interests include mobile relay networks and wireless sensor networks.

Heejung Yu received his BS in radio science and engineering from Korea University, Seoul, Republic of Korea, in 1999, and his MS and PhD in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea, in 2001 and 2011, respectively. He is currently an assistant professor with the Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, Republic of Korea. From 2001 to 2012, he was a senior researcher with ETRI, Daejeon, Republic of Korea. His areas of interest include statistical signal processing and communication theory. Prof. Yu was the recipient of the Bronze Prize in the 17th Humantech Paper Contest and the Best Paper Award in the 21st Joint Conference on Communications and Information (JCCI) in 2011.