Abstract—Multihop relaying in WiMAX networks is considered an increasingly attractive technology for providing throughput enhancement, coverage extension, or both. In this paper, we consider deploying non-transparent relay stations (RSs) that transmit on the same carrier frequency as the base station (BS) for the purpose of coverage extension. Since it is possible that the RSs and BS are transmitting/receiving simultaneously to/from the subscriber stations (SSs) during the access zone period, it is challenging to schedule resources optimally for serving SSs in a fair manner. We propose an optimal scheduling scheme that preserves fairness among active SSs, and compare its performance with two different resource allocation schemes, namely, the overlapped and orthogonal in terms of cell throughput and outage performance. Our numerical results show that proposed scheme can achieve more throughput than the orthogonal scheme, while maintaining the fairness and lower outage performance of the orthogonal one.

I. INTRODUCTION

Despite the promises on high speed data service by most broadband wireless service providers, in practice there are still problems such as dead spots due to shadowing, and non-uniformly distributed traffic in densely populated areas. Deploying additional base stations (BS) can be a simple solution to the problems above, however, due to the high cost of dense BSs deployment, using relay stations (RSs) is considered a cost-effective solution. IEEE 802.16j for WiMAX with RSs has been published by IEEE last year [1]. Furthermore, multihop relaying is being considered an essential feature in the IEEE 802.16m standard and 3GPP Long Term-Evolution (LTE) Advanced standard.

According to the IEEE 802.16j standard, RSs are classified into two categories: transparent mode and non-transparent mode. Transparent RSs do not provide framing information to the associated subscriber stations (SSs), hence, they cannot provide coverage extension and only aim to enhance the throughput within the BS cell coverage, whereas non-transparent RSs transmit framing information as well as data traffic to the SSs, and hence, are appropriate for coverage extension and/or throughput improvement. The standard also allows non-transparent RSs to transmit either on the same carrier frequency or on different carrier frequencies than that of superordinate station. In this paper, we focus on deploying non-transparent RSs transmitting on the same carrier frequency as the BS in order to maximize the spectrum reuse. A non-transparent mode frame structure [1] consists of a downlink (DL) subframe followed by an uplink (UP) subframe. Each subframe is further divided into access zone and relay zone. The BS and the RSs are allowed to transmit to their associated SSs simultaneously during the access zone period. In other words, frequency reuse is possible during the access zone.

Since the standard does not specify any particular scheduling scheme, it is of interest to explore how the well known max-min fairness scheme [2] can be implemented in 802.16j networks, especially in the non-transparent RS mode. Under this max-min fairness constraint, we propose an optimal scheduling scheme by assigning transmission opportunities to each link in the network in order to maximize cell throughput. We also evaluate the performance of the proposed optimal scheduling scheme by comparing its performance with two previously proposed schemes in the literature: the overlapped and orthogonal allocation schemes. The overlapped scheme increases throughput by fully reusing resources, i.e., focusing on downlink, the BS and RSs can be active during the entire access zone period, but it can lead to a significant interference between the BS to SS and RS to SS links. Meanwhile, the orthogonal scheme does not allow any frequency reuse, but minimizes interferences, i.e., the BS and RSs are transmitting to the SSs by using different time slots or different subchannels. The overlapped scheme can achieve a higher throughput than the orthogonal one, but at the same time its outage performance is far worse than that of the orthogonal scheme. The main goal of this work is to maximize the frequency reuse efficiency, while avoiding outage due to interference. Our numerical results show that the proposed optimal scheduling scheme can achieve more throughput than the orthogonal case, while maintaining the lower outage of the orthogonal case.

The rest of this paper is organized as follows. In the next section, we discuss related work. In Section III, we present the system model including SINR analysis. In Section IV, we present the proposed optimal resource allocation scheme, as well as the overlapped and orthogonal schemes under max-min fairness constraints. Numerical results and analysis are shown in Section V. Finally, Section VI concludes the paper.

II. RELATED WORK

Recent research efforts have been devoted to quantifying the benefits of using relays in multihop WiMAX systems. The work by Hoymann et al. [3] show an analytical approach for dimensioning cellular multihop WiMAX networks and
analyze the network capacity by placing RSs at the border of the BS’s transmission range for coverage extension. However, the scheduling problem was not considered. Several studies have considered the problem of scheduling for OFDMA based WiMAX networks [4], [5]. Sayenko et al. [4] present a round-robin based scheduling solution for the 802.16 BS that ensures the QoS requirements of SSs in the uplink and downlink directions. Work by Deb et al. [5] studies the proportional fair scheduling problem, while taking into account frequency selectivity and multiuser diversity. However, the scheduling schemes in [4], [5] restrict the transmission opportunity to one node at a time and, hence they do not optimally utilize the capacity of the network.

According to the standard [1], the BS and RSs can transmit simultaneously for the non-transparent RS mode. Park et al. [6] have recently investigated bandwidth efficiency of a multihop relay along with outage performance by comparing two different resource allocation scenarios named overlapped and orthogonal. However, the boundary between access and relay zone intervals was not determined dynamically based on the traffic load and the fairness scheme was not considered. Fairness has been extensively studied in various wireless network areas [2], [7], [8]. Max-min fairness is the most common fairness definition in both wired and wireless communication networks. In our previous paper [9], we show how well-known fairness schemes such as max-min and proportional fairness can be implemented in 802.16j systems.

In contrast with the prior work, we propose an optimal scheduling scheme under a max-min fairness constraint. Thus, the boundary between the access and relay zone intervals is determined optimally based on the traffic load. Although we consider only the downlink analysis, our scheme can be easily extended to include the uplink.

### III. System Model

We consider a WiMAX network enhanced with non-transparent RSs and each cell having a three-tiered MMR network architecture comprised of a BS, RSs, and SSs. We assume that every node has a single omni-directional antenna, hence, no terminal can transmit and receive simultaneously. Although the standard does not place a limit on the number of hops from the BS to the SSs, in this paper the maximum number of hops is limited to two since an increased number of hops causes an increased delay and decreased bandwidth efficiency. We assume a reuse factor of 7, one sector per cell, and hilly terrains with heavy tree densities (i.e., terrain type A [10]). For the path loss model, we use the Erceg-Greenstein model, which is recommended by the IEEE 802.16 working group [10].

Depending on the link quality, a variety of modulation and coding schemes (MCS) are supported in WiMAX networks. Table I shows the achievable data rates denoted as $d_1, d_2, ..., d_7$ and the corresponding MCS; the last column represents the minimum required threshold values of signal to interference and noise ratio (SINR), $\gamma_m$, computed by the bit error rate expression for M-QAM [11] when bit error rate is $10^{-6}$. With the assumption of a Rayleigh fading channel, the received SINR, $\gamma$, is an exponential random variable [12]. Therefore, the probability that a transmitter can achieve data rate $d_m$ can be expressed as:

$$p(d_m) = \int_{\gamma_m}^{\gamma_{m+1}} \frac{1}{\gamma^*} \exp \left( -\frac{\gamma}{\gamma^*} \right) d\gamma,$$

where $\gamma^*$ is the average SINR. Consequently, the average achievable data rate, $d_s$, can be computed by:

$$d_s = \sum_{m=1}^{7} d_m \cdot p(d_m).$$

The relay data rate (BS-RS-SS) is influenced by the link capacities of both hops involved. In order to utilize channel bandwidth efficiently, i.e., avoid wasting resource and over-flowing data, the incoming and outgoing data at the relays should be equal:

$$d_{BS-RS} \cdot t_A = d_{RS-SS} \cdot t_B,$$

where $d_{BS-RS}$ and $d_{RS-SS}$ are the capacities of BS to RS and RS to SS links respectively when each link is allocated the entire bandwidth, and $t_A$ and $t_B$ are the durations of BS to RS and RS to SS link allocations respectively. The average data rate of an SS using a relay is equal to the amount of data received divided by the time required to receive it:

$$d_{BS-SS} = \frac{d_{BS-RS} \cdot t_A}{t_A + t_B},$$

as the RS cannot receive from the BS while transmitting to the SS. Using (3), (4) can be rewritten as:

$$\frac{1}{d_{BS-SS}} = \frac{1}{d_{BS-RS}} + \frac{1}{d_{RS-SS}}.$$
Each contour line represents the achievable average data rate of an SS according to its location inside a cell. Since the impact of both placement and number of RSs on throughput and coverage of the cell were shown in our previous paper [14], we consider only three RSs deployed at the edge of the BS’s coverage in this work.

Due to the fact that the original tile (two-dimensional time×frequency) scheduling problem is NP-hard, and is hard to approximate [5], we shall not deal with multiuser resource allocation over the frequency domain. In other words, we do not consider frequency selectivity, thus the entire spectrum is allocated to each node whenever they are allowed to transmit, i.e., scheduling is done by assigning time slots to every node.

A. Optimal Scheme

A key task of an optimal scheduling scheme is to determine the time duration of the transmissions for each SS and RS in the access and relay zones in order to maximize cell throughput under the max-min fairness constraint. There should not be any wasted resources under an optimal scheme. Thus, for the optimal scheduling scheme, none of active SSs can achieve more throughput without decreasing the throughput of other SSs. In other words, the max-min fairness under optimal scheme is equivalent to absolute fairness. We formulate the optimization problem as a linear programming problem. We denote with $\mathcal{R}$ and $\mathcal{S}$ the set of RSs and SSs respectively. All RSs and the BS are referred to as service nodes, and the set of the service nodes is denoted as $\mathcal{R}^+$ (i.e., $|\mathcal{R}^+| = |\mathcal{R}| + 1$). Multiple service nodes may be active simultaneously during the access zone to maximize frequency reuse, but in this case some areas of the cell will not be able to be served by those nodes due to insufficient SINR. To determine the optimal schedule we need to compute the achievable data rate for each SS for each combination of active service nodes in $\mathcal{R}^+$. Therefore we need to consider all possible combinations of active service nodes ($\mathcal{U}$ the power set of $\mathcal{R}^+$ excluding the empty set) and determine the signal strengths for each SS from each active service node.

For each combination of active service nodes $u \in \mathcal{U}$, an SS can be served by only the service node that has highest link capacity; it is possible that some SSs will not be served in some subsets of $\mathcal{R}^+$ due to interference. For example, if an SS $s_1$ receives the same signal strength from two RSs $r_1$ and $r_2$, then $s_1$ can be served by either $r_1$ or $r_2$ but not both at the same time, i.e., $s_1$ cannot be served in the subsets containing both $r_1$ and $r_2$.

Scheduling objective:

Since every SS will achieve an equal throughput due to the equivalence of the absolute and max-min fairness constraint for this system, our goal of maximizing cell throughput corresponds to maximizing the throughput, $T_s$, of any subscriber $s \in \mathcal{S}$:

$$\max_{s \in \mathcal{S}} T_s. \tag{6}$$

The achievable data rate of an SS, $d_s$, varies for each subset of active service nodes because intra cell interference changes according to the number of active service nodes in each subset. Thus, we denote the achievable data rate of an SS in each transmission subset $u$ by $d_u^s$. Also, the time fraction allocated to an SS, $\lambda_s$, in each transmission subset $u$ will be denoted as $\lambda_u^s$. The throughput of an SS, $T_s$, during the current DL subframe is the summation of throughputs received by $s$ in each transmission subset $u$ when $s$ was allocated the time fraction $\lambda_u^s$.

$$T_s = \sum_{u \in \mathcal{U}} d_u^s \lambda_u^s, \quad \forall s \in \mathcal{S}. \tag{7}$$

Scheduling constraints:

The constraints that the schedule should satisfy are as follows:

1. Max-min fairness constraints: this constraint ensures that every active SS in a cell achieves an equal throughput:

$$T_{s_1} = T_{s_2}, \quad \forall s_1, s_2 \in \mathcal{S} \quad (s_1 \neq s_2). \tag{8}$$

2. Relaying constraints: let $d_r$ and $\lambda_r$ be the achievable data rate and time fraction allocated for transmission from the BS to RS $r \in \mathcal{R}$ respectively. The set of SSs associated with RS $r \in \mathcal{R}$ in each transmission subset $u$ is denoted as $\mathcal{S}_r^u$. To ensure that there is no data loss at the RSs, the data transferred
from BS to RS $r \in \mathcal{R}$ is equal to the data transferred from RS $r$ to the associated SSs:

$$d_r \lambda_r = \sum_{u \in \mathcal{U}} \sum_{s \in \mathcal{S}_u} d_s^u \lambda_s^u, \quad \forall r \in \mathcal{R}. \quad (9)$$

3. Time sharing constraints 1: this constraint ensures that resources within the duration of each subset $u$ are fully utilized by the associated SSs. Let $\mathcal{S}_u^r$ denote the set of SSs associated with service node $r^+ \in \mathcal{R}^+$ in each transmission subset $u$. To ensure that every service node $r^+$ fully use resources, the summation of time fractions allocated to SSs associated with $r^+_1$ should be equal to the summation of time fractions allocated to SSs associated with $r^+_2$ when $r^+_1$ and $r^+_2$ are active service nodes in subset $u$:

$$\sum_{s \in \mathcal{S}_u^r} \lambda_s^u = \sum_{s \in \mathcal{S}_u^r} \lambda_s^u, \quad \forall r^+_1, r^+_2 \in u, \forall u \in \mathcal{U}, |\mathcal{S}_u^r| > 0, |\mathcal{S}_u^r| > 0. \quad (10)$$

4. Time sharing constraints 2: this constraint captures the fact that the DL subframe consists of an access zone and a relay zone. We denote with $\lambda^u$ the time fraction of each subset $u$. From (10), $\lambda^u$ can be expressed as:

$$\lambda^u = \sum_{s \in \mathcal{S}_u^{r^+}} \lambda_s^u, \quad \forall r^+ \in u, \forall u \in \mathcal{U}, |\mathcal{S}_u^{r^+}| > 0. \quad (11)$$

Thus, the summation of time fractions of every subset will be equivalent to access zone time fraction, and the summation of time fractions allocated to RSs is the same as relay zone time fraction. The sum of access and relay zone time fractions should be less than or equal to 1:

$$\sum_{u \in \mathcal{U}} \lambda^u + \sum_{r \in \mathcal{R}} \lambda_r \leq 1. \quad (12)$$

5. Lower and upper bound constraints: finally, the lower bound and upper bound of each element should be set up to solve the optimization problem by using linear programming. The time fractions allocated to SSs and RSs have to be positive and smaller than one:

$$0 \leq \lambda_s^u, \lambda_r \leq 1, \quad \forall u \in \mathcal{U}, \forall s \in \mathcal{S}, \forall r \in \mathcal{R}. \quad (13)$$

By using scheduling constraints (8), (9), (10), (12), (13), the objective function (6) can be maximized. The solution produced by solving linear programming is always feasible. Once the throughput of an SS (6) is maximized, the cell throughput can be computed by:

$$\text{Cell Throughput} = \sum_{s \in \mathcal{S}} T_s. \quad (14)$$

**B. Orthogonal Scheme**

The essential consideration in the orthogonal scheduling scheme is to avoid interference between BS to SS and RS to SS links by restricting transmission opportunities to one service node at a time during the access zone interval. However, the radio resource efficiency can be significantly reduced by precluding frequency reuse. The optimization problem presented above can also be used for this orthogonal scheme by modifying the definition of set $\mathcal{U}$. In the optimal scheduling scheme, the cardinality of $\mathcal{U}$ is $2|\mathcal{R}^+| - 1$ because every possible transmit combination of active service nodes has to be considered. However, in the orthogonal scheduling scheme, only subsets of $\mathcal{R}^+$ that have a cardinality of 1 are considered. Let us denote with $\mathcal{U}' = \{u \in \mathcal{U} : |u| = 1\} = \{\{BS\}, \{r_1\}, ..., \{r_{|\mathcal{R}|}\}\}$ the set of subsets of $\mathcal{R}^+$ such that each element $u \in \mathcal{U}'$ has a cardinality of 1, and each element represents a service node that can be active during access zone.

By substituting $\mathcal{U}'$ for $\mathcal{U}$, the optimal scheme in Section IV-A can be formulated as an orthogonal scheme. An active SS can be associated with more than one service nodes at different times during the access zone. The objective of the orthogonal scheme is the same as that of optimal scheme (6), but the relaying constraint (9) can be simplified as:

$$d_r \lambda_r = \sum_{s \in \mathcal{S}_r} d_s^u \lambda_s, \quad \forall r \in \mathcal{R}. \quad (15)$$

Also, the time sharing constraints (10), (12) can be simplified as followings:

$$\sum_{r^+ \in \mathcal{R}^+} \sum_{s \in \mathcal{S}_u^{r^+}} \lambda_s^u + \sum_{r \in \mathcal{R}} \lambda_r \leq 1. \quad (16)$$

Once $T_s$ is maximized by using linear programming with objective (6) under constraints (8), (13), (15), (16), the cell throughput for the orthogonal scheme can also be computed using (14).

**C. Overlapped Scheme**

In contrast to the orthogonal scheme, the focus of overlapped scheme is to fully reuse radio resources during the access zone interval. In general, when there are many SSs distributed uniformly throughout the cell, all service nodes should be active during the entirety of the access zone interval to maximize frequency reuse. However, when there are few SSs or non-uniformly distributed in a cell, not every service node has to be active because some of the service nodes do not have associated SSs to transmit, hence, those are only introducing co-channel interference.

Unlike the optimal and orthogonal schemes, the overlapped scheme will consider only one subset of $\mathcal{R}^+$, in other words, once the set of active service nodes is determined at the beginning of a frame, it will be active for the entire current frame duration. Therefore, an active SS can be served by only one service node during the access zone. An interesting problem is determining which service nodes should be active at the beginning of a frame. For determining the set of active service nodes, we introduce additional set selection objectives such as maximizing cell throughput, maximizing the number of active service nodes or maximizing the number of served SSs. Depending on these set selection objectives, the maximum cell throughput achieved with the optimal scheduling scheme can be significantly different for each case. For each of these three cases the set of active service nodes will be optimally scheduled to maximize the throughput as the
scheduling objective by our scheme, but the composition of this set of active nodes will be different depending on the set selection optimization objective. Figures 2 and 3 show the cell throughput and outage rate as a function of the number of active SSs for different set selection optimization objectives. The maximum throughput can be achieved in the max throughput set selection objective case, but, at the same time, the outage rate for this case is much higher than in the other cases. Similarly, the lowest outage rate is achieved by max served nodes set selection objective case, but the cell throughput of that case is very low. From a service provider’s point of view, we consider max served nodes case as the most realistic set selection objective for the overlapped scheme.

In order to compute the maximum cell throughput after determining the active service nodes, we can also use the optimal scheduling scheme formulation in Section IV-A by considering only one subset \( u \) instead of the set \( U \). One unique aspect of overlapped scheme is that there may be wasted resources due to fairness constraint under only one subset scenario. Moreover, the equivalence of the absolute and max-min fairness constraint does not hold for the overlapped scheme, because the SSs associated with BS directly may achieve a higher throughput without decreasing the throughput of the rest of SSs associated with the RSs. The details on the fairness schemes can be found in our previous paper [9]. As a result, the scheduling objective of overlapped scheme is to maximize cell throughput:

\[
\max \sum_{s \in S} T_s. \tag{17}
\]

Due to the fact that an SS can be associated with only one service node during the access zone in the overlapped scheme, the set \( S \) can be divided into two disjoint subsets \( S_b \) and \( S_R \). The set \( S_b \) represents the SSs served by the BS and \( S_R \) represents the SSs served by RSs. The max-min fairness constraint ensures that every SS achieves equal throughput in each subset \( S_b \) and \( S_R \). However, the throughput of an SS in \( S_b \) could be higher than the throughput of an SS in \( S_R \). Therefore, the max-min fairness constraint for overlapped scheme can be expressed as:

\[
\begin{align*}
T_{s_1} &= T_{s_2}, \quad \forall s_1, s_2 \in S_b \ (s_1 \neq s_2) \\
T_{s_3} &= T_{s_4}, \quad \forall s_3, s_4 \in S_R \ (s_3 \neq s_4) \\
T_{s_5} &\geq T_{s_6}, \quad \forall s_5 \in S_b, \ \forall s_6 \in S_R.
\end{align*}
\tag{18}
\]

Since the overlapped scheme considers only one subset of \( \mathcal{R}^+ \), the relaying constraint will be the same as (15). We denote with \( S^* \) the set of SSs associated with one of RSs where the resources are fully utilized over the access zone interval. Only one of RSs can fully use resources due to max-min fairness under the overlapped scheme. Therefore, the time sharing constraint in (16) can be further simplified as:

\[
\sum_{s \in S^*} \lambda_s + \sum_{r \in R} \lambda_r \leq 1. \tag{19}
\]

The cell throughput for the overlapped scheme can be maximized by solving linear programming with the objective (17) under constraints (13), (15), (18), (19).

V. NUMERICAL RESULTS AND ANALYSIS

In this section we evaluate the performance of the proposed optimal resource allocation scheme by comparing it with the orthogonal and overlapped schemes. We analyze the cell throughput and outage rates as a function of the number of active SSs in a cell. From the optimization problems presented in Section IV, the cell throughput for each scheduling scheme can be computed by solving the corresponding linear programming problems. Figure 4 shows the cell throughput results from the three different scheduling schemes. To obtain the average cell throughput value, the simulation is repeated 10,000 times for each scenario with \( N \) active SSs randomly placed within a cell. We computed the 95% confidence intervals, but do not show them as they are very small and would clutter the graphs.
When there is only one active SS in a cell, hence no frequency reuse and co-channel interference, the average throughput of the SS under different scheduling schemes is equal to 10.67Mbps. As the number of active SSs increases, the cell throughput achieved by the orthogonal scheme decreases because it is more likely to have SSs with low link capacities consuming a large fraction of the time in order to preserve fairness. In contrast, the cell throughput for the proposed optimal scheme is growing as the number of active SSs increases since the optimal scheme maximizes frequency reuse, while minimizing co-channel interference to avoid outage. The cell throughput for the overlapped scheme with max served nodes set selection objective is lower than that of the proposed optimal scheme.

Figure 5 shows the outage rate as a function of the number of active SSs within a cell for the three different scheduling schemes. The outage rate from proposed optimal scheme is identical to the result from the orthogonal scheme. Although there is no interference between service nodes in the orthogonal scheme, about 5% of active SSs still encounter outage due to the Rayleigh fading channel. This result shows that the cell coverage probability under Rayleigh fading channel was improved from 90% when not using RSs to 95% when using three RSs in addition to coverage extension. In contrast, the outage rate for the overlapped scheme continues to rise significantly as more SSs join the cell because the number of active service nodes also increases leading to insufficient SINR for many SSs. Therefore, the cell throughput can be significantly enhanced by using the proposed optimal scheme compare to orthogonal scheme while preserving the low outage performance as orthogonal scheme and fairness.

VI. CONCLUSION

In this paper we studied the optimal resource allocation problem in 802.16j based networks enhanced with non-transparent relays. Since frequency reuse is allowed during the access zone, our objective was to maximize the frequency reuse efficiency under max-min fairness constraint. We formulated the optimization problem to maximize cell throughput and provide solutions by using linear programming. We evaluated the performance of our proposed optimal scheme by comparing it with well-known scheduling schemes such as the orthogonal and overlapped schemes. Our results showed that the cell throughput achieved by proposed optimal scheme is much higher than that of orthogonal scheme, while maintaining as low outage rate as orthogonal scheme.

REFERENCES