Uplink capacity of multi-class IEEE 802.16j relay networks with adaptive modulation and coding

Wang, Hua; Xiong, C; Iversen, Villy Bæk

Published in:

Link to article, DOI:
10.1109/ICC.2009.5199552

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Uplink Capacity of Multi-class IEEE 802.16j Relay Networks with Adaptive Modulation and Coding

Hua Wang†, Chunlin Xiong‡, and Villy B. Iversen†

†Department of Communications, Optics & Materials, Technical University of Denmark, Lyngby, Denmark
‡School of Electronic Science and Engineering, National University of Defense Technology, Changsha, China
Email: †{huw, vbi}@com.dtu.dk, ‡xchlzju@nudt.edu.cn

Abstract—The emerging IEEE 802.16j mobile multi-hop relay (MMR) network is currently being developed to increase the user throughput and extend the service coverage as an enhancement of existing 802.16e standard. In 802.16j, the intermediate relay stations (RSs) help the base station (BS) communicate with those mobile stations (MSs) that are either too far away from the BS or placed in an area where direct communication with BS experiences unsatisfactory level of service. In this paper, we investigate the uplink Erlang capacity of a two-hop 802.16j relay system supporting both voice and data traffics with adaptive modulation and coding (AMC) scheme applied in the physical layer. We first develop analytical models to calculate the blocking probability in the access zone and the outage probability in the relay zone, respectively. Then a joint algorithm is proposed to determine the bandwidth distribution between the access zone and the relay zone, and to derive the Erlang capacity region of the system. The numerical examples show that some capacity gains can be obtained with a relay-enhanced 802.16j system compared to the conventional single-hop 802.16e system.

I. INTRODUCTION

IEEE 802.16 standard, also known as WiMAX, has emerged as an advanced broadband wireless access technology and has attracted tremendous attention in both academia and industry. It is expected that future mobile communication systems are evolving to provide a wide range of services, including voice, data and multimedia applications with higher transmission rate and better quality of service (QoS).

However, due to significant loss of signal strength along the propagation path and the transmit power constraint of mobile stations (MSs), the coverage area of a legacy single-hop cellular network is often geographically limited. Traditional solution to this problem is to deploy additional base stations (BSs). Unfortunately, the high cost of BSs and potential aggravation of interference make this approach less favorable. Alternatively, relay stations (RSs) that receive and forward signals between the BS and MSs are introduced to help BS communicate with those MSs that are either too far away from the BS or placed in an area where direct communication with BS experiences unsatisfactory level of service. Recently, 802.16j task group was formed to amend the current 802.16e standard to support mobile multi-hop relay (MMR) operation. In an MMR network, MSs are allowed to route through intermediate RSs to reach the BS, which differs from the legacy single-hop 802.16e system.

Reference [1] & [2] show that with appropriate deployment of RSs in an MMR network, significant performance gains can be obtained in terms of the user throughput as well as the cellular coverage compared to conventional single-hop cellular networks. The capacity gain of an 802.16j system is investigated in [4] via simulations. In this paper, we propose an analytical model to calculate the uplink Erlang capacity of multi-class multi-hop relay networks with adaptive modulation and coding (AMC) scheme applied in the physical layer. With AMC, the allocation of bandwidth to each user is no longer deterministic, i.e., a fixed amount of bandwidth, but in a dynamic behavior. Therefore, the calculation of the Erlang capacity should take both the call admission control (CAC) module at the data link layer as well as the AMC module at the physical layer into considerations. We first present analytical models to calculate the blocking probability in the access zone and the outage probability in the relay zone, respectively. Based on that, a joint algorithm to determine the Erlang capacity of the system is proposed. The Erlang capacity here is defined as the maximum traffic load among different service classes that the system can support when both the blocking and the outage probabilities do not exceed certain thresholds.

The rest of the paper is organized as follows. In Section II, we introduce the system model as well as the new frame structure in 802.16j networks. In Section III, analytical models for calculating the blocking probability in the access zone and the outage probability in the relay zone are presented, followed by a joint algorithm to determine the system Erlang capacity. Numerical results are shown in Section IV. Finally, conclusions are drawn in Section V.

II. SYSTEM MODEL

In order to improve the capacity and extend the coverage range without compromising backward compatibility with the legacy MSs, the 802.16j standard has extended the 802.16e frame structure to support mobile multihop relay capability.

A high level frame structure of 802.16j systems operated in OFDMA-TDD mode is illustrated in Fig. 1. The frame structure supports a typical two-hop relay-enhanced communication, where some MSs communicate directly to the BS, while other MSs are attached to a RS and communicate to the BS via the RS. Frame sections in grey denote receive operation, whereas sections in white denote transmit operation. Even though extensions of the frame structure to support more than two hops have been defined by the work group, here we only concentrate on a two-hop relay scenario, as the
communication path between the MSs and BS is usually no more than two-hop away in conventional cellular networks.

At the physical layer, the time axis is divided into frames with fixed length, each of which consists of a downlink (DL) and an uplink (UL) subframe to support TDD operation. In order to enable multihop communication, both downlink and uplink subframes are further divided into the access and relay zones. The access zone is dedicated for communication that directly engages MSs, and it is backward compatible with the 802.16e standard. Specifically, MSs receive from or transmit to the BS or RS with which they are associated with in the access zone of downlink and uplink subframe, respectively. The access zone in both downlink and uplink is followed by a relay zone. In the relay zone, the RS receives from or transmits to the BS in the corresponding downlink or uplink subframe, respectively. For legacy MSs, the notion of relay zone remains transparent to them, as they will only become aware of the existence of some new zones following the access zone based upon the DL-MAP and UL-MAP, and thus simply stay idle during these relay zones. In an MMR network, the radio link between MSs and their superordinate BS or RS is defined as the access link, while the radio link between the RS and the BS is defined as the relay link.

In this paper, we only focus on the uplink capacity of a relay-enhanced system. In each UL subframe, assume there are $M$ subchannels in frequency domain and $L$ time slots in time domain, which results in a total of $K = M \times L$ OFDMA slots available in frequency-time domain for UL transmissions. A slot is the minimum resource allocation unit as defined in the 802.16 standard. Adaptive modulation and coding scheme is employed at the physical layer to adjust the transmission mode dynamically according to the time-varying channel conditions. Multiple transmission modes are available, with each mode representing a pair of specific modulation format and a forward error correcting code. The transmission mode is determined by the instantaneous signal-to-noise ratio (SNR).

Two types of service, voice service and data service, are supported by the system. Voice service is not tolerant to packet delay, thus requires a constant bit rate of $R_v$ bits per frame. Data service is more elastic in terms of being able to vary the transmission rate according to the channel conditions, but also requires a minimum throughput of $R_d$ bits per frame. Voice and data users arrive at the cell in a random order.

III. ANALYTICAL MODELS

A. PHY Layer Transmission with AMC

We assume that each subchannel is frequency flat, and remains constant within a frame, but may vary from frame to frame. We further assume that the channel quality of each subchannel is independent identically distributed (i.i.d.), and follows a Rayleigh fading. For flat Rayleigh fading channels, the received SNR is a random variable $\gamma$ with probability density function (pdf) [5]:

$$p_\gamma(\gamma) = \frac{1}{\pi} \exp\left(-\frac{\gamma}{\bar{\gamma}}\right)$$  \hspace{1cm} (1)

where $\bar{\gamma}$ is the average received SNR.

Adaptive modulation and coding scheme has been widely adopted in the physical layer of modern communication systems. The design objective of AMC is to maximize the data rate by adjusting the transmission parameters according to time-varying channel conditions, while maintaining a prescribed packet error rate (PER) $P_0$ over wireless links. Let $N$ denote the total number of transmission modes available (e.g., $N = 6$). Assuming constant power transmission, we partition the entire SNR range into $N$ non-overlapping consecutive intervals with boundaries denoted as $\{\Gamma_{n}\}_{n=1}^{N-1}$. Specifically, mode $n$ is chosen when $\gamma \in [\Gamma_{n-1}, \Gamma_n)$. Therefore, with Rayleigh fading, mode $n$ will be chosen with probability:

$$P_t(n) = \frac{\Gamma_n}{\Gamma_n - \Gamma_{n-1}} p_\gamma(\gamma) d\gamma$$

$$= \exp\left(-\frac{\Gamma_n}{\bar{\gamma}}\right) - \exp\left(-\frac{\Gamma_{n+1}}{\bar{\gamma}}\right)$$  \hspace{1cm} (2)

<table>
<thead>
<tr>
<th>Mode</th>
<th>Mode 2</th>
<th>Mode 3</th>
<th>Mode 4</th>
<th>Mode 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modulation:</td>
<td>BPSK</td>
<td>QPSK</td>
<td>QPSK</td>
<td>16QAM</td>
</tr>
<tr>
<td>Coding rate</td>
<td>1/2</td>
<td>1/2</td>
<td>3/4</td>
<td>3/4</td>
</tr>
<tr>
<td>$R_m$ (bits/sym)</td>
<td>0.5</td>
<td>1.0</td>
<td>1.5</td>
<td>3.0</td>
</tr>
<tr>
<td>$\alpha_n$</td>
<td>274.7229</td>
<td>90.2512</td>
<td>67.6181</td>
<td>53.3987</td>
</tr>
<tr>
<td>$g_n$</td>
<td>7.9932</td>
<td>3.4998</td>
<td>1.6883</td>
<td>0.3756</td>
</tr>
<tr>
<td>$\gamma_{\text{PRE}}$ (dB)</td>
<td>$-1.5331$</td>
<td>$1.0942$</td>
<td>$3.9722$</td>
<td>$10.2488$</td>
</tr>
</tbody>
</table>

TABLE I

TRANSMISSION MODES WITH CONVOLUTIONALLY CODED MODULATION [5]

The closed-form of the average PER corresponding to mode $n$ is obtained as [5]:

$$\text{PER}_n = \frac{1}{P_t(n)} \int_{\Gamma_{n-1}}^{\Gamma_n} \alpha_n \exp(-g_n \gamma) p_\gamma(\gamma) d\gamma$$  \hspace{1cm} (3)
where $a_n$, $g_n$ are the mode dependent parameters shown in Table I.

The algorithm for determining the thresholds $\{f_{n}\}_{n=1}^{N-1}$ with the prescribed $P_{FIR} = P_f$ is described in details in [5].

Let $R$ be a random variable with probability mass function $f_R$, denoting the number of bits that can be transmitted in one slot.

$$\mathcal{R} \in \{sR_0, sR_1, \cdots, sR_{N-1}\}$$

$$f_R(sR_n) = \mathbb{P}(\mathcal{R} = sR_n) = P_t(n)$$

(4)

where $R_n$ is the number of bits carried per symbol in transmission mode $n$ as shown in Table I, and $s$ is the number of symbols per time slot.

As mentioned in the previous section, uplink subframes are further divided into the access zones and relay zones. In UL access zone, MSs transmit to the BS or RS with which they are associated with, while in UL relay zone, the RS relays the traffic received from the MSs to the BS. Let random variables $\gamma_{M2B}$, $\gamma_{M2R}$, and $\gamma_{R2B}$ denote the average received SNR on link MS-to-BS, MS-to-RS, and RS-to-BS, respectively. Generally it is expected that $\gamma_{R2B}$ enjoys better quality than $\gamma_{M2B}$ and $\gamma_{M2R}$, as it is likely that the propagation conditions on the RS-to-BS link are less severe than that on the MS-to-RS and MS-to-BS links. From Eqn. (1) - (4), we can obtain the number of bits that can be transmitted in one slot over different links, denoted as random variables $R_{M2B}$, $R_{M2R}$, and $R_{R2B}$, with probability mass function $f_{R_{M2B}}$, $f_{R_{M2R}}$, and $f_{R_{R2B}}$, respectively.

**B. Blocking Probability in the Access Zone**

In UL access zone, MSs transmit to BS or RS with which they are associated with. It is assumed that voice and data users associated with the MS-to-BS access zone arrive following a Poisson process with rate $\lambda_{v}^{M2B}$ and $\lambda_{d}^{M2B}$, respectively, and voice and data users associated with the MS-to-RS access zone arrive following a Poisson process with rate $\lambda_{v}^{M2R}$ and $\lambda_{d}^{M2R}$, respectively. The service time for voice and data users is exponentially distributed with intensity $\mu_{v}$ and $\mu_{d}$, respectively. We assume that $K_{Access}$ slots are allocated to the UL access zone per frame.

For a given quality of service, $R_v$ ($R_d$) bits are allocated to a voice (data) user per frame. In AMC scheme, the modulation and coding rate is chosen according to time-varying channel conditions. As a consequence, the number of slots allocated to each user is varying from frame to frame. Let $N_{v}^{M2B}$ ($N_{d}^{M2B}$) be a random variable denoting the number of slots required by a voice (data) user associated with the MS-to-BS access zone per frame, and $N_{v}^{M2R}$ ($N_{d}^{M2R}$) be a random variable denoting the number of slots required by a voice (data) user associated with the MS-to-RS access zone per frame. The probability mass function of random variables $N_{v}^{M2B}$, $N_{d}^{M2B}$ $N_{v}^{M2R}$, and $N_{d}^{M2R}$ are denoted as $f_{N_{v}^{M2B}}$, $f_{N_{d}^{M2B}}$ $f_{N_{v}^{M2R}}$, and $f_{N_{d}^{M2R}}$, correspondingly, and can be obtained from $f_{R_{M2B}}$ and $f_{R_{M2R}}$:

$$N_{v}^{M2B} = \frac{R_{v}}{R_{M2B}}$$

$$N_{d}^{M2B} = \frac{R_{d}}{R_{M2B}}$$

$$N_{v}^{M2R} = \frac{R_{v}}{R_{M2R}}$$

$$N_{d}^{M2R} = \frac{R_{d}}{R_{M2R}}$$

(5)

where $R_{M2B}$ and $R_{M2R}$ are random variables denoting the number of bits that can be transmitted in one slot on MS-to-BS link and MS-to-RS link, respectively.

Let us denote the system state to be $(i, j, m, n)$, where $i$ and $j$ denote the number of admitted voice and data users associated with the MS-to-BS access zone, and $m$ and $n$ denote the number of admitted voice and data users associated with the MS-to-RS access zone. Assume that the system is in state $(i, j, m, n)$, then the total number of slots required by the admitted voice and data users in the access zone can be modeled by a random variable $D_{i,j,m,n}$:

$$D_{i,j,m,n} = \sum_{i=1}^{i} \sum_{j=1}^{j} \sum_{m=1}^{m} \sum_{n=1}^{n} N_{v}^{M2B} + \sum_{i=1}^{i} \sum_{j=1}^{j} \sum_{m=1}^{m} \sum_{n=1}^{n} N_{d}^{M2B} + \sum_{i=1}^{i} \sum_{j=1}^{j} \sum_{m=1}^{m} \sum_{n=1}^{n} N_{v}^{M2R} + \sum_{i=1}^{i} \sum_{j=1}^{j} \sum_{m=1}^{m} \sum_{n=1}^{n} N_{d}^{M2R}$$

(6)

with probability mass function $f_{D_{i,j,m,n}}$:

$$f_{D_{i,j,m,n}} = f_{N_{v}^{M2B}} \otimes \cdots \otimes f_{N_{v}^{M2B}} \otimes f_{N_{v}^{M2R}} \otimes \cdots \otimes f_{N_{v}^{M2R}}$$

$$\otimes f_{N_{d}^{M2B}} \otimes \cdots \otimes f_{N_{d}^{M2B}} \otimes f_{N_{d}^{M2R}} \otimes \cdots \otimes f_{N_{d}^{M2R}}$$

(7)

where $a \otimes b$ denotes discrete convolution.

Outage in the access zone is defined to occur when the total number of slots required by the admitted users exceeds the total number of available slots allocated to the UL access zone. If outage occurs, the admitted users may not get the prescribed data rate and the QoS level might be degraded. If the system is in state $(i, j, m, n)$, the outage probability in the access zone can be calculated as:

$$P_{outage}^{Access}(i, j, m, n) = P(D_{i,j,m,n} > K_{Access})$$

$$= 1 - P(D_{i,j,m,n} \leq K_{Access})$$

$$= 1 - \sum_{d \leq K_{Access}} f_{D_{i,j,m,n}} (d)$$

(8)

The call admission control (CAC) module at the data link layer is responsible for preventing the system capacity from being oversubscribed by limiting the number of ongoing connections. When a new connection arrives at the system, it will be accepted under the condition that the acceptance of the new connection will not cause the outage above a predefined threshold. Therefore, the blocking probability of a new connection depends on both the current state of the system and the bandwidth requirement of the new connection.

Specifically, assume that a voice call associated with the MS-to-BS access zone arrives at the time instant when the system is in state $(i, j, m, n)$. The CAC module estimates the outage probability of the next state, which is $P_{outage}^{Access}(i+1, j, m, n)$. If the estimated outage probability is smaller than the threshold, the new call is accepted, otherwise, the new call is rejected.

Let $a_{i,j,m,n}^{v,M2B}$ be the acceptance probability of a voice call associated with the MS-to-BS access zone arriving in state $(i, j, m, n)$, determined as follows:

$$a_{i,j,m,n}^{v,M2B} = \begin{cases} 
1 & \text{if } P_{outage}^{Access}(i+1, j, m, n) \leq Out_{Th} \\
0 & \text{if } P_{outage}^{Access}(i+1, j, m, n) > Out_{Th}
\end{cases}$$

(9)
where $\text{Out}_{\text{Th}}$ is the predefined outage threshold in the access zone. Similar calculations can be performed to obtain the acceptance probability of $a_{i,j,m,n}$, $b_{i,j,m,n}$, and $d_{i,j,m,n}$.

The call-level characteristics of a multi-class (voice/data) 802.16j relay system with the CAC policy described above can be modeled by a Continuous Time Markov Chain (CTMC). Based on the acceptance probability derived in each system state, we can construct a four-dimensional state transition diagram. Since it is a reversible Markov process (Kolmogorov’s criteria is fulfilled for all possible paths), we can apply the local balance equations to calculate the relative state probabilities, denoted as $q(i, j, m, n)$, with reference to state $(0, 0, 0, 0)$ [7]. The absolute state probabilities can be obtained after normalization of the relative state probabilities:

$$p(i, j, m, n) = \frac{q(i, j, m, n)}{\sum q(j)} \quad (10)$$

The state space of the system is denoted as:

$$S := \{i, j, m, n\} \in \{I \times J \times M \times N\} \quad (11)$$

where $I$, $J$, $M$, $N$ are the maximum number of users that can be accommodated in each traffic class.

Let $\Omega_{v}^{M2B}$ be a set of marginal states for voice users associated with the MS-to-BS access zone:

$$\Omega_{v}^{M2B} := \{(i^{*}, j^{*}, m^{*}, n^{*})\} \quad (12)$$

where the marginal states for voice users associated with the MS-to-BS access zone $(i^{*}, j^{*}, m^{*}, n^{*})$ are defined satisfying the following conditions:

$$p(i^{*}, j^{*}, m^{*}, n^{*}) \neq 0$$

$$p(i^{*}, j^{*} + 1, m^{*}, n^{*}) = 0 \quad (13)$$

$$p(i^{*}, j^{*}, m^{*} + 1, n^{*}) = 0$$

$$p(i^{*}, j^{*}, m^{*}, n^{*} + 1) = 0$$

Due to the PASTA-property in Poisson arrival process, the call congestion is equal to the time congestion, thus the call blocking probability of voice users associated with the MS-to-BS access zone ($P_{\text{block},v}^{M2B}$) can be calculated as:

$$P_{\text{block},v}^{M2B} = \sum_{(i^{*}, j^{*}, m^{*}, n^{*})} p(i^{*}, j^{*}, m^{*}, n^{*}) \quad (14)$$

Similar calculations can be performed to obtain the call blocking probability of the other three traffic classes ($P_{\text{block},d}^{M2R}$, $P_{\text{block},v}^{M2R}$, and $P_{\text{block},d}^{M2R}$).

### C. Outage Probability in the Relay Zone

In UL relay zone, the RS relays the traffic received from the MSs to the BS. To satisfy the flow conservation constraint, the link capacity of RS-to-BS should be high enough to accommodate all the traffic received by the RS in the UL access zone. We assume that $K_{\text{Relay}}$ slots are allocated to the UL relay zone per frame.

Assume that there are $C_{v}^{M2R}$ voice users and $C_{d}^{M2R}$ data users associated with the MS-to-BS access zone admitted in the system, respectively. $C_{v}^{M2R}$ and $C_{d}^{M2R}$ are random variables with probability mass function calculated as follows:

$$P(C_{v}^{M2R} = a) = \sum_{i=0}^{I} \sum_{j=0}^{J} \sum_{m=0}^{M} \sum_{n=0}^{N} p(i, j, a, n) \quad 0 \leq a \leq M \quad (15)$$

$$P(C_{d}^{M2R} = b) = \sum_{i=0}^{I} \sum_{j=0}^{J} \sum_{m=0}^{M} \sum_{n=0}^{N} p(i, j, m, b) \quad 0 \leq b \leq N \quad (16)$$

Let $N_{v}^{R2B}$ ($N_{d}^{R2B}$) be a random variable denoting the number of slots occupied in UL relay zone by a voice (data) user associated with the MS-to-RS access zone.

$$N_{v}^{R2B} = \frac{R_{v}}{R_{R2B}} \quad N_{d}^{R2B} = \frac{R_{d}}{R_{R2B}} \quad (17)$$

where $R_{R2B}$ is a random variable denoting the number of bits that can be transmitted in one slot on RS-to-BS link. The probability mass functions of random variables $N_{v}^{R2B}$ and $N_{d}^{R2B}$ can be easily obtained from the probability mass function of $R_{R2B}$.

Then the total number of required slots in UL relay zone per frame is given by:

$$Z = \sum C_{v}^{M2R} N_{v}^{R2B} + \sum C_{d}^{M2R} N_{d}^{R2B} \quad (18)$$

Outage in the relay zone is defined to occur when the total number of required slots exceeds the total number of available slots allocated to the UL relay zone. If outage occurs, the RS does not have enough capacity to relay the traffic received from MSs associated with the MS-to-RS access zone, thus increase the packet loss probability.

Since $Z$ is a sum of random variables, we can approximate $Z$ to be a Gaussian random variable by applying the central limit theorem. The outage probability in the UL relay zone can be approximated by:

$$P_{\text{outage}} = P(Z > K_{\text{Relay}}) \approx Q\left(\frac{K_{\text{Relay}} - E[Z]}{\sqrt{Var[Z]}}\right) \quad (19)$$

where the mean and variance of $Z$ are given by:

$$E[Z] = E\left\{E[Z \mid (C_{v}^{M2R}, C_{d}^{M2R})]\right\} = E[C_{v}^{M2R}] E[N_{v}^{R2B}] + E[C_{d}^{M2R}] E[N_{d}^{R2B}] \quad (20)$$

$$Var[Z] = \sqrt{\left\{E[Z \mid (C_{v}^{M2R}, C_{d}^{M2R})]\right\}^{2} - E[Z]} = \sqrt{E[C_{v}^{M2R}] Var[N_{v}^{R2B}] + E[C_{d}^{M2R}] Var[N_{d}^{R2B}]} \quad (21)$$

### D. Calculation of System Erlang Capacity

The Erlang capacity in this paper is defined as the maximum value of the offered traffic in each traffic class that the system can support when both the blocking probability in the access
zone and the outage probability in the relay zone do not exceed certain thresholds. The offered traffic $A$ is usually defined as the average number of connection attempts per mean service time. In this paper, the offered traffic of voice users is equal to $A_v = A_v^{MB} + A_v^{MR} = \frac{M^{MB} + M^{MR}}{\mu_v}$, and the offered traffic of data users is equal to $A_d = A_d^{MB} + A_d^{MR} = \frac{M^{MB} + M^{MR}}{\mu_d}$.

In 802.16j MMR networks, there are a total of $K$ slots available for UL transmissions per frame. $K_{Access}$ of them are allocated to the UL access zone, while the remaining $K_{Relay} = K - K_{Access}$ slots are allocated to the UL relay zone. If the system does not allocate enough bandwidth to the UL relay zone, the link capacity of RS-to-BS is not high enough for the RS to relay the traffic received from the MSs to the BS. On the other hand, if the system allocates too much bandwidth to the UL relay zone, not only the redundant capacity in the relay zone is wasted, but also the capacity in the access zone is decreased. Therefore, the value of $K_{Relay}$ is a critical factor on the overall system performance of a relay network and has to be carefully designed.

From the analytical models developed in the previous two subsections, we note that the outage probability in the relay zone is determined by the bandwidth allocated to the relay link $K_{Relay}$ and the distributions of voice users and data users associated with the MS-to-RS access zone $\{C_v^{MR}, C_d^{MR}\}$, which depend on the blocking probability. On the other hand, the blocking probability in the access zone is determined by the offered traffic $\{A_v, A_d\}$, and the bandwidth allocated to the access link $K_{Access} = K - K_{Relay}$, which depends on the outage probability. Therefore, the calculations of the blocking probability in the access zone and the outage probability in the relay zone are tightly coupled. In this subsection, we propose a joint algorithm to determine the optimal value of $K_{Relay}$ and obtain the UL Erlang capacity of an 802.16j MMR network.

The design objective of the proposed algorithm is to minimize the value of $K_{Relay}$ while maintaining a prescribed target outage probability in the UL relay zone, so that most of the bandwidth can be allocated to the UL access zone to accommodate more user traffic. The proposed algorithm consists of iterations of two steps. In the first step, given the traffic loads of voice users and data users $\{A_v, A_d\}$, the blocking probabilities in the access zone $\{P_{block}^v, P_{block}^d\}$ are obtained such that the outage probability in the relay zone $P_{outage}$ does not exceed certain threshold. Specifically, the proposed algorithm first sets an initial value of $K_{Relay}$, and calculates the blocking probabilities in the access zone, the user distributions, and the outage probability in the relay zone by using Exp. (14), (15) & (18). Then it checks whether $P_{outage}$ meets its requirement or not at the current value of $K_{Relay}$. If $P_{outage}$ is lower than the threshold, and is decreased otherwise. Unless $P_{outage}$ meets its requirement, $K_{Relay}$ is further increased or decreased. Once $P_{outage}$ is satisfied, the blocking probabilities are recorded under given traffic loads. Then the values of offered traffic $\{A_v, A_d\}$ are modified and the above procedure is repeated. After we have got the blocking probabilities under various traffic loads, the algorithm proceeds to the second step, where the Erlang capacity region is obtained by extracting the maximum values of $\{A_v, A_d\}$ when both requirements of blocking and outage probabilities are satisfied. The flow chart of the joint algorithm is shown in Fig. 2.

**IV. NUMERICAL RESULTS**

In this section, we present some numerical results based on the analytical models developed in Section III.

We consider an infrastructure-based two-hop cellular network, where $n$ RSs are placed uniformly at the same distance from the BS. Each RS is deployed to cover a fraction $\rho$ of the cell radius. Assuming uniform MS distribution, the arrival intensity of voice and data users associated with the MS-to-RS access zone is a fraction of the arrival intensity of voice and data users associated with the MS-to-BS access zone, which is...
equal to $\lambda_{v,d}^{M2R} = \lambda_{v,d}^{M2B} \cdot n \cdot p$. The system parameters are listed in Table II.

We set the target outage probability in the relay zone to be 3%, which is usually considered to be an acceptable QoS requirement. Fig. 3 shows the blocking probability of voice service versus different traffic loads of voice and data service in Erlangs, with 2 RSs deployed per cell. The figure shows that the blocking probability increases monotonically as the traffic load increases. Furthermore, the increase of data traffic will cause the blocking probability to rise more sharply than the increase of voice traffic. This is because data service has a higher bit rate requirement than voice service, thus occupies more system bandwidth than voice service does.

A two dimensional Erlang capacity region at 3% blocking probability is depicted in Fig. 4 with no RS, 1 RS, 2 RSs, and 3 RSs deployed per cell. The ‘no RS’ curve corresponds to the performance of conventional 802.16e UL, while the remaining three curves characterize the performance of the 802.16j UL. The figure shows that a significant higher user load can be accommodated by the relay-enhanced 802.16j networks, and the system Erlang capacity region increases as the number of RSs increases. The higher capacity gain with increasing number of RSs is due to the reason that with uniform user distribution, more users that are located in the edge of a cell area will be served by the RS with high link spectrum efficiency when the number of RSs deployed per cell increases, therefore significantly enlarge the system Erlang capacity region. However, the actual number of RSs to be deployed in a cell and the placement of RSs have to be carefully designed to guarantee a cost-effective RS solution. To capture the trade-off between the network capacity gain and the cost with RS deployment, a cost-benefit analysis of the relay deployment will be performed in future work.

V. CONCLUSIONS

In this paper, we have investigated the uplink capacity of a two-hop relay-enhanced 802.16j network with adaptive modulation and coding scheme applied in the PHY layer. Two types of service, voice service and data service, are supported by the system with a given quality of service. Based on the new frame structure introduced in 802.16j, we have developed analytical models to calculate the blocking probability in the access zone and the outage probability in the relay zone. Since the calculation of the blocking and the outage probabilities are tightly coupled, we proposed a joint algorithm to determine the optimal value of bandwidth distribution between the access zone and the relay zone, as well as to obtain the Erlang capacity that the system can support when both the blocking and the outage probabilities are within certain thresholds. Numerical results have shown that the UL Erlang capacity region for both voice and data traffic increases significantly when RSs are deployed in the system.

REFERENCES