Fairness-Aware Radio Resource Management in Downlink OFDMA Cellular Relay Networks

Mohamed Salem, Abdulkareem Adinoyi, Mahmudur Rahman, Halim Yanikomeroglu,
David Falconer, and Young-Doo Kim

***Abstract*— Relaying and orthogonal frequency division multiple access (OFDMA) are the accepted technologies for emerging wireless communications standards. The activities in many wireless standardization bodies and forums, for example IEEE 802.16 j/m and LTE-Advanced, attest to this fact. The availability or lack thereof of efficient radio resource management (RRM) could make or mar the opportunities in these networks. Although distributed schemes are more attractive, it is essential to seek outstanding performance benchmarks to which various decentralized schemes can be compared. Therefore, this paper provides a comprehensive centralized RRM algorithm for downlink OFDMA cellular fixed relay networks in a way to ensure user fairness with minimal impact on network throughput. In contrast, it has been observed that pure opportunistic schemes and fairness-aware schemes relying solely on achievable and allocated capacities may not attain the desired fairness, e.g., proportional fair scheduling. The proposed scheme is queue-aware and performs three functions jointly; dynamic routing, fair scheduling, and load balancing among cell nodes. We show that the proposed centralized scheme is different from the traditional centralized schemes in terms of the substantial savings in complexity and feedback overhead.**

***Index Terms*—RRM, OFDMA, relaying, routing, scheduling, fairness, load balancing, proportional fairness.**

I. INTRODUCTION AND MOTIVATION

ORTHOGONAL frequency division multiple access (OFDMA) is the envisioned air-interface for 4G and beyond wireless networks mainly due to its robustness to frequency selective multipath fading, and the flexibility it offers in radio resource allocation [1]. However, in order to truly realize ubiquitous coverage, the high data rate opportunity in OFDMA schemes has to reach to user terminals (UTs) in the most difficult channel conditions, for example, cell edge UTs. Therefore, relaying techniques have been earmarked as the best option to address this problem since relay stations (RSs), with less functionality than a base station (BS), can forward high data rates to remote areas of the cell, and thus overcome the high path losses, while maintaining low infrastructure cost [2]. Hence, the future network roll-out is expected to include various forms of relays. We consider networks enhanced with fixed digital relays deployed by service providers in strategic locations.

The combination of relaying and OFDMA techniques has the potential to provide high data rate to UTs everywhere, anytime. In contrast, conventional opportunistic schedulers will rarely serve UTs with bad channel conditions such as cell edge UTs; this defeats the notion of ubiquitous coverage targeted in future networks, and exposes the importance of fair RRM algorithms to facilitate location-independent service, especially when users subscribed to the same service class are charged similarly regardless of their channel conditions.

In the well-established literature of conventional cellular networks, several queue/traffic-aware fair scheduling algorithms have been proposed such as the channel state dependent packet scheduling (CSDPS), the channel independent packet fair queueing (CIF-Q) [3], and the OFDMA-based algorithms in [4]-[6]. However, such algorithms can not be directly applied to relay-enhanced networks since the problem is not just a scheduling problem. Rather, it is in principle, a joint routing and scheduling problem. In addition, the desired user fairness may not be attained through the fairness-aware schemes that rely solely on achievable and allocated capacities, e.g., proportional fair scheduling (PFS) [7], [8]. The relay-based RRM algorithms developed for single-cell system models along with their performance results are not applicable to multi-cell scenarios since inter-cell interference is not considered, e.g., [9]-[11].

An observed tendency in the literature is to maximize the total cell capacity, e.g., [12] and [13] whereas capacity does not map directly to throughput due to the burst traffic. As such, allocating fair shares of the cell capacity might not result in actual throughput fairness. Furthermore, the vast majority of RRM schemes presented in the literature decouple in-cell routing and resource allocation for simplicity. As such, limiting the opportunities in spatial diversity and channel dynamism the scheme could exploit. In fact, pathloss-based and distance-based relay selection are common and simple strategies, e.g., [14] and [15]. Finally, by oversimplifying channel models, a transmission scheme selection algorithm is proposed in [15] where selection and resource allocation are solely based on the number of required subcarriers. However, in order to exploit the multiuser and frequency diversities, an

---

Manuscript received November 20, 2008; revised July 14, 2009 and December 18, 2009; accepted January 14, 2010. The associate editor coordinating the review of this paper and approving it for publication is W. Lou.

M. Salem, M. Rahman, H. Yanikomeroglu, and D. Falconer are with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada. A. Adinoyi is currently with Swedtel Arabia (e-mail: {mrashad, adinoyi, mmrahman, halim, ddf}@sce.carleton.ca).

Y.-D. Kim is currently with SK Telecom, Korea (e-mail: young-doo.kim@gmail.com).

This work was done while A. Adinoyi and Y.-D. Kim were affiliated with Carleton University and Samsung Electronics, SAIT, Korea, respectively.

This work was presented in part at the IEEE International Conference on Communications 2009. International, US, and Korean patent fillings have been made by Samsung Electronics Co. Ltd., SAIT, Korea.

Digital Object Identifier 10.1109/TWC.2010.05.081548

RRM scheme has to cope with the channel variations.

Therefore, it is imperative to devise intelligent, dynamic, and fair RRM schemes to harness the potentials in OFDMA-based relay networks. In [16], a centralized joint scheduling and routing algorithm employing indefinite number of hops is proposed for a single-carrier relay network, based on CDMA (Evolution-Data Optimized). Therein, several necessary constraints are imposed for operability and reduction of complexity but preclude implementation in multi-carrier systems.

In this paper, we propose a novel formulation with a novel low-complexity centralized algorithm that achieves a ubiquitous coverage, high degree of user fairness and enables intra-cell load balancing in downlink OFDMA-based multi-cell fixed relay networks. The proposed scheme utilizes the opportunities provided in channel dynamism, spatial, and queue and traffic diversities. We show that the scheme provides an efficient tradeoff between network throughput and fairness to all UTs, even to those at the cell edge. We demonstrate the learning ability of the dynamic routing strategy. We also show how substantial savings in complexity and feedback overhead can be attained distinguishing the proposed scheme from traditional centralized schemes. To the best of the authors' knowledge, this contribution is unique among the works presented so far in the literature.

II. SYSTEM DESCRIPTION AND ASSUMPTIONS

In the multi-cellular network, the BS serves K UTs either directly or through M RSs in a cell. All resources are available in each cell resulting in aggressive resource reuse. The total bandwidth is divided into N subchannels, each composed of a set of adjacent OFDM data subcarriers[1]. The serving BS and each of the M RSs in a cell are equipped with K user-buffers. User packets arrive at the corresponding BS buffer according to the traffic model. The channel fading is assumed to be time-invariant within a frame duration. We first consider a generic scenario that is not restricted to a specific geographical deployment of RSs. Thus, potentially, any UT can be connected to any combination of the M RSs yet in only two hops as RSs are not allowed to exchange user data. Such unconstrained relay selection or 'open routing' exposes the ability of our routing strategy to dynamically settle for the best route(s) for each UT given an arbitrary relay deployment. We also present a constrained mode of operation for the routing strategy where geographical relay deployment can be exploited offering substantial savings in feedback overhead. In the proposed scheme, a UT can receive from a group of nodes (BS and/or RSs), and any node can transmit as well to multiple destinations, simultaneously, on different orthogonal subchannels. In addition, any RS is assumed to have the ability to receive and transmit concurrently on orthogonal subchannels. A practical concern might arise if orthogonal transmit and receive subchannels happen to be close in frequency band. Since RSs are fixed, they can be deployed with two antennas (if necessary, at different elevations); a directional antenna for the feeder link from the BS and an omni-directional antenna to the UTs, thus, alleviating such concern.

Load balancing is usually incorporated with the connection admission control mechanisms in conventional cellular networks and it refers to the hand over (hand-off) of some UTs between adjacent cells to distribute the traffic load among BSs network-wide while maintaining users' quality of service (QoS). Although this load balancing function will be an integral part of any prospective RRM scheme, in the literature of OFDMA-based relay networks, researchers often associate the term "load balancing" with a different function which aims at distributing the load evenly among the cell nodes. The number of OFDM subcarriers handled by a node is often employed in literature as a good estimate of its traffic load [9], [17]. As such, an even distribution of subcarriers balances the load among the nodes cell-wide [9], [18]. Although the scheme in [19] aims at achieving the conventional load balancing among cells using boundary RSs, it employs the number of subchannels as a measure of the traffic load. A balanced traffic load reduces the packet processing delays at the regenerative relays. Moreover, load balancing results in the so called 'relay fairness'; a fair utilization of the energy sources of the RSs if the network employs battery/solar-powered RSs [20]. The following section describes the proposed scheme in details and explains how the load balancing function is integrated.

III. THE BS'S JOINT ROUTING AND FAIR SCHEDULING

The objective is to maximize the total cell throughput while maintaining throughput fairness among users. The idea is to operate a throughput-optimal scheduling policy, that stabilizes user queues at all nodes, in a system that receives equal inelastic mean arrival rates at only one source node in the cell which is the BS, using two hops at most. Therefore, the fair behavior of such policy is a special case due to our cellular network system model where we consider that all users belong to the same service class and thus have the same mean arrival rates and the same QoS requirements. Such policy is perceived fair given a similar scenario in [16]. In [21], a congestion control mechanism is proposed with such policy employed to introduce user fairness, through traffic policing, if the arrival rates are elastic, i.e., the traffic sources can adapt their rates. Otherwise, the authors perceive throughput-optimal scheduling more adequate for inelastic traffic.

Let us define the 'demand' metric for any node$_m$-UT$_k$ link on subchannel n as the product of the achievable rate on that access link and the queue length of the user's buffer at that node, as follows

$$D_{n,m\rightarrow k} = R_{m,k,n} Q_k^m, \; m = 0, 1, \ldots, M, \tag{1}$$

whereas the demand of any BS-RS$_m$ feeder link on subchannel n incorporates the queues at the BS (node 0) and those at RS$_m$ and can be expressed as

$$D_{n,0\rightarrow m} = R_{0,m,n} \max_k \{(Q_k^0 - Q_k^m)^+\}, \; m \neq 0. \tag{2}$$

The function $(.)^+$ sets negative arguments to zero. Q_k^m is the queue length of UT$_k$ at node m in bits, bytes, or packets of equal length (shown in blue bars in Fig. 1). Whereas $R_{m,k,n}$ and $R_{0,m,n}$ are the achievable rates on the links node$_m$-UT$_k$

---

[1] The number of OFDMA subcarriers comprising a subchannel is such that its bandwidth is less than the expected coherence bandwidth of the channel.

and BS-RS$_m$, respectively, on subchannel n. These rates are calculated, without loss of generality, using the continuous rate formula for adaptive modulation and coding (AMC) given as $R_{i,j,n} = W \log_2\left(1+\frac{-1.5\beta_{i,j,n}}{\ln(5P_e)}\right)$ where $\beta_{i,j,n}$ is the the received signal-to-interference-plus-noise ratio (SINR) from source i at destination j on subchannel n considering all the dominant interference observed in the previous transmission. P_e and W are the target bit error rate and the OFDM subchannel bandwidth, respectively. As an alternative, either Shannon capacity formula (possibly with some practical SINR gap or penalty) or a discrete AMC lookup table can be used.

Modulated versions of this metric are used in non-relaying OFDMA [5] and SDMA/TDMA [6] networks. Although our results show outstanding performance employing the earlier metric definition, designing the mathematical structure of the metric is an interesting problem by itself, since different emphasizes can be imposed on the rate and link weight arguments. Nevertheless, it can be easily shown that any monotonically increasing function of the metric, in its composite form, will result in the same radio resource allocation (RRA).

A. Mathematical Formulation of the RRA at the BS

In order to maximize the total cell throughput while stabilizing user queues at all nodes, the RRA scheme needs to assign the subchannels with the highest capacities at any node to the outstanding queues at that node. This can be achieved by optimizing the assignment of subchannels to all links and the assignment of user buffers to feeder links so that the sum-demand is maximized at each allocation instant. The resource allocation at the BS can be formulated as a binary integer linear programming (BILP) problem as

$$\max_{\boldsymbol{\rho},\boldsymbol{\gamma}} \Bigg\{ \sum_{n=1}^{N}\sum_{m=0}^{M}\sum_{k=1}^{K} \rho_{m,k,n} R_{m,k,n} Q_k^m + \sum_{n=1}^{N}\sum_{m=1}^{M} \gamma_{0,m,n} R_{0,m,n} \max_k\{(Q_k^0 - Q_k^m)^+\} \Bigg\}, \quad (3)$$

subject to the constraints

$$\rho_{m,k,n} \in \{0,1\}, \forall (m,k,n), \;\; \gamma_{0,m,n} \in \{0,1\}, \forall (m,n), \quad (4)$$

$$\sum_{m=0}^{M}\sum_{k=1}^{K} \rho_{m,k,n} + \sum_{m=1}^{M} \gamma_{0,m,n} \le 1, \;\; \forall\, n, \quad (5)$$

$$\sum_{n=1}^{N}\sum_{k=1}^{K} \rho_{0,k,n} + \sum_{n=1}^{N}\sum_{m=1}^{M} \gamma_{0,m,n} \ge \mu,$$
$$\sum_{n=1}^{N}\sum_{k=1}^{K} \rho_{m,k,n} \ge \mu, \qquad \forall m \ne 0, \quad (6)$$

$$T\sum_{n=1}^{N}\left(\rho_{0,k,n} R_{0,k,n} + \sum_{m=1}^{M} \gamma_{0,m,n} R_{0,m,n} \kappa_k^m\right) \le Q_k^0, \forall k,$$
$$T\sum_{n=1}^{N} \rho_{m,k,n} R_{m,k,n} \le Q_k^m, \;\; \forall (m,k), \;\; m \ne 0. \quad (7)$$

In the above, $\rho_{m,k,n}$ is the k^{th} UT binary assignment variable to the m^{th} node, $m = 0, 1, 2, \ldots, M$, on the n^{th} subchannel ($m = 0$ corresponds to BS, and the rest correspond to relays). The variable $\gamma_{0,m,n}$ is the m^{th} relay binary assignment variable to the BS node on the n^{th} subchannel whereas T is the transmission time of the downlink frame and $\mu = \lfloor N/(M+1) \rfloor$ is the minimum number of subchannels to be assigned to any node (BS or RS), assuming for now uniform user distribution with respect to relay deploymet. The binary indicator κ_k^m is 1 if user k has the highest queue difference between the BS and RS$_m$, and 0 otherwise. The constraints in (4) forces the optimization variables to binary values while the constraints in (5) ensure that at most one link is active per subchannel. The constraints in (6) guarantee even distribution of subchannels among all nodes and hence balances the load. Finally, the constraints in (7), unlike the majority of works in the literature, e.g., [9]-[11] and [15], ensure efficient bit-loading and prevent scheduling errors which could occur if the total capacity of the links withdrawing from a particular buffer is greater than the queue length at that buffer. Therefore, solving the optimization problem in such a novel formulation, results in the joint routing and fair scheduling, guarantees efficient use of resources, and balances the load among cell nodes. A discussion on the routing strategy will follow in the next subsection. The unique aspects of the problem formulation leading to the outstanding performance of the proposed scheme are summarized as follows:

- No explicit non-linear fairness constraints or functions are imposed and thus a single linear objective function is maximized towards achieving a remarkable combination of both high ubiquitous throughput and user fairness, under the system model considered.
- The formulation does not imply any kind of preset routes, user partitioning, or resource partitioning, which are known to be suboptimal simplifying techniques.
- Dynamic routing and scheduling are performed jointly using the ‘differential backlog’ represented by the queue-length difference between BS and RSs [22]; this is analogous to the hydrostatic pressure between fluid tanks connected with pipes of different capacities, which are controlled by the on-off assignment variables, while UTs represent the relevant sinks of individual user flows.
- Traffic diversity (statistical multiplexing) is exploited through incorporating the buffer states; this does not require knowledge of the arrival process statistics.
- Load balancing between relay nodes is achieved jointly as well, as in [20], and not by rearranging the optimal allocation, e.g., [9].

The computational complexity, however, of such three-dimensional BILP problem is non-polynomial in time and can be approximated to $\mathcal{O}(((M+1)K)^N)$. As such, the complexity might reach prohibitive limits in a system with high density of UTs and RSs given the expected high number of subchannels. Therefore, in the next subsection, we propose a low-complexity iterative algorithm that virtually updates the buffer states between iterations while satisfying all of the aforementioned constraints.

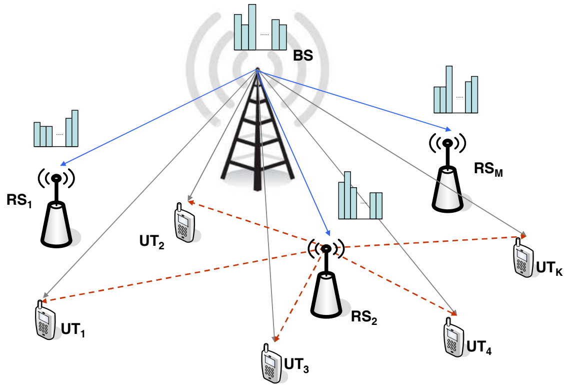
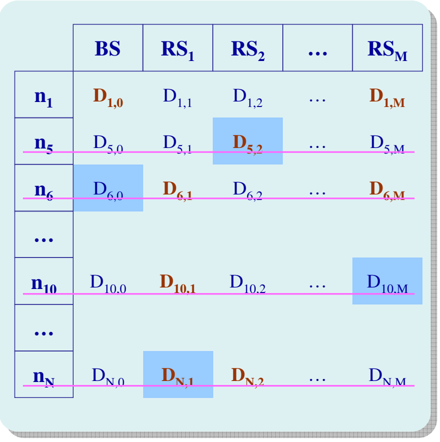


Fig. 1. Example partial network of BS and relays showing a snap shot of user queues and the potential links of the BS and RS$_2$ on subchannel n.



	BS	**RS$_1$**	**RS$_2$**	**…**	**RS$_M$**
n$_1$	**$D_{1,0}$**	$D_{1,1}$	$D_{1,2}$	…	**$D_{1,M}$**
~~**n$_5$**~~	~~$D_{5,0}$~~	~~$D_{5,1}$~~	~~**$D_{5,2}$**~~	~~…~~	~~$D_{5,M}$~~
~~**n$_6$**~~	~~$D_{6,0}$~~	~~**$D_{6,1}$**~~	~~$D_{6,2}$~~	~~…~~	~~**$D_{6,M}$**~~
…					
~~**n$_{10}$**~~	~~$D_{10,0}$~~	~~**$D_{10,1}$**~~	~~$D_{10,2}$~~	~~…~~	~~$D_{10,M}$~~
…					
~~**n$_N$**~~	~~$D_{N,0}$~~	~~**$D_{N,1}$**~~	~~**$D_{N,2}$**~~	~~…~~	~~$D_{N,M}$~~

Fig. 2. The demand matrix during one iteration. Rows with assigned entries are crossed out and eliminated. Bold red entries reflect on the queue updates due to the previous iteration.

B. A Low-complexity Iterative RRA Algorithm

The formulated problem can be viewed as a three-dimensional assignment problem in which even subchannel-to-node assignments are required. The Hungarian algorithm [23] is an efficient solver, with polynomial complexity, for similar two-dimensional assignment problems and thus has been used in different scheduling algorithms for non-relaying networks, e.g., [4], [24], and [25]. Before we discuss how the Hungarian algorithm is applied to our assignment problem, we highlight the following facts: 1) Applying the Hungarian algorithm to an N-by-$M+1$ profit matrix results in the optimal one-to-one assignment. 2) If all the N jobs ($> M+1$) are required to be assigned such that each worker (node) handles almost the same number of jobs (load balancing) while his assignments are interdependent, due to time (or buffer size) constraints, a close-to-optimal solution is attained by applying the Hungarian algorithm $N/(M+1)$ times (iterations) while eliminating the assigned jobs (subchannels). Note that each iteration is solved optimally. The algorithm executes the following steps each allocation instant:

1) The demand metric of each RS$_m$ on subchannel n is calculated as the maximum of K potential links as

$$D_{n,m} = \max_k \{R_{m,k,n} Q_k^m\}, \quad m = 1, 2, \ldots, M. \tag{8}$$

Thus, $D_{n,m}$ is the best proposal of RS$_m$ to use subchannel n while the UT associated with that maximum is marked as the candidate receiver. The demand metric for the BS node is the maximum metric of $M+K$ potential links and is expressed as

$$D_{n,0} = \max_j \{D_{n,0\to j}\}, \tag{9}$$

where $D_{n,0\to j}$ is calculated using (1) and (2), and j denotes any of the potential destinations. Thus, $D_{n,0}$ is the best proposal of the BS to use subchannel n. The destination associated with that proposal is marked as the candidate receiver. Note that if the destination is a RS, the UT that achieved the highest queue difference on that link is marked as well.

2) After calculating the $(M+1)$ demand metrics on each subchannel, the algorithm solves a one-to-one optimization problem to maximize the total demand by applying the Hungarian algorithm to the $N \times (M+1)$ demand matrix $[D_{n,m}]$ (see Fig. 2).

3) The algorithm virtually updates the affected UTs' queues according to the decisions of the previous iteration:

$$Q_k^{m^{(\imath+1)}} = (Q_k^{m^{(\imath)}} - \lfloor R_m^{(\imath)} T \rfloor)^+ . \tag{10}$$

In the above, $Q_k^{m^{(\imath)}}$ is the input queue length to iteration $\imath$ and $R_m^{(\imath)}$ is the rate of the link assigned by the Hungarian algorithm to node m as a result of iteration $\imath$. Note that the queues at destination RSs are not incremented between iterations because the transmissions on all subchannels occur simultaneously and the algorithm has to obey the causality law.

4) The rows with assigned subchannels are eliminated.

5) Steps 1-4 are repeated for the unassigned subchannels until all enqueued packets are scheduled or the subchannels are exhausted.

Due to the one-to-one assignment, each iteration will only assign $M+1$ subchannels to the $M+1$ nodes. As a result, each node is linked to only one destination per iteration; this prevents, along with step 3, the same queue length from being involved in the activation of more than one link as discussed earlier. Furthermore, if $N \bmod (M+1) = 0$, each node will be assigned exactly $N/(M+1)$ subchannels. Hence, load balancing is inherent in the algorithm.

Routing of user data is thus performed dynamically and jointly with its resource allocation. Such dynamic routing strategy uses the maximum differential backlog represented by $\max_k\{Q_k^0 - Q_k^m\}$ to establish the routes. Several works have employed this dynamic routing strategy such as [16]-[22], based on the throughput-optimal link scheduling policy developed in [26] for multihop packet radio networks where routes can comprise indefinite number of hops. However, by 'open routing' in the cellular network we mean that a UT can be connected to any set of RSs while the algorithm is not informed a priori of which RS(s) to use for that UT. Note that a route is comprised of two hops only as RSs are not allowed to exchange user data amongst them. Therefore, in the open routing mode, initial accumulation of the user's data may occur at some RS(s). For instance, let us assume that RS_M in Fig. 1 has a heavily shadowed link to UT_3 while the BS has forwarded some UT_3 data to RS_M, In this situation, these packets neither will be forwarded to UT_3 nor will they be absorbed by a neighboring RS. However, the algorithm exploits the presence of these trapped data, as they reflect on the quality of the second-hop link, by reducing the likelihood of further nominating UT_3 data on BS-RS_M feeder link, irrespective of the channel. That is, while some other user queues at RS_M are being discharged from one iteration to another, UT_3 may no longer achieve the maximum difference $Q_k^0 - Q_k^M$. The algorithm therefore, possesses the ability to learn from the previous forwarding mistakes to improperly selected RSs in previous iterations.

To further demonstrate the feasibility of the dynamic routing and its learning ability, another mode of operation for the dynamic routing strategy named 'constrained routing' is examined and compared to the open routing mode. In that mode, routing constraints are imposed on BS-RS transmissions accounting for the geographical distribution of the RSs and user locations. As such, the dynamic routing is allowed to operate on only $N_{cnst} (\leq M)$ closest RSs to each UT; this is done by ignoring the user buffers at the irrelevant (far) RSs while calculating the differential backlogs[2]. Intuitively, faster routing convergence is expected due to fewer forwarding mistakes. More interestingly, the improvement comes along with substantial savings in feedback overhead due to the eliminated links as discussed in Section VI.

We note that the iterative algorithm is the practical implementation of such joint policy. In a scenario where the distribution of UTs is not uniform with respect to the deployment of RSs, the algorithm should be run in the constrained routing mode. Since the Hungarian algorithm excludes the columns with zero entries (including those RSs with empty buffers), the one-to-one assignment will attempt to achieve the load balancing among only the active RSs (where UTs are clustered).

The computational complexity has been significantly reduced, compared to the BILP problem, using the iterative algorithm since each iteration has a polynomial complexity of $\mathcal{O}(N_u^3)$, where N_u is the number of unassigned subchannels which is usually greater than $M+1$. Given that $M+1$ subchannels are eliminated each iteration, the complexity of iterations rapidly decreases in cubic polynomial manner as N_u decreases. Since the total complexity of step 1 is $\mathcal{O}(MK)$, the complexity of the whole algorithm is loosely upper-bounded by $\mathcal{O}(\frac{N^4}{M+1})$, $M+1 \leq N$. A more precise complexity estimate is down to $\mathcal{O}(\frac{N^2(N+M+1)^2}{4(M+1)})$, $M+1 \leq N$; that approximation holds for reasonable K satisfying $MK \ll N^2$. Unlike the majority of algorithms, the complexity decreases as M, the number of RSs, increases. For the limiting case $M+1 = N$, both estimates coincide at the asymptote $\mathcal{O}(N^3)$ which implies the optimal one-shot Hungarian solution. For further illustration, a pseudocode for the algorithm follows

A Pseudocode for the Iterative Algorithm

```
Initialization: $\mathcal{U} = \mathcal{N}$
while $\|\mathcal{U}\| \neq 0$ and $\sum \mathbf{Q^m} \neq \mathbf{0}$
  for each $n \in \mathcal{U}$
    for $m = 1$ to $M$
      $D_{n,m} = \max_k\{R_{m,k,n}\ Q_k^m\}$
      $\kappa_{n,m} = \arg\max_k\{R_{m,k,n}\ Q_k^m\}$
      $D_{n,0\rightarrow m} = R_{0,m,n} \max_k\{(Q_k^0 - Q_k^m)^+\}, k \in \mathcal{K} - \mathcal{K}_{\text{cnst}}^m$
      $\kappa^m = \arg\max_k\{Q_k^0 - Q_k^m\}$
    end for
      $D_{n,0\rightarrow k} = R_{0,k,n} Q_k^0$
      $D_{n,0} = \max_j\{D_{n,0\rightarrow j}\}, \quad j \in \mathcal{K} \bigcup \mathcal{M}$
      $\kappa_{n,0} = \arg\max_j\{D_{n,0\rightarrow j}\}$
  end for
    % $\mathbf{D} = [D_{n,m}]$ is the demand matrix
    $(\hat{\mathbf{n}}, \hat{\mathbf{m}}) \Longleftarrow$ Hungarian($\mathbf{D}$) % Vectors of indices
    $\mathcal{U} = \mathcal{U} - \{\hat{\mathbf{n}}\}, N_{assigned} = \|\hat{\mathbf{n}}\| = \|\hat{\mathbf{m}}\|$
    %$N_{assigned} \leq \min\{M+1, \|\mathcal{U}\|\}$
  for $i = 1$ to $N_{assigned}$
    $\hat{n} = \hat{\mathbf{n}}(i), \hat{m} = \hat{\mathbf{m}}(i), \hat{r} = \kappa_{\hat{n},\hat{m}}$
      if $\hat{r} \in \mathcal{M}$ then
        $\hat{k} = \kappa^{\hat{r}}$,
        $Q_{\hat{k}}^0 = (Q_{\hat{k}}^0 - \lfloor R_{0,\hat{k},\hat{n}}\ T \rfloor)^+$
      else
        $Q_{\hat{k}}^{\hat{m}} = (Q_{\hat{k}}^{\hat{m}} - \lfloor R_{\hat{m},\hat{k},\hat{n}}\ T \rfloor)^+$
      end if
  end for
end while
```

where $\mathcal{U}$, $\mathcal{N}$, $\mathcal{K}$, $\mathcal{M}$, and $\mathcal{K}_{\text{cnst}}^m$ denote the sets of unassigned subchannels, all subchannels, UTs, RSs, and buffers ignored under constrained routing at RS_m, respectively.

IV. PFS-BASED RRM IN OFDMA RELAY NETWORKS

In this section we consider the Proportional Fair Scheduling (PFS) concept [7], [27]. PFS is known in literature to provide an efficient throughput-fairness tradeoff for conventional (non-relaying) networks. Therefore, integrating PFS with the most commonly adopted RRM techniques in the literature of OFDMA relay networks represents a reasonable reference scheme. The generic framework is to partition the UTs into clusters around the chosen serving nodes (BS and RSs), partition resources among nodes accordingly, then allow individual nodes to perform PFS as adopted in OFDMA-based systems [13], [28]. Some essential details of the PFS-based scheme are shown in Fig. 3 and summarized as follows:

- The closest serving node is chosen by the UT. As such, the direct BS transmission to K_0 UTs occurs within a radius of half the distance between the BS and RSs while $K_r = \sum_{m=1}^{M} K_m$ UTs are relayed, $K_0 + K_r = K$.
- Based on such connections, the BS reserves $N_0 = N\frac{K}{K_0+2K_r}$ subchannels to allocate among its direct UTs and feeder links. The remaining $N - N_0$ subchannels are partitioned among the RSs in proportion to the numbers of their connected UTs. The total power available at each

---

[2] This selection of relay sets is adopted for simplicity. A more appropriate selection could be based on pathloss rather than distance only.

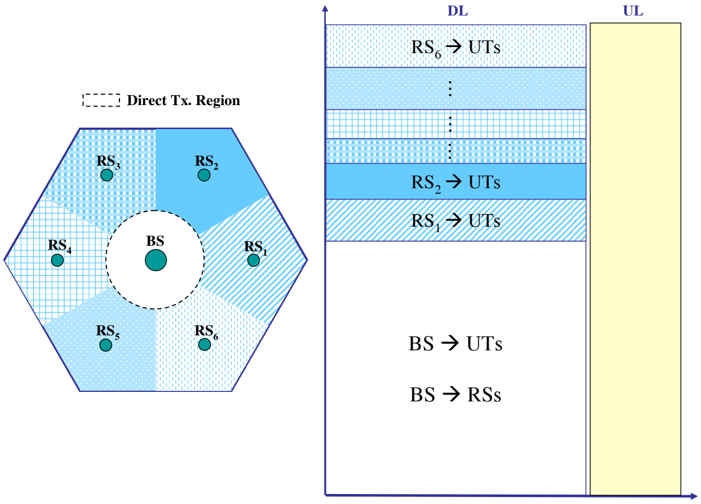


Fig. 3. User and frequency partitioning in the relay-aided PFS scheme.

node is divided equally among the subchannels of its allocated partition.

- The PFS at each node updates the average rates after allocating a subband of the available subchannels. The number of the subbands S is a parameter in the implementation of PFS in OFDMA-based systems [29], [30], and affects the choice of the averaging window size T_p in (11). This is a low-complexity implementation of PFS in multicarrier systems [8].
- At each RS$_m$, subchannel n is assigned to user $k^* = \arg\max_k \frac{R_{m,k,n}}{\overline{R_k}(\imath-1)}$, $k \in \mathcal{K}_m^a$, where $\mathcal{K}_m^a$ is the set of active UTs (with buffered data at RS$_m$), $\overline{R_k}(\imath)$ is the exponentially weighted average rate of user k after allocating the current subband and defined as

$$\overline{R_k}(\imath) = (1-\frac{1}{T_p})\overline{R_k}(\imath-1) + \frac{1}{T_p}\sum_{n\in\mathcal{C}_k} R_{m,k,n}, \qquad (11)$$

  where $\mathcal{C}_k$ is the set of the subchannels assigned to UT$_k$. The relevant user buffer is updated after the subchannel assignment.
- At the BS, subchannel n is assigned to one of the contending destinations (direct UTs and RSs); $j^* = \arg\max_j \frac{R_{0,j,n}}{\overline{R_j}(\imath-1)}$, $j \in \mathcal{K}_0^a \bigcup \mathcal{M}^a$, where $\mathcal{M}^a$ is the set of active RSs in which any RS has at least one connected UT with buffered data at the BS and $\overline{R_j}(\imath)$ is defined as

$$\overline{R_m}(\imath) = (1-\frac{1}{T_p})\overline{R_m}(\imath-1) + \frac{1}{T_p\|\mathcal{K}_{0,m}^a\|}\sum_{n\in\mathcal{C}_m} R_{0,m,n}. \qquad (12)$$

  In the above, $\mathcal{C}_m$ is the set of the subchannels assigned to the feeder of RS$_m$. If the subchannel is assigned to the feeder of RS$_m$, the buffered data at the BS of some relayed user $k \in \mathcal{K}_{0,m}^a \subseteq \mathcal{K}_m$ is scheduled on that feeder following a round-robin sequence. The relevant user buffer is updated after the subchannel assignment.

Through the strategy described earlier, some heuristics have integrated PFS into combined relay and OFDMA technologies such as in [31], [28], and the partial proportional fair (PPF) scheduler in [32]. Since the BS node is required to allocate the resources among the direct UTs and the feeder links of the RSs, a priority metric, similar to that in (12), for such feeders to contend with direct UTs has been proposed in [31]. For the relay-enhanced scheme proposed in [28], a potential improvement in proportional fairness sense can be realized through the clustering (routing) criterion of UTs which aims as well at maximizing the proportional fairness metric, $\sum_k \log R_k$.

V. SIMULATED NETWORK PERFORMANCE

A. *Simulation Models and Parameters*

The simulated network and channel parameters are given in Table I.[3] The cellular network consists of 19 non-sectorized hexagonal cells enhanced with 3 or 6 RSs per cell. These relays are placed at a distance of 0.65 of the cell radius from the BS and with a uniform angular spacing. UTs are uniformly distributed within the cell area. Independent Poisson packet arrival processes are assumed at BS queues. The average arrival rate is 632 packets (188 bytes each) per second per UT. The path-loss model used is $PL = 38.4 + A\log_{10}(d)$ where $A = 23.5$ for BS-RS links and $A = 35.0$ for all other links. RSs transmit to UTs with an omni-directional antenna and receives with a highly directive antenna from the BS. Independent lognormal shadowing is assumed for all links but with different standard deviations. Time-frequency correlated Rician fading is assumed for (LOS) BS-RS links while all other (NLOS) links are assumed to experience time-frequency correlated Rayleigh fading.

B. *Simulation Results and Discussion*

Figure 4 shows scatter plots of UT time-average throughput against UT distance from the BS for 6 and 3 RSs with 25 UTs/cell. Each point in the plot represents the time-average throughput (over 100 allocation time frames) for a particular UT within a drop with fixed location and shadowing. The time average is calculated over the downlink frame duration which is 2/3 of the total TDD frame duration. Statistics are collected from 7 cells (the center cell and the surrounding 6 cells) for each of 30 drops. The performance of the proposed algorithm in its open and constrained routing modes and that of the reference PFS scheme are compared. The distance-based conditional mean of user throughput is approximated by fitting curves of the scatter points as a means of averaging out shadowing. A 7^{th}-degree polynomial well captures the mean coverage behavior of the PFS scheme while only a 3^{rd}-degree polynomial is adequate for the proposed scheme. For the proposed scheme, the uniform average throughput across the cell area is clearly evident and demonstrated by the almost flat performance from BS to cell edge. This implies that a fair service and ubiquitous coverage are provided for all users regardless of their locations, channels, and interference conditions. Some throughput gain is further achieved when the algorithm operates with 6 RSs in its more practical constrained routing mode due to better routing convergence.

---

[3] Adopted from the WiMAX Forum based on IEEE 802.16e and 802.16m. The PL model, RS antenna pattern, and BS-RS channel PDP are adopted from the EU's 6^{th} framework project, IST WINNER: www.ist-winner.org.

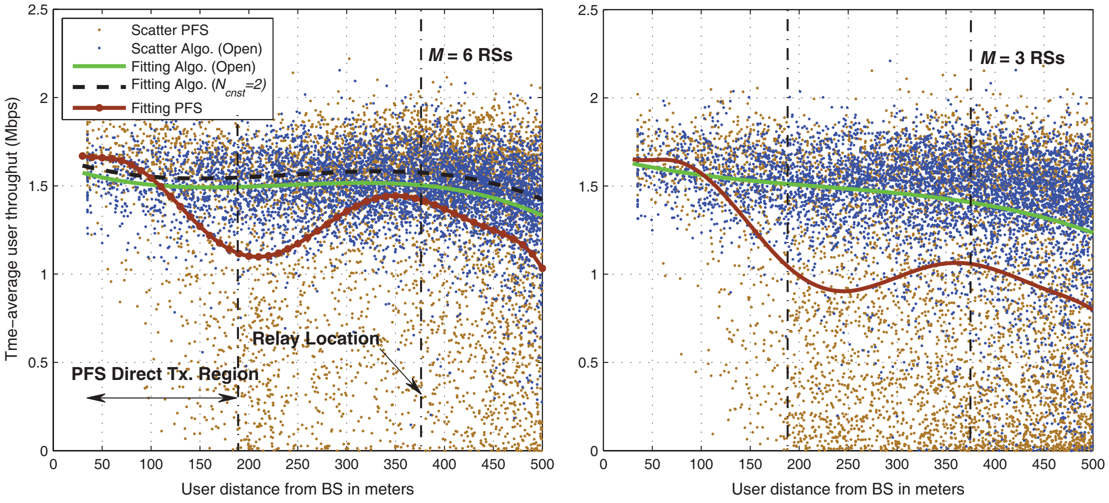


Fig. 4. Time-average user throughput as function of user location and shadowing with 25 UTs/cell using 3 and 6 RSs.

TABLE I
SIMULATION PARAMETERS

Parameter	Value
BS-BS distance	1 Km
RS distance from BS	0.65 × cell radius
UT min. close-in distance to BS	35 m
BS Tx. antenna gain	15 dB
RS Tx. antenna gain	10 dB
RS Rx. antenna $\theta_{3dB} = 20^\circ$	$\pi/9$
UT Rx. antenna gain	0 dB
Shadowing σ for NLOS links	8.9 dB
Shadowing σ, for LOS links (BS-RS)	4 dB
Rician K-factor for BS-RS links	10 dB
Carrier frequency	2.5 GHz
Total bandwidth	20 MHz
UT mobility	20-90 Km/hr
BS-RS links max. Doppler spread	4 Hz
Number of channel taps	6
Number of channel taps (BS-RS)	8
TDD frame length	2 msec
Downlink : Uplink ratio	2:1
DL Tx. time in OFDM data symbols	11 symbols
OFDM subcarrier bandwidth	10.9375 KHz
OFDM symbol duration	102.86 μsec
Subchannel width	18 subcarriers
Total number of subchannels	102
CR-QAM target BER	10^{-3}
Noise power density at Rx. nodes	-174 dBm/Hz
BS total Tx. power P_B	46 dBm
RS total Tx. power P_R	37 dBm
PFS averaging window size T_p	5
PFS number of subbands S	7
PFS radius of direct Tx. region	0.325 × cell radius

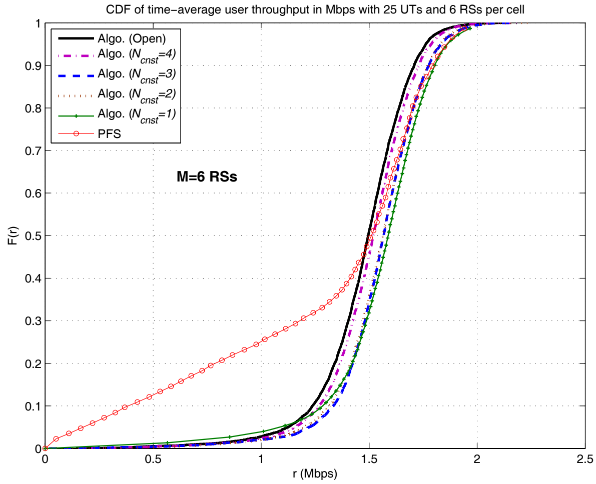


Fig. 5. CDF of the time-average user throughput with 25 users and 6 RSs per cell.

On the contrary, the coverage of the PFS reference scheme is significantly distance dependent as the mean throughput depreciates when users move away from the serving node, especially at the cell edge. That is mainly due the fact that spatial diversity is not exploited (due to static routing) while scheduling a UT on the available subchannels partially exploits the frequency diversity and, moreover, may not overcome large pathloss (e.g., due to heavy shadowing) which dominates all the UT's subchannels. This results in a very poor time-average throughput for such UT (i.e., points at the bottom of the scatter plot). Whereas, the scatter points for the proposed algorithm have high throughput and narrow spread. This indicates the ability of the dynamic routing strategy to find the appropriate path(s) for such UTs and to deliver a fair service. The difference in performance is further emphasized in the scenario with 3 RSs as more users are expected to have poor link qualities from their serving RSs in the PFS scheme. However, the proposed scheme still offers a reasonable ubiquity and substantial throughput gains over the reference scheme, especially at the cell edge.

Note that the traditional PFS is not queue-aware and expected to provide performance inferior to the shown here where the PFS at a serving node excludes the users with empty buffers. Although such practical constraint enables the reference PFS to partially exploit the traffic diversity, it does not prevent resource under-utilization.

Figure 5 shows the CDF of the time-average throughput with both the open and constrained routing modes. Different N_{cnst} are considered. The 5th-percentile throughput of such CDFs is associated with cell-edge users in LTE evaluation

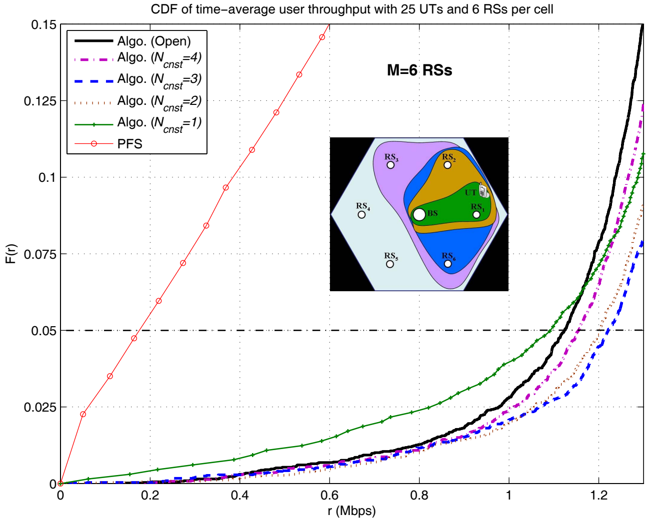


Fig. 6. Lower tail behavior of the CDF of the time-average user throughput with 25 users and 6 RSs per cell.

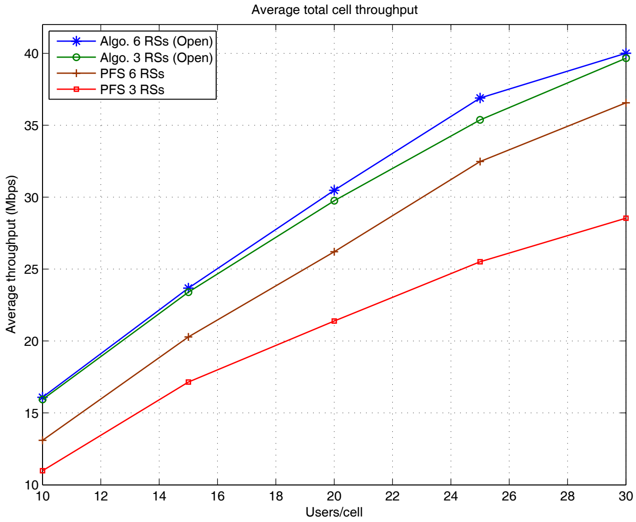


Fig. 7. Total average cell throughput for the proposed and PFS schemes.

methodology [28], [33]. First, a substantial 5th-percentile throughput advantage of 540% is realized for the proposed scheme in its open routing mode as compared to the PFS scheme which possesses a poor lower tail behavior. Second, the performance gaps between different cases of the constrained routing mode and the open mode of the dynamic routing strategy are generally narrow (8.6% at most). Such close performance implies that the open routing mode has an inherent capability of avoiding the poor routes to the UTs using the differential queue length information as discussed earlier in Subsection III-B. However, some throughput losses are inevitable due to initial forwarding mistakes and after occasional improvements of the poor links (due to small-scale fading and/or co-channel interference). Note that the constrained routing mode also utilizes the same learning ability to establish routes using fewer yet better candidate RSs for each UT. This can be observed in Fig. 6 (Fig. 5 with zoom-in).

In Fig. 6, a slight cell-edge throughput improvement of 2.76% is attained by excluding the farthest 2 RSs ($N_{cnst} = 4$) as compared to the open routing mode. Using only the 3 closest RSs ($N_{cnst} = 3$), yields the best improvement (8.6%) as the ‘far’ RSs with potentially poor links to UTs have been excluded along with the associated throughput loss. As expected, further elimination of RSs reduces the spatial diversity the dynamic routing exploits and thus the performance degrades slightly from $N_{cnst} = 3$ to $N_{cnst} = 2$ and significantly at $N_{cnst} = 1$, where only the closest RS is allowed, resulting in a degradation of 2.9% relative to the open mode. As an alternative interpretation of these results; at a target average throughput of 1 Mbps, the outage probability of the proposed scheme ranges between 2.8% in the open mode and 2% in the constrained mode $N_{cnst} = 2$ or 3 as compared to 25.6% with the PFS scheme.

Figure 7 shows the total average cell throughput, as function of the number of UTs per cell, employing 3 and 6 RSs. The behavior in these curves is in agreement with the multiuser diversity concept and emphasizes the ability of the proposed scheme to maximize the total cell throughput by exploiting the multiuser, frequency, spatial, and traffic diversities. We employ the IEEE 802.16m fairness index [34] and Jain’s fairness index [35] to assess the performance of the proposed and the reference schemes in terms of the time-average fairness and the long- and short-term fairness, respectively. Although the two metrics have different mathematical properties, both metrics are symmetric for all user rates, and in both, ideal fair situation will be indicated by exact unity values. This is in line with our system model considering users with same priority.

The IEEE 802.16m fairness index represents the ratio of a user’s throughput rate to the average of K users’ throughput rates and thus takes a value between 0 and K as

$$x_k(t) = \frac{r_k(t)}{\frac{1}{K}\sum_{i=1}^{K} r_i(t)}. \tag{13}$$

The time-average throughput rates, for 15 and 25 UTs with 3 and 6 RSs, are collected from all drops to plot the CDFs shown in Fig. 8. It can be observed that the same outstanding fairness behavior is achieved by the open and constrained routing modes with $M = 6$ at the different loading levels. While it becomes more difficult to maintain fairness as the number of users increases, there is insignificant degradation with the proposed scheme at 25 UTs/cell as opposed to the PFS scheme. Furthermore, reducing the number of RSs to 3 with 15 UTs has almost no impact on the fairness behavior of the proposed scheme and only a slight degradation with 25 UTs/cell; this however does not hold for the PFS scheme.

Jain’s index has been widely used in relevant works, e.g., [4], [28], and [36], and it is recommended by the WiMAX Forum [37] for fairness assessment of proponents’ algorithms. It is defined as

$$x_{w_j} = \frac{\left(\sum_{i=1}^{K} r_{i,w_j}\right)^2}{K\sum_{i=1}^{K} r_{i,w_j}^2}, \tag{14}$$

where r_{i,w_j} is the i^{th} user’s average throughput rate achieved during the j^{th} time window w_j. As such, the index is a positive fraction that is lower-bounded by $1/K$.

Therefore, in Fig. 9, the closer the CDF to a unit step at unity the more long-term fairness the scheme achieves after a time window of 20 frames. Although short- and

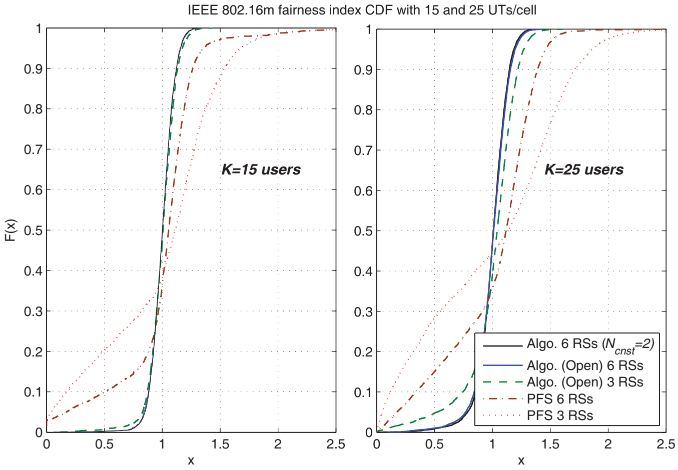


Fig. 8. Time-average fairness using the IEEE 802.16m index with different numbers of users and RSs.

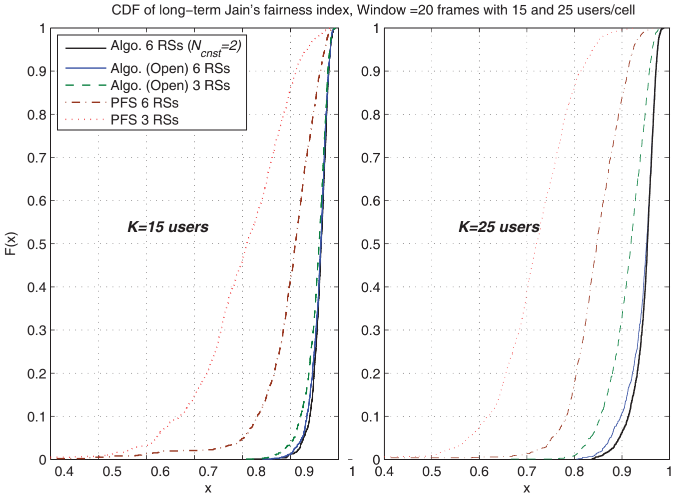


Fig. 9. Long-term fairness using Jain's index with different numbers of users and RSs.

long-term fairness are much more stringent than time-average fairness, the relative behavior observed in Fig. 8 matches that observed in Fig. 9 and Fig. 10 for long- and short-term (instantaneous) fairness, respectively. Generally, queue-awareness allows RRM schemes to compensate the overlooked user buffers, if any, and potentially improve user fairness, at least, in the long term sense. However, PFS relies on metrics based solely on allocated channel capacities. Furthermore, under static routing, as the number of users increases with fewer RSs employed, more UTs links to the serving nodes experience large pathlosses and thus only low achievable rates are left for the PFS to apply its fairness criterion. In contrast, the proposed scheme circumvents the problem of heavy shadowing and/or large pathloss through dynamic in-cell routing. This explains its ability to further improve fairness as time evolves and to exploit the spatial diversity when the number of RSs is increased, yielding such a wide gap in performance as compared to the reference PFS scheme.

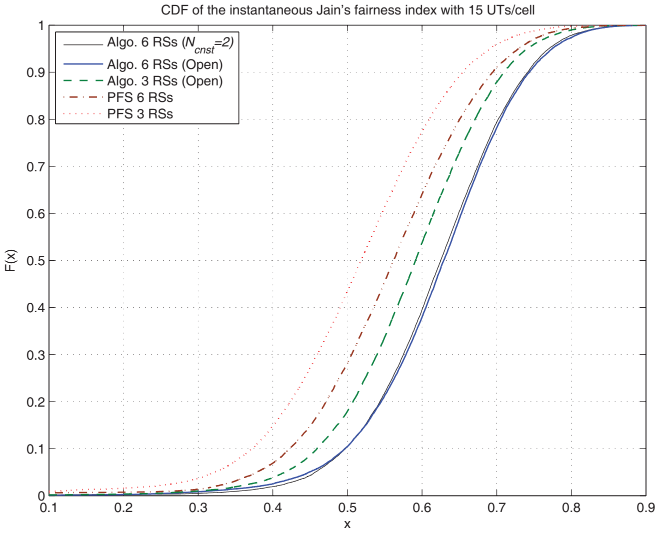


Fig. 10. Jain's instantaneous fairness index with 15 UTs/cell.

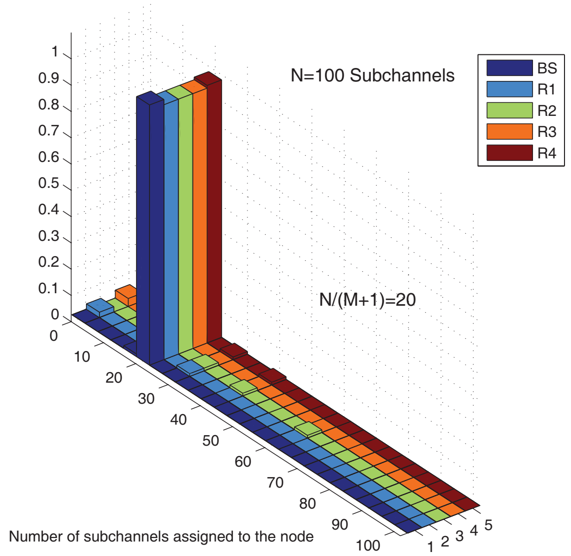


Fig. 11. Normalized histogram of the number of subchannels assigned to transmitting nodes under open routing and uniform distribution of UTs.

It is worth mentioning that such outstanding performance in terms of throughput and fairness is achieved without overloading any node in the system. We demonstrate the intra-cell load-balancing behavior of the proposed scheme through Fig. 11 which shows a normalized histogram for the number of subchannels assigned to each node in a cell of a BS and 4 RSs during a drop of 100 time frames. It can be observed that each node is persistently assigned 20 subchannels, what is equal to $\frac{N}{M+1}$. We note that this is achieved with the exception that if a node has no buffered data, then this node is excluded from the assignment and the load balancing is maintained among the active nodes only. Thus, the very limited perturbations shown deliberately in the figure occur only during the first few initialization frames when RSs start receiving data with empty buffers while resources are assigned among the BS and only the active RSs. received data to forward. In addition to distributing the traffic load among cell nodes and thus reducing the packet processing delays at the regenerative RSs, the load balancing feature also spatially spreads (randomizes) the co-channel interference across the network exploiting the uniform geographical relay deployment.

VI. SIGNALLING OVERHEAD AND DELAYS

In order for an RRM scheme to exploit the multiuser and frequency diversities, the allocation process should be conducted periodically with a period not greater than the shortest channel coherence time which is determined by the highest user mobility supported. Therefore, if the allocation is conducted at the beginning of each TDD frame, the feedback is required that frequent for maximum mobility users, e.g., 90 Km/hr based on our adopted frame duration as per the WiMAX Forum. However, for lower mobility, the feedback can be acquired as less frequently as each $\lfloor T_c/T_F \rfloor$ frames. That is the maximum integer number of TDD frames less than the user's coherence time of the channel (4 TDD frames in the simulated scenario). As such, the RRA algorithm can be invoked that often while the allocation result will be applied to the transmissions of the intermediate frames until the following allocation instant. Such relaxed resource allocation, however, less exploits the traffic diversity for highly burst traffic. We hereby discuss the following items to quantify the amount of feedback information required each allocation instant:

- Implementing the constrained routing mode provides substantial savings in feedback overhead. That is because no feedback is required from the UT for the eliminated RS-UT links.
- In practice, AMC lookup tables are used and therefore reporting the indices of the achievable AMC levels per link significantly saves in signalling overhead as compared to reporting a wide range of continuous SINRs; this applies to both UTs and RSs.
- Our dynamic routing strategy, either in open or constrained mode, allows the UT to be connected to more than one node; having many users per cell, this implies that only very few subchannels are used per each node-UT link. As such, with potentially marginal performance losses, further savings in overhead can be achieved if UTs report only the 'best' fraction of subchannels in term of achievable rates[4]. Although examining the impact of such partial feedback is outside the scope of this paper, auxiliary studies considering only the best 50% of link subchannels show no performance degradation, even when the number of reported links is limited by constrained routing.

The following formula can be used to estimate the feedback overhead per UT in the system taking into account the previous items:

$$T_{OH} = \frac{N_{subch}\ (N_{cnst}+1)\ n_{\text{AMC}}}{T_F\ \lfloor T_c/T_F \rfloor} \quad \text{bps.} \tag{15}$$

In the above, N_{subch} and N_{cnst} denote the number of reported subchannels per link and the number of RSs allowed for the constrained routing, respectively. Whereas n_{AMC} denotes the number of bits used to indicate the index of the achievable AMC mode on a subchannel. We note that if the algorithm needs to allocate all the subchannels to evacuate the system buffers while the UTs provide partial feedback, the parameter N_{subch} can be decreased as the number of admitted UTs increases.

Based on the frame structure, the minimum delay a relayed packet encounters is $T_F+2\tau$ where τ is the OFDM transceiver transmission time. Although the current simulation platform is quite advanced, individual packet delays are intractable. However, as the results show, the algorithm is designed to maximize throughput while stabilizing all queues and avoiding build ups. Hence, once the algorithm converges to the proper routes, it is expected to minimize the queuing delay as a consequence.

VII. INTRA-CELL RESOURCE REUSE

The vast majority of schemes in literature resort, in the first place, to suboptimal techniques in allocating the premium resources to simplify the problem. Among these techniques are static partitioning of users based on their locations, static routing or relay selection, partitioning of resources among different cell regions or nodes, and excluding the traffic and queue status. We observe that employing intra-cell reuse on top of such suboptimal start does not produce an optimal solution for any given objective. Rather, the gains from intra-cell reuse might improve the resource utilization and thus compensate for some of the losses. However, no significant additional gains are expected from further aggressive intra-cell reuse among RSs, beyond the reuse among BS and RSs [13]. On the contrary, by exploiting all the aforementioned degrees of freedom without planning or partitioning of any kind, our proposed scheme does not treat the fundamental resource allocation problem superficially. It seeks, in the first place, a ubiquitous and fair high-data-rate coverage through efficient management of the premium resources.

Moreover, since our proposed scheme does not rely on the geography of the network, the cell region in our scheme can be shrunk to a sector of the cell (with a directional BS antenna and fewer RSs and UTs) without intra-sector reuse as in [28]. Also, the cell can be shrunk to one of the 3 cells forming an eNB in LTE-A architecture. In both cases, resources are reused three times in the original service area providing opportunity for higher spectral efficiencies. This might be infeasible for other schemes that rely on the geographical distribution of RSs. It is however observed that the vast majority of works apply static intra-cell spatial reuse patterns based solely on user locations, e.g., reusing the channels assigned to RS-UT links in the BS-UT links within the close vicinity of the BS [11], [13]. Nevertheless, further increase in our system capacity could be achieved if opportunistic intra-cell reuse is employed utilizing instantaneous channel conditions and antenna directivity [38], rather than the static spatial reuse patterns commonly adopted in literature. In such case, least co-channel interference levels are attained, and under-utilized resources, if any, are used first.

VIII. CONCLUSIONS

Efficient RRM schemes are required to harness the opportunities in the future relay-enhanced OFDMA-based networks in which user fairness is crucial. This paper provides a novel fairness-aware joint routing and scheduling algorithm for such

---

[4] Either fixed number of the best subchannels or every subchannel whose quality is above a certain threshold.

networks in cellular environments. The proposed algorithm exploits the opportunities in the frequency, spatial, and traffic diversities irrespective of the geographical relay deployment. As such, its performance is superior to that of a proportional fair relaying scheme in terms of ubiquitous coverage, cell-edge throughput, short- and long-term user fairness, as well as load balancing. Simulation results prove the learning ability and the efficacy of the dynamic open routing strategy which converges to better routes, even under the challenging uniform relay deployment considered. The dynamic constrained routing is shown to be the practical mode of operation due to the substantial savings in feedback overhead. The inherent load-balancing feature works independently from the traffic load at adjacent cells, results in spatial spreading of the co-channel interference across the network, and minimizes the packet processing delays at the regenerative relays.

REFERENCES

[1] IEEE P802.16j/D1, "Draft IEEE standard for local and metropolitan area networks Part 16: air interface for fixed and mobile broadband wireless access systems: Multihop relay specification," pp. 1002-1007, Aug. 2007.

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

[3] Z. Han and K. J. Liu, *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications.* Cambridge University Press, 2008.

[4] D. Niyato and E. Hossain, "Adaptive fair subcarrier/rate allocation in multirate OFDMA networks: radio link level queuing performance analysis," *IEEE Trans. Veh. Technol.*, vol. 55, no. 6, pp. 1897-1907, Nov. 2006.

[5] P. Parag, S. Bhashyam, and R. Aravind, "A subcarrier allocation algorithm for OFDMA using buffer and channel state information," in *Proc. IEEE Veh. Technol. Conf.*, pp. 622-625, Sep. 2005.

[6] M. Kobayashi and G. Caire, "Joint beamforming and scheduling for a multi-antenna downlink with imperfect transmitter channel knowledge," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1468-1477, Sep. 2007.

[7] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277-1294, June 2002.

[8] H. Kim and Y. Han, "A proportional fair scheduling for multicarrier transmission systems," *IEEE Commun. Lett.*, vol. 9, no. 3, pp. 210-212, Mar. 2005.

[9] 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, Feb. 2007.

[10] R. Kwak and J. M. Cioffi, "Resource-allocation for OFDMA multi-hop relaying downlink systems," in *Proc. IEEE Global Commun. Conf.*, pp. 3225-3229, Nov. 2007.

[11] M. Kaneko and P. Popovski, "Radio resource allocation algorithm for relay-aided cellular OFDMA system," in *Proc. IEEE International Conf. Commun.*, pp. 4831-4836, June 2007.

[12] W. Nam, W. Chang, S.-Y. Chung, and Y. Lee, "Transmit optimization for relay-based cellular OFDMA systems," in *Proc. IEEE International Conf. Commun.*, pp. 5714-5719, June 2007.

[13] Ö. Oyman, "Opportunistic scheduling and spectrum reuse in relay-based cellular OFDMA networks," in *Proc. IEEE Global Commun. Conf.*, pp. 3699-3703, Nov. 2007.

[14] V. Sreng, H. Yanikomeroglu, and D. Falconer, "Relayer selection strategies in cellular networks with peer-to-peer relaying," in *Proc. IEEE Veh. Technol. Conf.*, pp. 1949-1953, Oct. 2003.

[15] J. Lee, S. Park, H. Wang, and D. Hong, "QoS-gurarantee transmission scheme selection for OFDMA multi-hop cellular networks," in *Proc. IEEE International Conf. Commun.*, pp. 4587-4591, June 2007.

[16] H. Viswanathan and S. Mukherjee, "Performance of cellular networks with relays and centralized scheduling," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2318-2328, Sep. 2005.

[17] J. Lee, H. Wang, S. Lim, and D. Hong, "A multi-hop user admission algorithm for fixed relay stations with limited capabilities in OFDMA cellular networks," *IEEE 18th International Symp. Personal, Indoor Mobile Radio Commun.*, Sep. 2007.

[18] T. Kim, T.-Y. Min, and C. Kang, "Opportunistic packet scheduling algorithm for load balancing in a multi-hop relay-enhanced cellular OFDMA-TDD System," *IEEE 14th Asia-Pacific Conf. Commun.*, Oct. 2008.

[19] J. Fan, T. Hui, X. Haibo, and W. Xijun, "A novel relay based load balancing scheme and performance analysis using Markov models," *IEEE Veh. Technol. Conf.*, Apr. 2009.

[20] G. Li and H. Liu, "Resource allocation for OFDMA relay networks with fairness constraints," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 11, pp. 2061-2069, Nov. 2006.

[21] A. Eryilmaz and R. Srikant, "Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control," *IEEE/ACM Trans. Networking*, vol. 15, no. 6, pp. 1333-1344, Dec. 2007.

[22] M. Neely, E. Modiano, and C. Rohrs, "Dynamic power allocation and routing for time-varying wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 89-103, Jan. 2005.

[23] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistic Quarterly,* vol. 2, no. 1, pp. 83-97, 1955.

[24] H. Yin and H. Liu, "An efficient multiuser loading algorithm for OFDM-based broadband wireless systems," in *Proc. IEEE Global Commun. Conf.*, pp. 103-107, Dec. 2000.

[25] M. Rahman and H. Yanikomeroglu, "Interference avoidance through dynamic downlink OFDM subchannel allocations with intercell coordination," in *Proc. IEEE Veh. Technol. Conf.*, pp. 1630-1635, May 2008.

[26] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Automatic Control,* pp. 1936-1948, Dec. 1992.

[27] F. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *J. Operational Research Society*, vol. 49, no. 3, pp. 237-252, Mar. 1998.

[28] W.-G. Ahn and H.-M. Kim, "Proportional fair scheduling in relay enhanced cellular OFDMA systems," in *Proc. IEEE 19th International Symp. Personal, Indoor Mobile Radio Commun.*, pp. 1-4, Sep. 2008.

[29] W. Anchun, X. Liang, Z. Shidong, X. Xibin, and Y. Yan, "Dynamic resource management in the fourth generation wireless systems," in *Proc. International Conf. Commun. Technol.*, pp. 1095-1098, Apr. 2003.

[30] R. Srinivasan and J. Zhuang, "IEEE 802.16m evaluation methodology document (EMD)," IEEE 802.16m-08/004r5, 15 Jan. 2009. [Online]. Available: http:/www.ieee802.org/16/tgm/core.html.

[31] L. Xiao and L. Cuthbert, "Improving fairness in relay-based access networks," in *Proc. 11th International Symp. Modeling, Analysis Simulation Wireless Mobile Syst.*, Oct. 2008.

[32] L. Huang, M. Rong, L. Wang, Y. Xue, and E. Schulz, "Resource scheduling for OFDMA/TDD based relay enhanced cellular networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, pp. 1544-1548, Mar. 2007.

[33] E. Dahlman, H. Ekstrom, A. Furuskar, Y. Jading, J. Karlsson, M. Lundevall, and S. Parkvall, "The 3G long-term evolution—radio interface concepts and performance evaluation," in *Proc. IEEE Veh. Technol. Conf.*, pp. 137-141, May 2006.

[34] IEEE C802.16m-07/306, "Protocol structure to support cooperative transmission," IEEE 802.16 Broadband Wireless Access Working Group. [Online]. Available: http://www.ieee802.org/16/tgm/contrib/C80216m-07_306.pdf.

[35] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation and Modeling.* New York: Wiley, 1991.

[36] M. Moretti and A. Todini, "A resource allocator for the uplink of multi-cell OFDMA systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2807-2812, Aug. 2007.

[37] K. Ramadas and R. Jain, "Mobile WiMAX—part I: a technical overview and performance evaluation," *WiMAX Forum,* Sept. 2007. [Online]. Available: http://www.wimaxforum.org/technology/downloads/

[38] M. Salem, A. Adinoyi, H. Yanikomeroglu, and Y.-D. Kim, "Radio resource management in OFDMA-based cellular networks enhanced with fixed and nomadic relays," to appear in *IEEE Wireless Commun. Netw. Conf.*, Apr. 2010.

Mohamed Rashad Salem (S'06) received his B.Sc. in Communications and Electronics from the department of Electrical Engineering, Alexandria University, Egypt, in 2000. He was nominated and hired as a faculty member in the department of Engineering Mathematics, Alexandria University, from which he received his M.Sc. degree and was promoted to the position of Assistant Lecturer in Feb 2006. Within this period he has gained wide experience in Research & Development and collaboration with industrial parties. In June 2006, Mr. Salem has received a scholarship from the department of Systems and Computer Engineering, Carleton University, Ottawa, Canada, to pursue his Ph.D. Toward that end, he has been conducting research in collaboration with Samsung Electronics on advanced radio resource management in next generation wireless networks. His research interests encompass stochastic modeling, congestion control, and optimization techniques. He has been granted the 2009/2010 Ontario Graduate Scholarship in Science and Technology (OGSST).

Abdulkareem Adinoyi received a B.Eng degree from the University of Ilorin, Nigeria, in 1992, a M.S degree from the King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia, in 1998 and a Ph.D degree from Carleton University, Ottawa, Canada, in 2006. All the degrees are in electrical engineering. From April 1993 to August 1995 Mr. Adinoyi was with Dubi Oil Limited, Port Harcourt, Nigeria as an instrument/electrical engineer. Between September 1995 and October 1998 he was with KFUPM as a research assistant and between January 1999 and August 2002 he held the position of a lecturer at the department of electrical engineering, KFUPM. He was a senior research associate at the Department of Systems and Computer Engineering (SCE) at Carleton University. Between January 2004 and December 2006, he participated in the European Union 6th Framework integrated project - the WINNER. From January 2007 to August 2007, he was briefly at Qassim University, Saudi Arabia where he held the position of an assistant professor of electrical and electronic engineering. Between September 2007 and July 2009, he worked, as project technical manager, on Samsung's advanced radio resource management project for OFDMA-based relay-enhanced broadband wireless networks under the auspices of SCE. He is currently at Swedtel Arabia as consultant for Saudi Telecommunication Company. Dr. Adinoyi's current research interests are in infrastructure-based multihop and relay networks, cooperative diversity schemes and protocols, efficient radio resource management techniques for the next-generation wireless networks, and technology evolution.

Mahmudur Rahman received his B.Sc. degree in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 1991. He obtained an M. Eng. degree in Telecommunications from Asian Institute of Technology (AIT), Bangkok, Thailand and an M.A.Sc. degree in Electrical Engineering from Carleton University, Ottawa, Canada, in 1994 and 2004, respectively.

He worked for Bangladesh Atomic Energy Commission (Bangladesh), Johnson Electric Industrial Ltd. (Thailand), and ACE Electronics Industries Co., Ltd. (Thailand) at different times of his career. He was involved in the World INitiative NEw Wireless Radio (WINNER) research project from 2004 to 2007. He is currently pursuing a Ph.D. degree in Electrical Engineering at Carleton University and working as a Research Engineer in Communications Research Centre (CRC), Ottawa, Canada. His current research interests include radio resource management, multi-hop wireless networks, inter-cell coordination, interference management, and cognitive radio systems.

Halim Yanikomeroglu received a B.Sc. degree in Electrical and Electronics Engineering from the Middle East Technical University, Ankara, Turkey, in 1990, and a M.A.Sc. degree in Electrical Engineering (now ECE) and a Ph.D. degree in Electrical and Computer Engineering from the University of Toronto, Canada, in 1992 and 1998, respectively. He was with the R&D Group of Marconi Kominikasyon A.S., Ankara, Turkey, from 1993 to 1994.

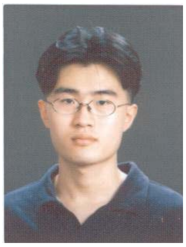
Since 1998 Dr. Yanikomeroglu has been with the Department of Systems and Computer Engineering at Carleton University, Ottawa, where he is now an Associate Professor with tenure. Dr. Yanikomeroglu's research interests cover many aspects of the physical, medium access, and networking layers of wireless communications with a special emphasis on multihop/relay/mesh networks and cooperative communications. Dr. Yanikomeroglu's research is currently funded by Samsung (SAIT, Korea), Huawei (China), Communications Research Centre of Canada (CRC), Research in Motion (Canada), and NSERC. Dr. Yanikomeroglu is a recipient of the Carleton University Research Achievement Award 2009.

Dr. Yanikomeroglu has been involved in the steering committees and technical program committees of numerous international conferences; he has also given 18 tutorials in such conferences. Dr. Yanikomeroglu is a member of the Steering Committee of the IEEE Wireless Communications and Networking Conference (WCNC), and has been involved in the organization of this conference over the years, including serving as the Technical Program Co-Chair of WCNC 2004 and the Technical Program Chair of WCNC 2008. Dr. Yanikomeroglu is the General Co-Chair of the IEEE Vehicular Technology Conference to be held in Ottawa in September 2010 (VTC 2010-Fall).

Dr. Yanikomeroglu was an editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS [2002-2005] and IEEE COMMUNICATIONS SURVEYS & TUTORIALS [2002-2003], and a guest editor for WILEY JOURNAL ON WIRELESS COMMUNICATIONS & MOBILE COMPUTING. He was an Officer of IEEE's Technical Committee on Personal Communications (Chair: 2005-06, Vice-Chair: 2003-04, Secretary: 2001-02), and he was also a member of the IEEE Communications Society's Technical Activities Council (2005-06).

Dr. Yanikomeroglu is an Adjunct Professor at Prince Sultan Advanced Technologies Research Institute (PSATRI) at King Saud University, Riyadh, Saudi Arabia; he is also a registered Professional Engineer in the province of Ontario, Canada.

David Falconer received the B.A. Sc. degree in Engineering Physics from the University of Toronto in 1962, the S.M. and Ph.D. degrees in Electrical Engineering from M.I.T. in 1963 and 1967 respectively, and an honorary doctorate of science from the University of Edinburgh in 2009. After a year as a postdoctoral fellow at the Royal Institute of Technology, Stockholm, Sweden he was with Bell Laboratories from 1967 to 1980 as a member of technical staff and group supervisor. During 1976-77 he was a visiting professor at Linköping University, Linköping, Sweden. Since 1980 he has been with Carleton University, Ottawa, Canada, where he is now Professor Emeritus and Distinguished Research Professor in the Department of Systems and Computer Engineering. His current research interests center around beyond-third-generation broadband wireless communications systems. He was Director of Carleton's Broadband Communications and Wireless Systems (BCWS) Centre from 2000 to 2004. He was the Chair of Working Group 4 (New Radio Interfaces, Relay-Based Systems and Smart Antennas) of the Wireless World Research Forum (WWRF) in 2004 and 2005. He received the 2008 Canadian award for Telecommunications Research, a 2008 IEEE Technical Committee for Wireless Communications Recognition Award, and the IEEE Canada 2009 Fessenden Award (Telecommunications). He is an IEEE Life Fellow.



Young-Doo Kim (S'99–M'06) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 1999, 2001, and 2006, respectively.

He was with Samsung Advanced Institute of Technology (SAIT), Gyeonggi-Do, Korea, from 2006 to 2009. He is currently with SK Telecom, Korea. His primary research interests include synchronization, detection and estimation for communication systems, and predistortion linearization of nonlinear power amplifiers for wireless applications. His current research activities include radio resource management in multihop and relaying networks and interference management in cellular radio systems.