

Generalized Proportionally Fair Scheduling for Multi-User Amplify-and-Forward Relay Networks

Alireza Sharifian*, Petar Djukic*, Halim Yanikomeroglu*, Jietao Zhang†

*Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada

e-mail: {alireza,djukic,halim}@sce.carleton.ca

†Wireless Research Department, Huawei Technologies Co., Ltd., Shenzhen, China

e-mail: jtzhang@huawei.com

Abstract—Providing ubiquitous very high data rate coverage in next generation wireless networks is a formidable goal, requiring cost-effective radio access network (RAN) devices, such as multi-user enabled amplify-and-forward (AF) relays, and fair radio resource management (RRM). To further this goal, we investigate multi-user enabled AF relays which multiplex user’s data in orthogonal frequency division multiple access (OFDMA). These relays are cost-effective, simpler to implement, and introduce less delay in comparison to other relay based routers. We devise a generalized proportionally fair (GPF) RRM framework for multi-user enabled AF relays. In GPF scheduling a single parameter is used to gradually change schedules from throughput optimal to proportionally fair. We formulate the GPF scheduling problem and due to its complexity devise a low complexity heuristic to solve it. We evaluate the performance of the heuristic with extensive simulations and show that the heuristic performs well.

Index Terms—Radio Resource Management, Amplify-and-Forward, Sub-Carrier Pairing, Fairness.

I. INTRODUCTION AND MOTIVATION

Current state-of-the-art standardization activities are leading towards peak data rates in the order of one gigabit per second (Gbps) in the downlink. While it is still early for the standardization bodies to consider much higher data rates of tens of Gbps, this is clearly a timely and important research topic due to the exponential growth of user traffic on existing networks. Providing very high data rate coverage, when and where required, is a formidable goal, requiring dense cost-effective radio access network (RAN) architectures. Since path loss, fading, and transmit power limitations prevent high spectral efficiency even for moderately long links, it is necessary to consider advanced RANs, such as relay networks, which effectively collect and distribute wireless signals. However, to achieve the full potential of the advanced RANs, efficient RRM techniques are also necessary to match the demand with limited wireless resources in a fair way.

We consider RANs with multi-user enabled amplify-and-forward (AF) relays, which multiplex user data. OFDMA-based AF relays buffer quantized samples of the symbols until they are amplified and transmitted at a later time.

As evident in today’s wired networks, implementing hop-by-hop routing is a huge challenge at high data rates due to the hardware complexities of fast packet header inspection.

Without AF relaying, similar issues would rise in very high-data rate wireless networks. In addition, since the AF relays do not decode packets, channel decoder delays are eliminated, reducing its impact on higher layers.

Previous work shows that scheduling for AF relay networks holds great promise [1]–[4]. With a single user, the scheduling problem becomes matching the input sub-carriers to output sub-carriers to maximize the sum-rate capacity – process called sub-carrier pairing. Since this problem is equivalent to the assignment problem, it can be solved by the Hungarian algorithm [3]. However, due to the special structure of the problem, a solution can also be obtained by matching input and output carriers, which were first sorted according to their spectral efficiency [1]. This technique is known as ordered sub-carrier pairing. The ordered sub-carrier pairing also minimizes BER in a high SNR regime [5]. Extensions taking interference information into account are also possible [1], as well as extensions that include power allocation [6].

AF scheduling in the multiuser setup is not as simple as scheduling in the single-user setup. Complications arise from the need to provide end-to-end fairness among user rates. In addition since the problem is also an end-to-end scenario in the physical layer, it is more difficult than the DF scheduling counterpart.

This paper presents two contributions to multi-user AF relay scheduling. First, we devise a generalized proportionally fair (GPF) scheduling framework to make digital AF relay scheduling fair to users. GPF scheduling is also known as γ -fair scheduling in the wired networking literature [7]. Users are assigned utility functions, which take user’s rate and a parameter γ as inputs. According to the value of the γ parameter, the utilities are able to gradually change resource allocation from throughput optimal to proportionally fair. Second, since finding GPF schedules is computationally hard, we propose a heuristic to quickly find schedules in each frame. Our heuristic is based on the gradient of the γ -fair utility functions, so it is similar to the proportionally fair scheduling algorithm [8], which was proposed for conventional cellular networks. Unlike [8], which finds *long-term* fair rates, our heuristic finds *short-term* fair rates in each frame.

We evaluate the performance of GPF scheduling and our heuristic with extensive simulations and show that the heuristic is close to optimum.

II. SYSTEM AND NETWORK MODEL

We assume that the resources are assigned using OFDMA technology. The orthogonal sub-carriers are grouped in time and frequency into resource blocks (RBs), with duration of T_b seconds and a frequency span of W_b Hertz. We say that RBs using the same frequency span are on the same sub-channel. There are T RBs in the frame and N available sub-channels.

We consider RANs with OFDMA-based AF relay stations (RSs), which are capable of multiplexing different users' data. The RS receives the signal from the base-station (BS), samples it, performs the Fast-Fourier Transform (FFT) to get the received modulation symbols on each sub-carrier, and stores them in its buffer. After receiving the signal for T_b seconds, the RS has one RB in its buffer for each sub-channel, so it may re-map the RBs to different sub-channels, before performing the inverse FFT to obtain the output signal, similar to single-user "chunk-based" sub-carrier coupling [3]. We note that because of the buffering, the RS has up to T RBs before re-transmitting them to user terminals, allowing for scheduling of multiple users in the same frame and on the same sub-channel (Fig. 1).

We assume a network of M users connected to the BS through a predetermined RS at any given time. A higher layer process determines which M users are connected to the BS through this RS; the other users are connected to other RSs. We note that each "user" may correspond to an application.

The AF RS amplifies the received symbols before multiplexing and retransmitting them. The highest adaptive modulation and coding (AMC) available on the combined link from the BS to the RS and from RS to the users in an RB depends on the combined signal-to-noise ratio (SNR) in its time interval and frequency span. The number of bits in each RB is given by $b_{ij}^{(m)}$, which is obtained from a look-up table containing a mapping of channel conditions to the maximum number of bits in an RB. We only consider scheduling on the downlink; uplink scheduling is identical.

Radio resources are assigned to the users in terms of RBs; each RB carries data of only one user at a time. The rate of a user is determined from the number of RBs it is allocated in the frame and the AMC used in each RB. The rate of user m is

$$r_m = \frac{1}{T_b} \sum_{i=1}^N \sum_{j=1}^N b_{ij}^{(m)} x_{ij}^{(m)}, \quad (1)$$

where $b_{ij}^{(m)}$ is the number of bits transmitted to user m in RBs on the sub-channel coupling (i, j) , in which (i, j) refers to the coupling when BS transmits RBs on first-hop sub-channel i and the relay re-transmits them on the second-hop sub-channel j , $x_{ij}^{(m)}$ is the number of RBs assigned to user m on sub-channel coupling (i, j) , and T_b is the duration of the resource block in time (Fig. 1). We use $x_{ij}^{(m)}$ to indicate that the slot allocations are the unknowns the RRM algorithm is searching for.

From a networking perspective the user rates should satisfy some type of "fairness", otherwise the network operator may have too many unhappy customers who are starved out by the

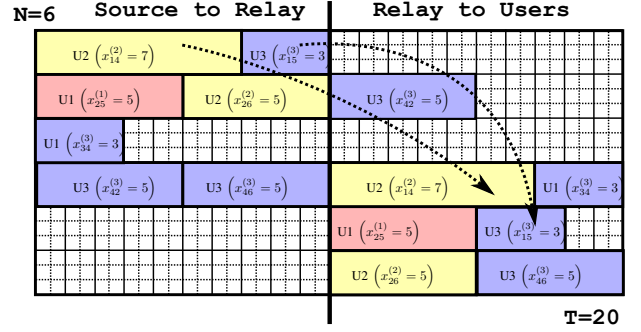


Fig. 1. A Simple AF Relay Schedule.

users with high AMC (spectral efficiency). Instead of requiring strict proportional fairness [9] among the user rates, we wish to have a more general proportional fairness, where the network operator has the flexibility to modify the scheduler to exchange fairness for throughput.

We take the approach from networking research [7], where each user is assigned a utility function, such that when the network utility is maximum for a given set of user rates, over all other user rates, those user rates are "fair" with respect to the utilities. Different fairness goals are achieved with different utility functions.

A family of utility functions, which result in γ -fair user rates is defined with

$$U_m(\dots, x_{ij}^{(m)}, \dots, \gamma) = \begin{cases} \frac{1}{1-\gamma} \left(\frac{1}{T_b} \sum_{i=1}^N \sum_{j=1}^N b_{ij}^{(m)} x_{ij}^{(m)} \right)^{1-\gamma} & \text{if } 0 \leq \gamma < 1 \\ \log \left(\frac{1}{T_b} \sum_{i=1}^N \sum_{j=1}^N b_{ij}^{(m)} x_{ij}^{(m)} \right) & \text{if } \gamma = 1 \end{cases}$$

, where γ is the parameter influencing the kind of fairness we expect and the term in the brackets is the user rates, (1). The sum utility over the user rates is the network utility

$$U_N(\dots, x_{ij}^{(m)}, \dots, \gamma) \triangleq \sum_{m=1}^M U_m(\dots, x_{ij}^{(m)}, \dots, \gamma), \quad (2)$$

User rates, which maximize the sum rate utility for a specific γ are said to be γ -fair. Different types of fairness can be achieved by changing the parameter γ [7]. For $\gamma \rightarrow 0$ the network utility corresponds to *throughput*,

$$U_N(\dots, x_{ij}^{(m)}, \dots, \gamma) \underset{\gamma \rightarrow 0}{=} \frac{1}{T_b} \sum_{m=1}^M \sum_{i=1}^N \sum_{j=1}^N b_{ij}^{(m)} x_{ij}^{(m)}.$$

For $\gamma \rightarrow 1$, the user rates maximizing the network utility are *proportional fair* [10]

$$U_N(\dots, x_{ij}^{(m)}, \dots, \gamma) \underset{\gamma \rightarrow 1}{=} \sum_{m=1}^M \log \left(\frac{1}{T_b} \sum_{i=1}^N \sum_{j=1}^N b_{ij}^{(m)} x_{ij}^{(m)} \right).$$

Since changing the parameter γ allows for a spectrum of scheduling algorithms, from optimizing for maximum through-

put to optimizing for proportional fairness, the resulting scheduling provides GPF.

III. γ -FAIR AF RELAY SCHEDULING

We now formulate the optimization problem, that finds time allocations for the AF relay resulting in γ -fair user rates. We call a set of rates γ -fair, if for a given γ they maximize the network utility (2) over all possible user rates (sub-channel coupling assignments).

The optimization, which maximizes network utility over all feasible user rates to find the γ -fair AF relay rates is

$$\max. \quad \sum_{m=1}^M \frac{1}{1-\gamma} \left(\frac{1}{T_b} \sum_{i=1}^N \sum_{j=1}^N b_{ij}^{(m)} x_{ij}^{(m)} \right)^{1-\gamma} \quad (3a)$$

$$\text{s.t.} \quad \sum_{m=1}^M \sum_{j=1}^N x_{ij}^{(m)} \leq \frac{T}{2}, \quad 1 \leq i \leq N, \quad (3b)$$

$$\sum_{m=1}^M \sum_{i=1}^N x_{ij}^{(m)} \leq \frac{T}{2}, \quad 1 \leq j \leq N, \quad (3c)$$

$$x_{ij}^{(m)} \in \left\{ 0, \dots, \frac{T}{2} \right\}, \quad 1 \leq i, j \leq N, 1 \leq m \leq M, \quad (3d)$$

where M is the number of users, N is the number of sub-channels, T_b is the time duration of the resource block, γ is the parameter which sets the type of fairness, $b_{ij}^{(m)}$ is the number of bits that can be transmitted to user m on sub-channel coupling (i, j) , and $x_{ij}^{(m)}$ is the number of RBs assigned to user m on sub-channel coupling (i, j) .

The objective function maximizes the network utility, which is the sum utility of all users. Depending on the parameter γ , the optimization results in different types of rate fairness. The constraints (3b) and (3c) ensure that the total number of allocated blocks does not exceed what is available in the frame (Fig. 1). The constraint (3d) ensures the integrality of sub-channel coupling.

The integrality of time allocations, which is exhibited by the discrete nature of constraint (3d) makes the problem computationally hard. However, if the integrality of time allocations is relaxed, by replacing the constraint (3d) with

$$0 \leq x_{ij}^{(m)} \leq \frac{T}{2}, \quad 1 \leq i, j \leq N, \quad 1 \leq m \leq M, \quad (4)$$

the optimization becomes a convex problem, which can be solved with an off-the-shelf convex optimization package.

A nice feature of the relaxed optimization is that it is also an upper bound on the integer solution of (3), so we can use it to verify the performance of heuristics.

Finding the time-allocations with convex programming suffers from several deficiencies. First, the optimal solution consists of real-numbers, which should somehow be converted to integers. Second, the size of the optimization can quickly get out of control. The optimization has MN^2 variables and $2N$ constraints. For a 30 user network with 50 sub-channels, there are 75,000 variables in the optimization, which

Algorithm 1 ALGORITHM-GPF(b_{ij}^m, M, N, T)

Initialize: $1 \leq i, j \leq N : T_i^{(BS)} = T/2, T_j^{(RS)} = T/2$

1: $\forall i, j, m : \tilde{b}_{ij}^m \leftarrow b_{ij}^m$

2: **while** $\exists T_i^{(BS)} > 0$ **and** $\exists T_j^{(RS)} > 0$ **do**

3: $(i^*, j^*, m^*) \leftarrow \arg \max_{\substack{1 \leq m \leq M, \\ 1 \leq i, j \leq N}} \frac{\tilde{b}_{ij}^{(m)}/T_b}{\left(\frac{1}{T_b} \sum_{i=1}^N \sum_{j=1}^N b_{ij}^{(m)} x_{ij}^{(m)} \right)^\gamma}$

4: $x_{i^*j^*}^{(m^*)} \leftarrow x_{i^*j^*}^{(m^*)} + 1$

5: $T_{i^*}^{(BS)} \leftarrow T_{i^*}^{(BS)} - 1$

6: $T_{j^*}^{(RS)} \leftarrow T_{j^*}^{(RS)} - 1$

7: **if** $T_{i^*}^{(BS)} = 0$ **then**

8: $\tilde{b}_{i^*j^*}^{(m^*)} \leftarrow 0, \quad 1 \leq m \leq M, \quad 1 \leq j \leq N$

9: **end if**

10: **if** $T_{j^*}^{(RS)} = 0$ **then**

11: $\tilde{b}_{i^*j^*}^{(m^*)} \leftarrow 0, \quad 1 \leq m \leq M, \quad 1 \leq i \leq N$

12: **end if**

13: **end while**

challenges even the best solvers. Therefore we devise an integer-based heuristic for the network utility maximization, with relatively simple complexity. The heuristic is based on the fact that the maximum change in the objective function, that can be obtained from increasing one time-allocation by one, is obtained by adding time allocation in the direction of the steepest gradient of the objective function.

Suppose we want to add one slot to user m on sub-channel coupling (i, j) . By Taylor's expansion,

$$U_N(\dots, x_{ij}^{(m)} + 1, \dots) \approx U_N(\dots, x_{ij}^{(m)}, \dots) + \frac{\partial}{\partial x_{ij}^{(m)}} U_N(\dots, x_{ij}^{(m)}, \dots),$$

where

$$\frac{\partial}{\partial x_{ij}^{(m)}} U_N(\dots, x_{ij}^{(m)}, \dots) = \frac{b_{ij}^{(m)}/T_b}{\left(\frac{1}{T_b} \sum_{i=1}^N \sum_{j=1}^N b_{ij}^{(m)} x_{ij}^{(m)} \right)^\gamma}$$

is strictly positive. So, if we are given a choice of increasing any one of $x_{ij}^{(m)}$, we should increase the time allocation of the sub-channel coupling (i, j) for user m with the highest partial derivative, to maximize the incremental change in the objective function.

Using the previous observation about the objective function, we devise an iterative greedy heuristic algorithm (**Algorithm GPF**). In each iteration, the user with the highest partial derivative is allocated an RB on its highest available sub-channel coupling (Steps 3-4). Variables $T_i^{(BS)}$ and $T_j^{(RS)}$ keep track of the available slots on each channel in the first and second parts of the frame, respectively (3b), (3c). After each iteration $T_i^{(BS)}$ and $T_j^{(RS)}$ are updated if any slots are allocated on their channels (Steps 5-6). The bits-per-slot values $\tilde{b}_{ij}^{(m)}$ are also updated (set to zero) according to the availability of RBs, to ensure that allocated slots are not considered in the next iteration (Steps 7-12). Note that our algorithm is

TABLE I
SIMULATION PARAMETERS

Parameter	Value
BS-RS Channel	Rician, K=10 dB ([11])
BS-RS Shadowing	Log-normal, variance 3 dB
BS-RS Doppler Shift	4 Hz
RS-Users Channel	Rayleigh ([11])
RS-Users Shadowing	Log-normal, variance 5 dB
RS-Users Doppler Shift	37 Hz
Path Loss	$38.4 + 2.35 \log_{10}(d)$ dB
Sub-Carrier Bandwidth	10.9375 kHz
Sub-Carriers per Sub-Channel	18
Number of Users	$M = 30$
Number of Sub-Channels	$N = 50$
Slots per Frame	$T = 20$
Cell Radius	1000 m
BS-RS Distance	500 m
Transmit Power	40 dBm BS, 30 dBm RS
Antenna Gain	10 dB BS, 5 dB RS, 0 dB Users
Noise Figure	2 dB RS, 2 dB Users

independent of the choice of utility function.

The algorithm's complexity depends on the implementation of the search in Step 3. We do not get into the specifics of the algorithm's implementation. However, we note that the search step can be implemented with multiple sorted lists holding SNRs; namely M lists for second-hop SNR measured at users, and one list for first-hop SNR measured at RS. It takes $N \log(N)$ steps to sort each list. With the sorted lists, in Step 3, we know the best available coupling for each user without any computation so we can find the user with the best partial derivative in M steps. Taking into account that there are $N \frac{T}{2}$ RBs in each hop, the algorithm goes through $N \frac{T}{2}$ iterations. Finally the worst-case complexity of the algorithm is $\mathcal{O}((M+1)N \log(N) + MN \frac{T}{2})$.

We note the relationship between the algorithm and one of the procedures proposed for single channel, single-hop, networks [8]. In contrast, our algorithm is for two-hop AF networks. The connection is not unexpected given the fact that both our approach and [8] use the same utility functions to achieve fairness. The difference is that our utility function takes the instantaneous frame rate, while in [8] the utility function takes in the current average of the rates. So, our optimization is performed in every frame for *short-term* fairness among user rates, whereas the optimization in [8] is performed in every frame to obtain *long-term* fairness among the rates.

IV. SIMULATION RESULTS

We ran a Monte-Carlo simulation for a network 30 users connected to BS through a predetermined RS. In each iteration of the Monte-Carlo simulation we randomly "drop" the users with a uniform density in the area around the relay. From users' locations, we calculate each user's path-loss to the relay and use a detailed channel model to find the number of bits carried in an RB on each sub-channel for 20 frames. There are a total of 40 drops for a total of 800 distinct inputs to the optimization. Even this modest number of drops took

about 20 hours to run due to the time it takes to find the upper bound values with the convex solver. Details of the simulation parameters are shown in Table I. We have also run our simulations with other parameters with similar results.

For each drop, we calculate user's time allocations using the relaxed optimization \hat{x}_{ij} and the heuristic \tilde{x}_{ij} . Since the relaxed optimization is the upper bound on the integer solution of the heuristic, the heuristic's gap is given by

$$\Delta_H \triangleq \frac{|U_N(\dots, \tilde{x}_{ij}^{(m)}, \dots) - U_N(\dots, \hat{x}_{ij}^{(m)}, \dots)|}{|U_N(\dots, \hat{x}_{ij}^{(m)}, \dots)|},$$

where $|\cdot|$ is the absolute value of its operand. Since the relaxed optimization upper bounds the value of the integer optimization, this is the maximum gap between the optimal integer solution of the heuristic.

For a more detailed comparison, we also find the difference between the individual user rates

$$\delta_m = \left| \frac{1}{T_b} \sum_{i=1}^N \sum_{j=1}^N b_{ij}^{(m)} (\hat{x}_{ij}^{(m)} - \tilde{x}_{ij}^{(m)}) \right|, \quad (5)$$

so that we can find the similarity between the way the heuristic and the optimal solution allocate user rates

$$\delta_H = \frac{(\sum_{i=1}^M \delta_m)^2}{(M \sum_{i=1}^M \delta_m^2)}. \quad (6)$$

Table II shows the heuristic's gap (Δ_H) and the similarity of the allocated rates (δ_H). We see that the gap is relatively small. In fact for $\gamma = 0$, there is no gap within this precision since the heuristic finds the rates maximizing the system throughput in this case. The similarity between the heuristic and the optimal solution is relatively small, indicating the presence of a large amount of multiuser diversity in the network which causes the heuristic and the upper bound to have the same objective rate with different rate allocations. For $\gamma = 0.45$ and $\gamma = 1.0$ the heuristic's gap (Δ_H) is still smaller than 5% on average. In these cases, the similarity between the rates (δ_H) increases since the rates are allocated fairly despite the available multiuser diversity.

Figure 2 shows the portion of the system rate allocated to users as a function of distance from the relay. We included all rates from all drops and fitted an exponential function ke^{-at} through all of them. We observe that as γ decreases from 1 to 0, more system resources are assigned to users close to the RS, increasing the system throughput. So, the heuristic scheduler achieves generalized proportional fairness.

V. CONCLUSION

We consider multi-user enabled AF relays, which multiplex user data. We looked at sub-carrier pairing problem in the multi-user case and formulated an GPF frame based scheduling problem for AF OFDMA relay networks.

The scheduler allows adjustable fair scheduling, which can find schedules from rates maximizing throughput to propor-

TABLE II
SUB-OPTIMALITY AND SIMILARITY OF THE HEURISTIC

	$\gamma = 0.00$	$\gamma = 0.20$	$\gamma = 0.45$	$\gamma = 1.00$
Δ_H (mean)	0.00%	3.43%	2.92%	0.45%
Δ_H (st. d.)	0.00%	0.68%	0.43%	0.06%
δ_H (mean)	11.10%	34.70%	49.08%	58.63%
δ_H (st. d.)	5.95%	9.01%	9.60%	7.53%

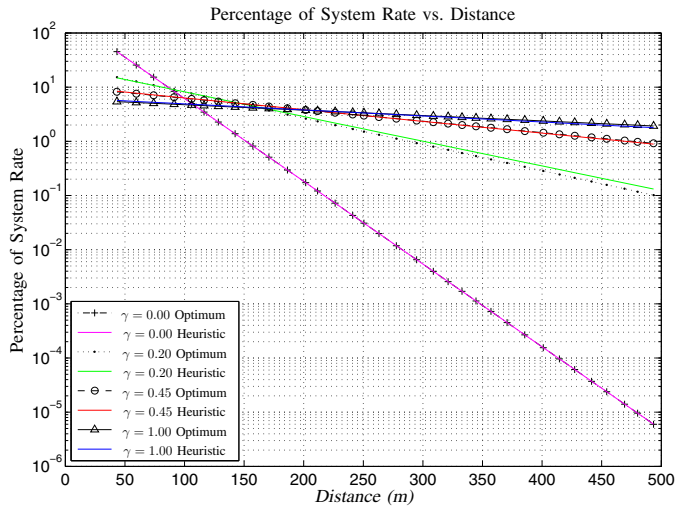


Fig. 2. Distribution of Network Resources.

tionally fair rates. Since finding the optimal schedules has high complexity we devise a greedy heuristic scheduler. Our simulations show that the network operator is able to adjust the parameter of the fairness to move between throughput optimal scheduling, which takes full advantage of multiuser diversity, and proportionally fair scheduling which is the result of considering both multiuser diversity and fairness. The simulations also show that the proposed heuristic is very close to the optimum solution.

REFERENCES

- [1] A. Hottinen and T. Heikkinen, "Optimal subchannel assignment in a two-hop OFDM relay," in *IEEE 8th Workshop on Signal Processing Advances in Wireless Communications*, June 2007, pp. 1–5.
- [2] —, "Subchannel assignment in OFDM relay nodes," in *40th Annual Conference on Information Sciences and Systems (CISS)*, March 2006, pp. 1314–1317.
- [3] M. Herdin, "A chunk based OFDM amplify-and-forward relaying scheme for 4G mobile radio systems," in *International Conference on Communications (ICC)*, vol. 10, 2006, pp. 4507–4512.
- [4] T. Riihonen, R. Wichman, and A. Hottinen, "Analysis of subcarrier pairing in a cellular OFDMA relay link," in *International ITG Workshop on Smart Antennas*, 2008, pp. 104–111.
- [5] C. K. Ho and A. Pandharipande, "BER minimization in relay-assisted OFDM systems by subcarrier permutation," in *IEEE Vehicular Technology Conference (VTC)*, May 2008, pp. 1489–1493.
- [6] Y. Guan-Ding, Z. Zhao-Yang, C. Yan, C. Shi, and Q. Pei-liang, "Power allocation for non-regenerative OFDM relaying channels," in *International Conference on Wireless Communications, Networking and Mobile Computing*, vol. 1, Sept. 2005, pp. 185–188.
- [7] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Transactions on Networking*, vol. 8, no. 5, pp. 556–567, Oct 2000.

- [8] H. Kushner and P. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions," *IEEE Journal on Wireless Communication*, vol. 3, no. 4, pp. 1250–1259, July 2004.
- [9] F. P. Kelly, "Fairness and stability of end-to-end congestion control," *European Journal of Control*, vol. 9, pp. 159–176, 2003.
- [10] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, no. 3, pp. 237–252, March 1998.
- [11] L. Hentilä, P. Kyästi, M. Köske, M. Narandžić, and M. Alatosava, "Matlab implementation of the winner phase II channel model ver1.1," http://projects.celtic-initiative.org/winner+/phase_2_model.html, December 2007.