Queue-Aware Resource Allocation for Multi-cell OFDMA Systems with QoS Provisioning

Ioannis G. Fraimis · Stavros A. Kotsopoulos

Published online: 9 January 2013 © Springer Science+Business Media New York 2013

Abstract This paper proposes a queue-aware resource allocation algorithm which provides quality of service (QoS) guarantees. The proposed solution adopts a cross-layer design approach since it is aware of both users' queue buffer states (data link layer) and channel quality state (physical layer). Main advantages of the proposed resource allocation algorithm are: the low computational complexity and its capacity of maintaining lower QoS violation probability than other multi-cellular schemes. The proposed solution can also result in enhanced cell-edge data rate and improved fairness performance. User minimum data rate and target bit error rate as considered as QoS parameters. Validation of the proposed algorithm is achieved through various simulation scenarios wherein QoS violation probability, system fairness, user average data rate and cell-edge throughput are investigated. Numerical results and complexity analysis demonstrate the efficiency and the feasibility of the proposed QoS-oriented approach.

Keywords OFDMA · Resource allocation · Cross-layer · QoS

1 Introduction

Orthogonal frequency division multiplexing (OFDM) has been widely acknowledged as the ideal radio access platform to support high-speed wireless services in next-generation cellular networks, such as the 3GPP long therm evolution (LTE) [1] and the worldwide interoperability for microwave access (WiMAX) [2]. The multi-user version of the OFDM is the orthogonal division multiple access (OFDMA) scheme, whereby multiple access is achieved by dividing the available bandwidth into a number of orthogonal bands and then

I. G. Fraimis (🖂) · S. A. Kotsopoulos

Wireless Telecommunications Laboratory (WTL), Department of Electrical and Computer Engineering, University of Patras, 26500 Rio, Greece e-mail: ifraimhs@gmail.com assigning subsets of them to individual users. OFDMA helps exploit multi-user diversity over frequency-selective channels, since it is very likely that some carriers that are in deep fade for some users, can provide great channel gain to other users [3] and in systems where adaptive modulation and coding (AMC) is employed, greater channel gain results in higher data rate. The optimality of the OFDMA has also been proved in [4,5], where it is shown that the optimal policy is to allocate each subcarrier to only one user in order to eliminate the intra-cell co-channel interference.

Future cellular networks along with the adoption of the OFDMA as the air-interface technology are expected to provide ubiquitous high data rate coverage to large number of users. On the other hand, bandwidth is a limited resource and should be judiciously utilized. Hence, the need for introducing new efficient resource allocation schemes is becoming very strong [6–12].

Next-generation system users should enjoy services and applications with high level of QoS, irrespective of location and wireless link quality. In other words, a service has to reach even the most disadvantaged users like cell-edge users which experience high path-loss and strong inter-cell interference. Recently proposed radio resource allocation schemes which solely aim at maximizing the aggregate system data rate [6] tend to overlook cell-edge users, while this problem is considered in other approaches like [7–12].

On the other hand, static reuse partitioning schemes like soft frequency reuse (SFR) scheme [7] and partial frequency reuse (PFR) scheme [8] have recently attracted much of the researchers' attention. In those schemes, different degrees of frequency reuse factors and higher power levels for cell-edge are used. However, such schemes need prior frequency planning which makes them unsuitable for future cellular systems which are expected to incorporate femtocell networks [13].

Therefore, dynamic resource allocation schemes which rely on dense spectrum reuse have recently been introduced. Authors in [9] apply interference avoidance by restricting the transmission through certain channels in each cell in order to enhance cell-edge performance, while the scheme in [10] gives greater trunking gain to users in cell boundary which results in higher cell-edge bit-rate. In [11], minimum rate constraints are imposed to cell-edge users, while in [12] a dynamic fractional frequency reuse is proposed with the objective to maximize system data rate while satisfying individual users' lower data rate requirements. Although schemes like in [9,10] show an improved cell-edge performance, they do not guarantee a certain level of QoS to users and algorithms like in [11,12] are not queue-aware and fail to exploit traffic diversity which is unrealistic.

Taking into account the afore-mentioned issues, in this paper we present a queue-aware resource allocation algorithm towards ubiquitous high data rate coverage. The objective of the proposed scheme is to maximize the sum of users' data rates while satisfying a certain level of QoS. The contributions and the features of the proposed algorithm are given below:

- We propose a dynamic resource allocation scheme with the objective of maximizing the total system throughput while providing to all serving users a minimum level of QoS.
- We provide a linear-complexity iterative algorithm in order to solve the formulated problem.
- The proposed scheme not only exploits multiuser diversity but also is capable of exploiting traffic diversity since it is a queue-aware scheme contrary to schemes in [10–12].
- The proposed scheme applies dense frequency reuse while it does not require frequency planning like static schemes do.

• Proportional fair scheduler (PFS) is used which is widely accepted as a good solution, since is provides a favorable tradeoff between the maximum average throughput and user fairness [9,14,15].

2 System Model and Problem Formulation

2.1 System Model

In this section we provide the problem formulation, from which our algorithm is derived. The resource allocation algorithm collects cross-layer information through channel state information (CSI) and queue state information (QSI) signalling and allocates resources according to the obtained information.

We consider the downlink of a multi-cellular network configuration of 19 base stations (BSs), whereby each BS is located in the center of each cell. Each BS of the cell serves *K* single antenna users via *N* subchannels and maintains *K* separate user buffers. Poisson traffic model is assumed and bit streams arrive from higher layers with arrival rate λ_k for user *k*.

2.2 Channel and Data Rate Models

The time-varying channels between the BS and different users are assumed to be frequency selective with independent Rayleigh fading and the coherence time of the multipath fading channel is assumed to be greater than the downlink scheduling interval time duration.

The signal-to-noise plus interference ratio (SINR) of user *k* on subcarrier *n* is denoted as γ_{kn} and it is given by:

$$\gamma_{kn} = \frac{p_c g_{kn}}{\sum_j p_c g_{jn} + \sigma^2}, \quad j \neq k,$$
(1)

where p_c is the transmit power applied on subcarrier *n* derived from equal power distribution, that is $p_c = P_t/N$, where P_t is the BS total transmission power. Note that water-filling power allocation only brings marginal performance improvement over fixed power allocation with adaptive coding and modulation (ACM) [4,6]. Parameter σ^2 is the noise power spectral density and parameter g_{kn} is the link gain between user *k* and its serving BS, on subcarrier *n* and includes large scale path-loss, fading and shadowing. The resulting achievable rate is r_{kn} and using the continuous rate formula for ACM:

$$r_{kn} = \frac{B}{N} \log_2 \left(1 + \frac{-1.5\gamma_{kn}}{\ln(5P_e)} \right),\tag{2}$$

where B is the available bandwidth and P_e is the target bit error rate.

We also define the subchannel allocation matrix $X_{K \times N} = [x_{kn}]$, whereby $x_{kn} = 1$ if user k is assigned subchannel n and $x_{kn} = 0$ if not. Each BS attains its own $X_{K \times N}$ matrix.

2.3 Problem Formulation

The objective of the proposed radio resource allocation scheme is to improve the cell capacity while ensuring a minimum level of QoS and a target P_e to all serving users.

Taking into account the above assumptions, the resource allocation during a scheduling interval duration or time slot is formulated as a binary integer linear programming problem as follows:

$$\max\sum_{k=1}^{K}\sum_{n=1}^{N}r_{kn}x_{kn}$$
(3)

subject to:

$$\sum_{k=1}^{K} x_{kn} = 1, \forall n \tag{4}$$

$$x_{kn} \in \{0, 1\}, \forall \{k, n\}$$

$$(5)$$

$$\sum_{n=1}^{N} x_{kn} r_{kn} \ge C_k, \forall j$$
(6)

Constraints (4), (5) ensure a disjoint subcarrier assignment, wherein one subcarrier can only be given to at most one user in a cell (intra-cell interference elimination). Constraint (6) denotes the minimum QoS level of each user expressed through the minimum data rate C_j (in bits per second) of user *j* for a specific service.

In order the algorithm to avoid serving an empty queue and wasting resources, it should be:

$$T\sum_{n=1}^{N} x_{kn} r_{kn} \le q_k, \quad \forall k,$$
(7)

where T denotes the time slot duration and q_k is the queue length in the BS, of user k in Mbps.

It should be pointed out that finding optimal solution would face prohibitive computational burden (i.e., Sect. 4), where its optimal solution must also be obtained within a designated time period due to quick channel variations. Hence, suboptimal algorithms with low complexity are preferable which are more practical, particularly for systems with large number of subchannels and users which is often the case in practise.

3 The Proposed Cross-Layer Resource Allocation Scheme

In this section, we present the proposed radio resource allocation algorithm. The steps of the proposed algorithm are as follows:

- Step 1: Form the set of available subchannels \mathcal{N} and users \mathcal{K} . Determine the set \mathcal{C} which contains the minimum data rate requirements of users and set all elements of $X_{K \times N}$ to 0.
- *Step 2*: Assign subchannels to users according to the standard PF scheduling [14, 15], until users' minimum data rate requirements are reached.
- *Step 3*: Schedule the remaining resources to users according to the standard PF scheduling algorithm.

It is worth to mention that Proportional fair (PF) scheduler is widely considered as an attractive compromise between maximum average throughput and user fairness [14,15]. Optimal criteria for PF scheduling for multicarrier systems has been discussed in [14], but its practical implementation is very complex, considering systems with numerous users and hundreds of subcarriers. With PF scheduling, at a depiction epoch t + 1, if the feasible rate of user k is $R_k(t + 1)$ and it's *m*-point moving average throughput up to epoch t is $\bar{R}_k(t)$,

then the user k^* with $k^* = \arg \max_k \frac{R_k(t+1)}{\bar{R}_k(t)}$ will be selected. The moving average throughput of that user is updated by:

$$\bar{R}_k(t+1) = (1 - 1/m)\bar{R}_k(t) + I_k(t+1) \cdot R_k(t+1)/m,$$
(8)

where $I_k(t + 1)$ is the indicator function of the event that user k is selected in decision epoch t + 1 and $I_k(t + 1) = 1$, if user k is selected and $I_k(t + 1) = 0$ if not. As a result, once user k has been selected, his chance of being given other subcarriers will be decreased, or other users have higher priorities to select resources.

In the following we give the pseudo-code of the proposed cross-layer resource allocation algorithm, whereby for simplicity we omit parameter t.

(1) Initialization:

- Set $\mathcal{K} = \{1, 2, \dots, K\}, \mathcal{N} = \{1, 2, \dots, N\}, \mathcal{C} = \{C_1, C_2, \dots, C_K\}$
- $x_{kn} = 0, \forall k, n$
- $\mathbf{q} = (q_1, q_2, \dots, q_K)$

(2) While $C \neq \{\}$:

• Find user
$$\tilde{k} \in \mathcal{K}$$
 and subcarrier $n^* \in \mathcal{N}$, with
 $\tilde{k} = \arg \max_k \frac{r_{kn^*}}{\tilde{R}_k}, \forall k, n$
• If $\sum_{n=1}^N x_{\tilde{k}n} r_{\tilde{k},n} < C_{\tilde{k}}$
- $x_{\tilde{k}n^*} = 1$
- $q_{\tilde{k}} = \max\{q_{\tilde{k}} - (r_{\tilde{k},n^*} \cdot T)\}$
- If $q_{\tilde{k}} = 0$ then $\mathcal{K} \setminus \{\tilde{k}\}$
- $\mathcal{N} \setminus \{n^*\}$
- Calculate (3)
• If $\sum_{n=1}^N x_{\tilde{k}n} r_{\tilde{k}n} \ge C_{\tilde{k}}$
- $\mathcal{C} \setminus \{C_{\tilde{k}}\}$

(3) While $\mathcal{N} \neq \{\}$ and $\sum_{k=1}^{K} q_k \neq 0$:

- Find subcarrier $n' \in \mathcal{N}$ and user $k^* \in \mathcal{K}$ with $k^* = \arg \max_k \frac{r_{kn'}}{\bar{R}_k}$
- $x_{k^*n'} = 1s$
- $q_{k^*} = \max\{q_{k^*} (r_{k^*n'} \cdot T)\}$
- If $q_{k^*} = 0$ then $\mathcal{K} \setminus \{k^*\}$
- $\mathcal{N} \setminus \{n^*\}$
- Calculate (3)

In step 1, the algorithm initializes the involved variables. Forms the user and channel set while also the BS defines each user's minimum data rate service requirement and initializes the queue vector. The elements of channel allocation matrices are set to 0 and in this part.

In step 2, the algorithm performs PF scheduling in each cell. It finds the pair of user and subcarrier which maximizes the ratio of instantaneous rate over average rate and allocates the subcarrier to the user. Updates the queue state of the users and if the queue is empty the user is removed from the user set. In the following, if the user has reached its service

minimum data rate requirement, it is removed from set C and it is excluded from this part of the algorithm.

In step 3, The algorithm examines if there are unallocated subcarriers left from the previous part and if there is traffic to be sent to users. While this situation is true the algorithm distributes the remaining resources according to PF scheduler.

4 Complexity Analysis

Problem (3)–(7) is a binary integer linear programming problem and the computational complexity for the optimal solution is prohibitively high. In a cell with *K* users and *N* subchannels, there are K^N possible subchannel assignments. As a result the direct search has an exponential complexity of $\mathcal{O}(K^N)$. In our case, step 1 requires constant time for initialization and in the step 2, the algorithm needs to compare *K* users at most, in order to allocate a subchannel. Thus, this part is upper-bounded by complexity $\mathcal{O}(KN)$. In step 3, assuming that $N' \neq 0$ unallocated subchannels exist after the end of step 2, the algorithm searches for the user $k^* = \arg \max_k \frac{r_{k,n'}}{R_k}$ among *K* users for the remaining N' unallocated subchannels, whereby $n' \in N'$. Hence, this part shows complexity of order $\mathcal{O}(KN')$. Note that always $N' \leq N$ and the equality is true in case no minimum rate requirements exist. Thus, the overall complexity of the algorithm is $\mathcal{O}(KN)$. It has a form of $a\mathcal{O}(z^b)$ with $0 < a < \infty$ and b = 1 which is a linear time algorithm [6] and efficient compared to the number of comparisons K^N for the original problem.

5 Simulation Parameters

The available spectrum is B = 20 MHz with carrier frequency 2.0 GHz. The available subcarriers are N = 144. The BS available power is $P_t = 46$ dBm and the variance of the additive noise is $N_0 = -174$ dBm/Hz. Target bit error rate is equal to $P_e = 10^{-7}$ and the pathloss model is described by $L_k(d) = 128.1 + 37.6 \cdot \log(d_k)$ [dB] with standard deviation 8 dB for shadowing, where d is the distance from the serving BS (km). All users are assumed to fall in the same service class and they are randomly distributed in various distances and angles from the serving BS. The minimum distance far from the BS is 35 m and maximum is 1 km which is defined as the cell radius.

The frequency-selective channel consists of six independent Rayleigh multipath components (taps) each one generated using the Jakes model with an exponentially decaying power delay profile [10, 16]. The ratio of the energy of the *l*th tap to the first tap is e^{-2l} and maximum doppler of 30 Hz is assumed. We consider two user density scenarios. A scenario for a light loaded system with 30 users per cell and a second one for a heavy loaded system with 65 users per cell. Performance results have been collected from 150 user drops whereby each drop is simulated over a time of 100 channel realizations. For simplicity of simulations, a simple queuing model with Poisson arrival is considered and the arrival data rate is assumed to be 5 Mbps.

In SFR and PFR schemes, cell-edge area is considered we consider the outer one third of the total cell area. We consider the power per subcarrier in the cell-edge band, which is also the one to be $a_p P_t/N$, while in the cell-center band is $(3 - a_p)P_t/2N$ with $a_p = 1.5$. The cell-edge band is the one third of the total band for the SFR scheme. For the PFR scheme the cell-edge band has a reuse factor of 3 and the cell-center band is considered to be the half of the total available bandwidth with reuse factor of 1. Channel power profiles of PFR are the same with SFR (please see [9] and [11]).



Fig. 1 Cumulative distribution function of user throughput for K=30 user density per cell



Fig. 2 Cumulative distribution function of user throughput for K = 65 user density per cell

6 Numerical Results

The Proposed algorithm is contrasted against traditional schemes applicable in multi-cell OFDMA systems which apply dense frequency reuse and all reference schemes exploit multiuser diversity since they employ PF scheduling and they are used as performance benchmarks



Fig. 3 Zoomed cdf of user data rate to show cell-edge throughput for K = 30 user density per cell



Fig. 4 Zoomed cdf of user data rate to show cell-edge throughput for K = 65 user density per cell

for our proposed scheme. We consider that all users fall in the same service class and we also assume two different service classes with two different minimum data rate requirements, 750 Kbps and 1.25 Mbps respectively. Figures 1 and 2 plots the cumulative distribution function (CDF) of user throughput for the service models of the proposed algorithm as well as the reference schemes. The lower tail of the CDF of user throughput for the system with K = 30 user density is zoomed in Fig. 3 and for the system with user density K = 65 user



Fig. 5 QoS violation probability versus service minimum data rates for K = 30 user density per cell



Fig. 6 QoS violation probability versus service minimum data rates for K = 65 user density per cell

density is zoomed in Fig. 4 respectively, in order to view the cell-edge throughput clearly. According to the LTE evaluation methodology, the cell-edge throughput is defined as the 5th percentile of the lower tail of the user throughput [17]. we can see that enhanced cell-edge throughput can be achieved by regulating properly the users' lower data rate bounds with impact to the average user throughput irrespective of system user density which reveals the algorithm's flexibility. More specifically, in the case of K = 30 user density, the proposed



Fig. 7 System fairness for K = 30 user density per cell

algorithm with $C_k = 750$ Kbps and $C_k = 1.25$ Mbps gives improved cell-edge throughput by 29 and 69.7% compared to Reuse-1 respectively. Moreover, the variants of the proposed algorithm show higher cell-edge area throughput than SFR and PFR whose design priority is the enhanced cell-edge throughput. In case of K = 30 users per cell, variant with $C_k = 750$ Kbps gives 18% higher cell-edge throughput than PFR and 5% than SFR. Similarly, variant with $C_k = 1.25$ gives 55.8 and 39.7% better performance than PFR and SFR respectively.

When the system user density is K = 65 users per cell, then the variant $C_k = 750$ Kbps shows improvement by 58.3% compared to Reuse-1 and 16.7% when the minimum data rate per user is $C_k = 1.25$ Mbps. Cell-edge data rate is enhanced by 18.75 and 26.67% when $C_k = 750$ Kbps in comparison with PFR and SFR respectively.

Quality of service (QoS) violation probability for K = 30 is depicted in Fig. 5 against various QoS requirements and we can see that our scheme results in better QoS provisioning than the other reference schemes and the same happens for K = 65 (Fig. 6). Finally from Figs. 7 and 8 we see that the proposed scheme gives higher Jain's fairness index [18,19]. Therefore, the proposed algorithm is capable of distributing the available resources in a fairer manner than other reference schemes.

7 Conclusion

A novel queue-aware resource allocation algorithm which provides QoS guarantees to users has been proposed. The proposed solution is a network-distributed scheme which has been compared against other popular distributed schemes.

We applied a cross-layer design approach, since the algorithm used information for both data link and physical layer in order to select a user and allocate resources. Furthermore, the proposed solution showed reduced complexity compared to the original problem.



Fig. 8 System fairness for K = 65 user density per cell

Finally, numerical results revealed the superiority of the proposed scheme. The proposed algorithm not only achieved lower levels of QoS violation probability, but it also distributed resources in fairer manner than existing multi-cell schemes.

References

- E-UTRA and E-UTRAN overall description; stage 2 (release 8). 3GPP technical specification TS 36.300 V8.7.0, December 2008 (Online). Available: http://www.3gpp.org.
- 2. Mobile WiMAX-part I: A tecnical overview and, performance evaluation, August 2006.
- Tse, D., & Viswanath, P. Fundamentals of wireless communication. Cambridge: Cambridge University Press.
- Jang, J., & Lee, K. B. (2003). Transmit power adaptation for multiuser OFDM systems. *IEEE Journal on Selected Areas in Communications*, 21(2), 171–178.
- Li, G., & Liu, H. (2005). On the optimality of the OFDMA network. *IEEE Communications Letters*, 9, 438–440.
- Li, G., & Liu, H. (2006). Downlink resource allocation for multi-cell OFDMA system. *IEEE Transactions* on Wireless Communications, 5(12), 3451–3459.
- Soft frequency reuse scheme for UTRAN LTE. 3GPP project document R1–050 507, May 2005 (online). Available: http://www.3gpp.org.
- OFDMA downlink intercell interference mitigation. 3GPP project document R1–060 291, February 2006 (Online). Available: http://www.3gpp.org.
- Rahman, M., & Yanikomeroglu, H. (2010). Enhancing cell-edge performance: A downlink dynamic interference avoidance scheme with inter-cell coordination. *IEEE Transactions on Wireless Communications*, 9(4), 1414–1425.
- Fraimis, I. G., Papoutsis, V. D., & Kotsopoulos, S. A. (2010). A decentralized subchannel allocation scheme with inter-cell interference coordination (ICIC) for multi-cell OFDMA systems. In *Proceedings* of the IEEE global communications conference (Globecom).
- Quek, T. Q. S., Lei, Z., & Sun, S. (2009). Adaptive interference coordination in multi-cell OFDMA systems. In Proceedings of the IEEE 20th international symposium on personal, indoor and mobile radio communications (PIMRC).

- Ali, S. H., & Leung, V. C. M. (2009). Dynamic frequency allocation in fractional frequency reused OFDMA networks. *IEEE Transactions on Wireless Communications*, 8(8), 4286–4295.
- Chandrasekhar, V., Andrews, J. G., & Gatherer, A. (2008). Femtocells networks: A survey. *IEEE Com*munications Magazine, 56–59.
- Kim, H., & Han, Y. (2005). A proportional fair scheduling for multicrrier transmission systems. *IEEE Communications Letters*, 9(3), 201–212.
- Liu, E., & Leung, K. (2010). Expected throughput of the proportional fair Scheduling over Rayleigh fading channels. *IEEE Communications Letters*, 14(6), 515–517.
- Jakes, Jr. W. C. Microwave mobile communications. A Wiley-Interscience publication. New York, Chichester, Brisbane, Toronto: Wiley.
- Dahlman, E., Ekstrom, H., Furuskar, A., Jading, Y., Karlsson, J., Lundevall, M., et al. (2006). The 3G longterm evolution-radio interface concepts and performance evaluation. In *Proceedings of IEEE vehicular* technnology conference (VTC).
- Jain, R., Chiu, D. M., & Hawe, W. (1984). A quantitave measure of fairness and discriminaton for resource allocation in shared systems. DEC, research report TR-301.
- Ramadas, K., & Jain, R. (2007). Mobile WiMAX part I: A technical overview and performance evaluation. WiMAX Forum. http://www.wimaxforum.org/technology/downloads/.

Author Biographies



Ioannis G. Fraimis was born in Patras, Greece, in 1982. He received his Diploma in Electrical and Computer Engineering from the University of Patras, Greece and he is a Ph.D. holder from the Wireless Telecommunications Laboratory of the same University. He has published 14 papers and he is also reviewer at International Journals and Conferences. He has also worked in the "National Organization of Telecommunications of Greece" and in the "Center for Computing, Information and Communication Systems-Dept. of Electrical and Computer Eng., Univ. of Patras'. His research interests include power control in wireless systems, green communication networks and cognitive radio communication systems with emphasis on QoS and resource allocation algorithms in multiuser environments.



Stavros A. Kotsopoulos was born in Argos, Greece, in 1952. He received his B.Sc. in Physics in the year 1975 from the Aristotle University of Thessaloniki, and in the year 1984 got his Diploma in Electrical and Computer Engineering from the University of Patras. He did his postgraduate studies in the University of Bradford in United Kingdom. He is a M.Phil. and Ph.D. holder since 1978 and 1985 respectively. Currently he is a member of the academic staff of the Department of Electrical and Computer Engineering of the University of Patras and holds the position of Professor. Since 2004, he is the Director of the Wireless Telecommunications Laboratory and develops his professional life teaching and doing research in the scientific area of Telecommunications, with interest in cellular mobile communications, wireless networks, interference, satellite communications, telematics applications, communication services and antenna design. Moreover, he is the (co)author of the book titled "mobile telephony". He is member of the Greek Physics Society and member of the Technical Chamber of Greece.