Channel Aware and Queue Aware Scheduling in LTE Uplink

Javad Hajipour[1], Amr Mohamed[2], and Victor C.M. Leung[1]

[1] University of British Columbia
[2] Qatar University
`{hajipour,vleung}@ece.ubc.ca`, `amrm@qu.edu.qa`

Abstract. Long Term Evolution (LTE) uses single carrier frequency division multiple access (SCFDMA) technique as the multiple access scheme in the uplink, due to its low peak to average power ratio (PAPR) compared to orthogonal frequency division multiple access (OFDMA). This advantage is achieved when the Resource Blocks (RBs) allocated to a user are contiguous in frequency domain. Considering this constraint we devise Channel and Queue Aware Scheduling (CQAS) algorithms to keep the users' queue sizes low and at the same time utilize the system resources efficiently. Using extensive simulations we verify the performance of these algorithms in terms of system throughput and queue size probability.

Keywords: LTE, SCFDMA, Resource Allocation, Channel and Queue Aware Scheduling.

1 Introduction

Long Term Evolution (LTE) standardization [1], proposed by the Third Generation Partnership Project (3GPP), aims at providing high speed data and multimedia services to mobile users and therefore has recently attracted a lot of attention. To achieve its goals, LTE employs orthogonal frequency division multiple access (OFDMA) as the downlink (DL) multiple access solution. OFDMA is a promising technique which provides high spectral efficiency, scalable bandwidth and robustness against multipath impairment. However, since it modulates data on multiple carriers, it leads to a high peak to average power ratio (PAPR). This makes it unattractive for the uplink (UL) where mobile handsets have power and amplifier limitations. Therefore as an alternative to OFDMA, LTE has selected single carrier frequency division multiple access (SCFDMA) for UL. SCFDMA, also known as discrete Fourier transform spread OFDMA (DFT-Spread OFDMA) can be considered as a modified version of OFDMA in which using DFT, data symbols are transformed into frequency domain before being mapped onto orthogonal subcarriers. SCFDMA can lead to lower PAPR while providing OFDMA benefits at the same time. This becomes possible when all the subcarriers allocated to a user are adjacent.

Both in DL and UL, multiple access is performed by allocating frequency and time resources in the units of Resource Blocks (RB) [2] to users. Each RB is

X. Zhang and D. Qiao (Eds.): QShine 2010, LNICST 74, pp. 116–127, 2011.

composed of several adjacent subcarriers in frequency domain and symbols in time domain. Resource allocation and scheduling play an essential role in the efficient use of network resources. Since channel conditions for different users independently vary over the time and frequency domains, channel aware resource allocation methods that utilize multiuser diversity by allocating RBs to the users with favourable channel conditions are desirable. In OFDMA, frequency domain packet scheduling (FDPS) has flexibility in allocating RBs to users, whereas in SCFDMA, RBs allocated to a user need to be contiguous in the frequency domain [2] . This constraint makes it difficult for FDPS to exploit multiuser diversity efficiently.

Recently several papers have worked on packet scheduling in SCFDMA and proposed different heuristic and suboptimal algorithms. The channel dependent scheduling algorithms in [3,4] do not consider any contiguity constraint on allocated RBs. In [5], authors have shown that applying the contiguity constraint results in about 2.7 dB decrease in PAPR. They have proposed three suboptimal algorithms which have been verified through simulations. Similar work has been performed in [6] with the proof to show the NP-hardness of the problem. Both [5] and [6] have taken into account the fact that channel conditions of users over RBs are correlated in the frequency domain, which makes it possible to cope with the contiguity constraint and at the same time utilize channels efficiently. A more general case is considered in [7], which defines a utility function to represent various scheduling policies and devises two approximation algorithms with polynomial runtime. However, how to choose the set of contiguous RBs used in the algorithms has not been explained.

Most of the previous work has considered an infinitely backlogged model, where users' queues are permanently backlogged and therefore always have data to transmit. In this paper we consider a finite-queue model for users where data packets arrive based on a random process and are buffered in users' queues before transmissions. To the best of our knowledge this is the first time that joint Channel and Queue Aware Scheduling (CQAS) is addressed in the LTE UL and the contiguity constraint for allocation of RBs in SCFDMA is considered together with finite data packets available in users' queues for transmissions. We propose different CQAS algorithms for allocation of RBs considering these two constraints. Using simulations, we investigate the performance of the proposed algorithm in terms of queue length and system throughput.

The remainder of this paper is organized as follows. Section 2 introduces the system model and CQAS. CQAS algorithms for SCFDMA are presented in section 3. Section 4 provides the performance evaluation. We conclude the paper in Section 5.

2 Preliminaries

2.1 System Model

We consider UL in a single cell of an LTE system, where a base station (BS) is located in the center and there are K active users inside the cell. Transmission bandwidth is divided into N RBs indexed by $n, n = 1...N$. In each time slot, FDPS decides about the allocation of RBs to the users. More than one RBs can be allocated to a single user with contiguity constraint applied. However each RB can be allocated to at most one user. We assume each user has a single flow of data where packets arrive randomly and are queued in a buffer before they are transmitted in the assigned RBs. We assume that BS has perfect knowledge about the queue size of users and channel condition of every user in all the RBs. Channel information can be obtained through Sounding Reference Signals (SRS) [8] which are sent by users as Channel Quality Indicator (CQI) and queue size information can be achieved based on the users' Buffer Status Reports (BSR)[9], together with the history of their scheduled rates. It is also assumed that channel state remains constant during each time slot. Based on this information and according to the scheduling algorithms, FDPS determines which user to transmit in each RB.

2.2 Channel and Queue Aware Scheduling

In wireless networks, users experience different channel conditions due to different path loss, shadowing and fading. Considering these, it is possible to use system resources efficiently. Opportunistic channel aware scheduling algorithms like Max C/I [10] allocate the channels to the user with the best condition, to obtain the maximum possible throughput in the system. However this method leads to the starvation of the users who experience deep fading in their channels. To prevent this, Proportional Fair Scheduling (PFS) [11] considers the history of the users access to the system in addition to their instantaneous achievable rates. Therefore it is possible to allocate resources to all the users whenever they have higher ratio of achievable rate to average rate. PFS is mostly considered for services that have infinitely backlogged queues. For services with finite buffer sizes it is important to maintain queue lengths in a reasonable range to keep the buffer overflow probability as low as possible. For this purpose, CQAS policies [12],[13],[14] take the length of users' queues into account as well as their channel conditions and give higher priority for transmission to the users with good channel condition or larger queue size. One of the metrics used, is MaxWeight [14]. One version of it is as follows

$$\rho_k = Q_k \min(Q_k, r_k) \tag{1}$$

where ρ_k, Q_k and r_k are, respectively, the MaxWeight metric, the queue length and achievable channel rate for user k. Other than the constraints mentioned in the previous section, UL CQAS has one other challenge. It is due to the fact that each user might have different maximum transmission power and based on the number of RBs allocated to it, different power levels will be available for transmission on each RB. This will have an effect on the modulation scheme and coding rate used for transmission on each RB given the transmissions quality objective, e.g., bit-error rate (BER). Therefore it will affect the number of bits the user can transmit on each RB and as a result the number of RBs needed to transmit a specific number of bits. In the following, we summarize these challenges

- RBs allocated to a user should be contiguous in frequency domain.
- Queues of users have a finite amount of data to send.
- Different users might have different power constraints and based on the number of RBs allocated to each user, transmission power on each RB can be different, and therefore different number of bits can be transmitted on each RB while satisfying the BER objective.

In this paper, considering these challenges, we propose resource allocation algorithms in SCFDMA and evaluate their performance.

3 CQAS in SCFDMA

In this section we propose three algorithms. In the first one we aim to calculate the number of needed RBs for each user to empty its queue and allocate the needed RBs continuously to the users that utilize them better. In the next two algorityms, we use the MaxWeight metric to allocate the RBs to users that have higher values for this metric on the corresponding RB. In all of the following algorithms it is assumed that the users use their maximum power and this power will be divided equally among all the subcarriers of allocated RBs.

3.1 Algorithm 1

This algorithm is illustrated in **Algorithm 1**. Available users are the users that have data in their queues to transmit. $P_{max,k}$ in line 6 is the maximum power of user k and $r_k(j, p)$ indicates the achievable channel rate of user k on RB j if the power used for this RB is p. The algorithm starts from the first RB and performs the following for each available user.

Algorithm 1

```
 1: Let $S$ be the set of available users
 2: Initialize $n = 1$.
 3: Do:
 4:   For each $k \in S$ compute the following
 5:     $M_k = min(N - n + 1,$ number of RBs needed to send $Q_k$ bits$)$
 6:     $C_k = min(Q_k, \sum_{j=n}^{n+M_k-1} r_k(j, \frac{P_{max,k}}{M_k}))$
 7:     $\alpha_k = \frac{C_k}{M_k}$
 8:   select $k^* = \arg\max_k \alpha_k$
 9:   Allocate $RBs : n...n + M_{k^*} - 1$ to user $k^*$
10:   $n = n + M_{k^*}$.
11:   $S = S - \{k^*\}$
12: while $(n \leq N)$ and $(\sum_{k \in S} Q_k > 0)$
```

It first assumes that one RB will be enough and therefore all the power of the user will be given to that RB. Based on this assumption, the channel achievable rate is calculated. If this rate does not empty the queue in one slot, the number of needed consecutive RBs will be considered two and the achievable rate over two RBs will be computed assuming maximm power divided equally among them. This calculation is continued to compute the number of RBs needed to empty the queue. Then user k^* is selected according to line 8. α_k is a measure of utilization of RBs by user k. The larger this parameter is, the better the RBs are used. In other words it shows that a user has better channel conditions on its allocated RBs. After selecting k^*, its needed RBs are assigned to it and the user will be deleted from the available users list. The similar procedure will continue for the rest of RBs to allocate them to available users.

3.2 Algorithm 2

In this algorithm, illustrated in **Algorithm 2**, we consider a MaxWeight metric for user k over each RB n, as follows

$$\rho_{k,n} = Q_k \min(Q_k, r_k(n, \frac{P_{max,k}}{N})) \tag{2}$$

where Q_k is the updated queue length of user k **before** allocation of RB n, which depends on the allocation of previous RBs. Starting from first RB, RBs are assigned one by one, based on the MaxWeight metric, as follows.

For each RB n, $n = 1...N$, this metric is calculated for the available users, and the user with highest metric value is selected and assigned to RB n. Then the information about the queue sizes and the number of assigned RBs are updated and considered in the allocation of next RBs.

Since the number of RBs that are going to be allocated to each user is not known a priori, therefore for calculating achievable channel rate of a user on an

Algorithm 2

```
Let $S$ , $U_k$ respectively be the set of available users and the set of RBs already allocated user $k$
Initialize $k_0 = 0$ ,$n = 1$ and $U_k = \emptyset$ for every user $k$
Do:
  For each $k \in S$ compute the following
    $\rho_{k,n} = Q_k min(Q_k, r_k(n, \frac{P_{max,k}}{N}))$
  Select $k^* = \arg\max_k \rho_{k,n}$
  $Q_{k^*} = Q_{k^*} - min(Q_{k^*}, \sum_{j \in \{U_{k^*} + \{n\}\}} r_{k^*}(j, \frac{P_{max,k^*}}{|U_{k^*}| + 1}) - \sum_{j \in U_{k^*}} r_{k^*}(j, \frac{P_{max,k^*}}{|U_{k^*}|}))$
  $U_{k^*} = U_{k^*} + \{n\}$
  If $k^* \neq k_0$ then $S = S - \{k_0\}, k_0 = k^*$
  If $Q_{k^*} = 0$ then $S = S - \{k^*\}$
  $n = n + 1$
while $(n \leq N)$and$(\sum_{k \in S} Q_k > 0)$
```

RB, we assume the allocated power by user k for each RB is $\frac{P_{max,k}}{N}$, as in (2). However after selection of the user according to line 6, the power for the RB is considered equal to $\frac{P_{max,k}}{|U_k|+1}$, where U_k is the set of RBs already assigened to user k. Based on this power, achievable channel rates are computed and queue length of selected user is updated according to line 7.

In case that the selected user's queue will be emptied after allocation of the RB, that user will be deleted from the list of available users. Also in order to comply with contiguity constraint, if the selected user for the current RB is different from the user assigned to the previous RB, the user of previous RB will be deleted from the list of available users and will not be considered in the allocation of the next RBs.

3.3 Algorithm 3

This algorithm, illustrated in **Algorithm 3**, utilizes the channel correlation in frequency domain, using similar idea as in [5] and [6]. Figure 1 demonstrates this idea.

Algorithm 3

```
Let $S$, $A$ and $U_k$ respectively be the set of available users, the set of available RBs and the set of RBs already assigned to user $k$
Calculate the metric $\rho_{k,n} = Q_k \; min(Q_k, r_k(n, \frac{P_{max,k}}{N}))$ for each user $k$, on each RB $n$ and sort them in descending order of values in the set $V$
Do:
  Pick $k, n$ corresponding to first$\rho_{k,n}$ in $V$
  Initialize $n_e = n, n_h = n - 1$
  Do:
    check($k, n_e, 1, t1$)
    check($k, n_h, -1, t2$)
  while $t_1$ or $t_2$
  $S = S - \{k\}$
  $V = V - \{\rho_{i,x} | i = k\}$
while $A \neq \emptyset$ and $(\sum_{k \in S} Q_k > 0)$
Check($k$,$m$,$direction, result$)
  if $m \in A$ and $Q_k > 0$ and $\arg\max_i \rho_{i,x} = k$
    $Q_k = Q_k - min(Q_k, \sum_{j \in \{U_k + \{m\}\}} r_k(j, \frac{P_{max,k}}{|U_k| + 1}) - \sum_{j \in U_k} r_k(j, \frac{P_{max,k}}{|U_k|}))$
    $U_k = U_k + \{m\}$
    $A = A - \{m\}$
    $m = m + direction$
    $result = true$
  else
    $result = false$
  end if
end check
```

At first, it computes MaxWeight metrics for all the users on all RBs. Then it starts from the RB that has the highest value of metric among all RBs and allocates it to the corresponding user. After updating the queue size of that user, it moves one RB towards the head of the RB sequence and if the same user has the highest metric, it is allocated to the same user. Then the similar procedure is performed one RB towards the end of the RB sequence. This is continued as long as the user has data in its queue and the metric of user on either sides are the highest among other users. Then these RBs and the user are deleted from the list of available RBs and available users.

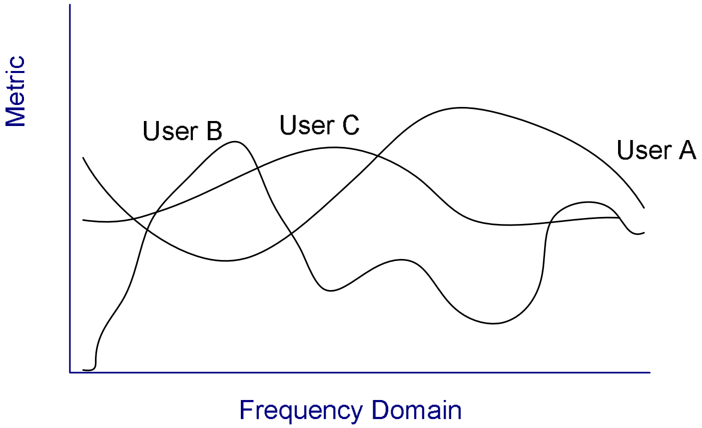

Fig. 1. Utilizing channel correlations for resource allocation

During allocation of adjacent RBs to the user, power allocation to RBs and queue size information are updated in the same way as in **Algorithm 2**. The same procedure is performed for the remaining RBs as long as there are data in queues of available users.

4 Performance Analysis

4.1 Simulation Parameters

To evaluate the performance of the proposed algorithms we have conducted Matlab simulations over 2000 time slots. Simulation parameters are shown in table 1.

Results are presented in terms of system throughput as well as average and maximum queue sizes in the system. Rates are computed, assuming a minimum mean square error (MMSE) equalizer at the receiver and using the following equation [4]:

$$r_k = \frac{BW|U_k|}{N}.\log_2(1+\beta\gamma_k) \tag{3}$$

Table 1. Simulation Parameters

Parameter Name	Setting
Cell Radius	1000m
Min UE-BS distance	50m
Pathloss and Lognormal Shadowing	From [15]
UE speed	3 km/h
Fast Fading Model	Vehicular A [15]
System Bandwidth	3MHz
Subcarrier Spacing	15 kHz
RB size	12 subcarriers
Number of RBs:N	15
Number of Users:K	10
Time Slot Duration	1ms
User Maximum power	24dBm
Noise Power Spectral Density	-174dBm/Hz
Traffic Model	Poisson
Packet Inter Arrival Time	8ms
Packet Size	2kbit

where BW is the system bandwidth, $|U_k|$ is the number of RBs allocated to user k, β is a constant related to the BER by $\beta = \frac{-1.5}{ln(5BER)}$, and γ_k is the SNR for the data delivered with RBs in U_k defined as

$$\gamma_k = \left(\frac{1}{\frac{1}{|U_{k,sub}|}\sum_{i\in U_{k,sub}}\frac{\gamma_{k,i}}{\gamma_{k,i}+1}} - 1\right)^{-1} \tag{4}$$

where $|U_{sub,k}|$ is the number of subcarriers allocated to user k which is equal to $12|U_k|$, $\gamma_{k,i}$ is the SNR of user k on subcarrier i which is defined as

$$\gamma_{k,i} = \frac{P_k^{(sub)} H_{k,i}}{\sigma_n^2 \triangle f} \tag{5}$$

where $P_k^{(sub)}$ is the allocated power of user k to each of its assigned subcarriers, $|H_{k,i}|$ is the channel gain of user k on subcarrier i including path loss, shadowing and fast fading, σ_n^2 is the noise power per Hz and $\triangle f$ is the subcarrier spacing in Hz.

4.2 Numerical Results

In this subsection, we investigate the performance of our proposed algorithms. As a refernce for our comparisions, we use Alg3(riding peaks) in [6], by considering PF metric in **Algorithm** 3 as follows

$$\rho_{k,n} = \frac{r_k(n, \frac{P_{max,k}}{N})}{R_k} \tag{6}$$

where R_k is the average rate of user k up to current time slot.

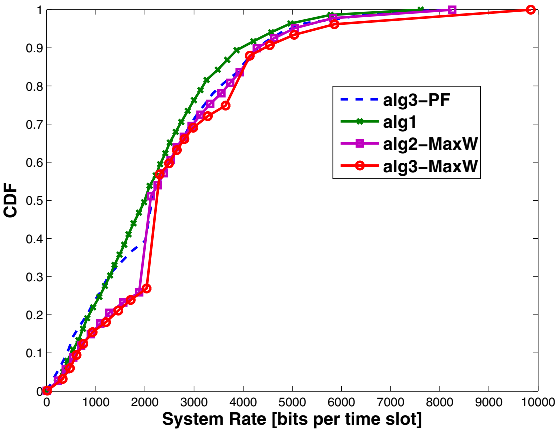

Fig. 2. Distribution of system throughput in each time slot; $K = 8$

Figure 2 shows the probability distribution of system throughput in each time slot. Reference algorithm is shown with *alg3-PF*, to distinguish it from *alg3-MaxW*, which is the **Algorithm** 3 with MaxWeight metric. Also to indicate the use of MaxWeight metric in **Algorithm** 2, it is shown as *alg2-MaxW*.

It is observed that *alg3-MaxW* has the best performance, as it has higher probability for higher throughputs than the other algorithms. *alg1* has the lowest throughput performance. The reason is that it tries to empty the queues and therefore allocates RBs as much as needed. This does not allow multiuser diversity to be utilized over several RBs as much as the other algorithms. *alg2-MaxW* utilizes multiuser diversity by allocating RBs to users that have higher MaxWeight metric. Unlike *alg3-PF*, it considers queue size of users in the allocation of RBs and therefore is able to allocate RBs to the users that have good channel and at the same time have data to transmit. However because *alg2-MaxW* does not start allocation from the RBs with highest metric, it sometimes removes the users from the available users list in the first steps of allocation. Hence it cannot utilize their probable better channel conditions on the next RBs.

Figures 3 and 4 illustrate the probability distribution of the average queue size and maximum queue size in the system, respectively.

It is observed that *alg1* has the lowest performance in terms of users' queue sizes. The reason is that although it aims at emptying the queues, it does not reach its goal as it does not allow more users to utilize RBs to transmit data. The opposite is true with *alg3-MaxW*, which utilizes system resources more efficiently by allocating channels with the best metrics to the corresponding users. Therefore it is able to allow more bits to be sent from the users queues in each time slot.

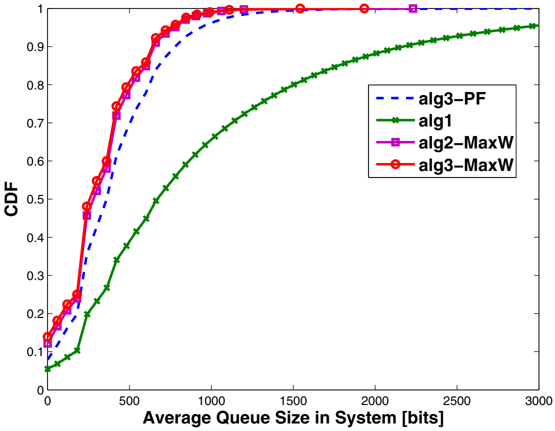

Fig. 3. Distribution of average queue size in each time slot; $K = 8$

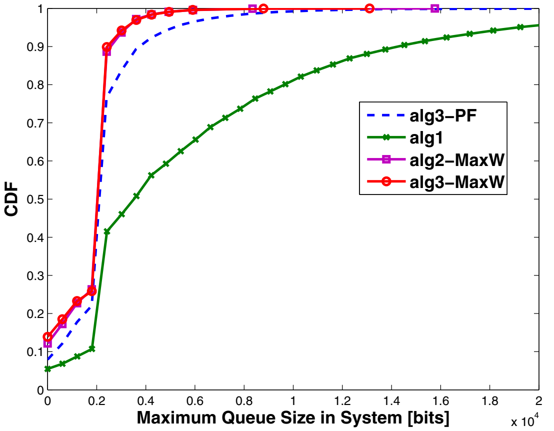

Fig. 4. Distribution of maximum queue size in each time slot; $K = 8$

Figure 5 shows the average queue size of users over time, for different number of users in different algorithms. We observe that as the number of users increases, the queue sizes of users increases rapidly with *alg1* and *alg3-PF* and cannot be kept within a reasonable range, while the average queue sizes for *alg3-MaxW* and *alg2-MaxW* has negligible increase. This shows that *alg3-MaxW* is best able to control the queue size, followed by *alg2-MaxW*, whereas *alg1* and *alg3-PF* are least able to do so.

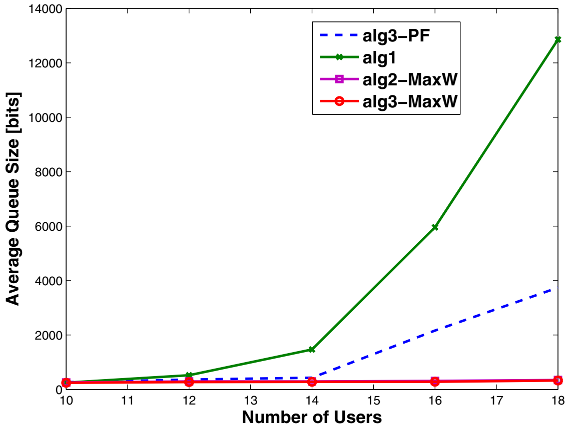

Fig. 5. Average queue size over time v.s. number of users

5 Conclusion

In this paper we have studied channel and queue aware scheduling in LTE UL. Considering the contiguity constraint in allocation of RBs to users, we have addressed the finite queue model for the first time in SCFDMA bandwidth and power allocation. We have proposed three algorithms and using simulations, we have evaluated their performance in terms of system throughput and queue sizes of users in the system. Numerical results show the best performance when using MaxWeight metric and starting allocation from the RBs with the highest metric value. In our future work we will try to improve these algorithms to provide quality of service support for different types of services.

Acknowledgement. This work is supported in part by the Qatar Telecom (QTel) Grant no. QUEX-Qtel-09/10-10 and grants from TELUS and the Natural Sciences Engineering Research Council of Canada (NSERC).

References

1. 3GPP TS36.300: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN): Overall Description (October 2010)
2. Holma, H., Toskala, A.: LTE for UMTS-OFDMA and SC-FDMA Based Radio Access (2009)
3. Lim, J., Myung, H.G., Oh, K., Goodman, D.J.: Proportional Fair Scheduling of Uplink Single-Carrier FDMA systems. In: IEEE PIMRC (September 2006)
4. Lim, J., Myung, H.G., Oh, K., Goodman, D.J.: Channel-Dependent Scheduling of Uplink Single Carrier FDMA Systems. In: IEEE PIMRC (September 2006)

5. Angel, L., Berardinelli, G., Frattasi, S., Mogensen, P.: Channel-Aware Scheduling Algorithms for SC-FDMA in LTE Uplink. In: IEEE PIMRC (September 2008)
6. Lee, S., Pefkianakis, I., Meyerson, A., Xu, S., Lu, S.: Proportional Fair Frequency-Domain Packet Scheduling for 3GPP LTE uplink. In: IEEE INFOCOM, pp. 2611–2615 (April 2009)
7. Yang, H., Ren, F., Lin, C., Zhang, J.: Frequency-Domain Packet Scheduling for 3GPP LTE Uplink. In: IEEE INFOCOM (2010)
8. 3GPP TR 25.814: Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA) (October 2006)
9. 3GPP TS 36.321: Technical Specification Group Radio Access Network; Medium Access Control (MAC) protocol specification (Release 8) (March 2008)
10. Shariat, M., Quddus, A., Ghorashi, S., Tafazolli, R.: Scheduling as an important cross-layer operation for emerging broadband wireless systems. IEEE Communications Surveys and Tutorials 11, 74–86 (2009)
11. Bender, P., Black, P.J., Grob, M., Padovani, R., Sindhushyana, N., Viterbi, S.: CDMA/HDR: a bandwidth efficient high speed wireless data service for nomadic users. IEEE Communications Magazine 38, 70–77 (2000)
12. Tassiulas, L., Ephremides, A.: Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. IEEE Transactions on Automatic Control 37(12), 1936–1948 (1992)
13. Andrews, M., Kumaran, K., Ramanan, K., Stolyar, A., Whiting, P., Vijayakumar, R.: Providing quality of service over a shared wireless link. IEEE Communications Magazine 39, 150–153 (2001)
14. Andrews, M., Zhang, L.: Scheduling algorithms for multi-carrier wireless data systems. In: ACM Mobicom (2007)
15. Jeruchim, M.C., Balaban, P., Shanmugan, K.S.: Simulation of Communication Systems, 2nd edn. Kluwer Academic/Plenum, New York (2000)