Green Resource Allocation for OFDMA Wireless Cellular Networks

Emilio Calvanese Strinati and Paolo Greco

CEA, LETI, MINATEC, F-38054 Grenoble, France Email: emilio.calvanese-strinati@cea.fr, paolo.greco@cea.fr

ABSTRACT

Abstract—Packet scheduling algorithms are viewed as one of the key mechanisms for increasing the diversity order, robustness and effectiveness of a wireless multi-user communication system. Traditional packet scheduling are mainly designed to increase the system capacity. In this paper we present a novel scheduling algorithm that improves the scheduling energy efficiency in OFDMA based wireless cellular networks. Our goal is to reduce the overall downlink energy consumption while adapting the target of spectral efficiency to the actual load of the system and meeting the Quality of Service (QoS). Our analysis reveals how the proposed approach permits to achieve notable energy gain over traditional scheduling algorithm especially in not saturated scenarios.

I. INTRODUCTION

Telecommunication has experienced a tremendous success causing proliferation and demand for ubiquitous heterogeneous broadband mobile wireless communications. Up to now, innovation has targeted to improve wireless networks coverage and capacity while meeting the QoS for users admitted in the system. Nowadays, the number of mobile subscribers equals more than half the global population. Forecast on telecommunication market assume an increase in subscribers, per subscriber's data rate and, the roll out of additional base stations for next generation mobile networks. The undesired consequence is the growth of wireless network's energy consumption which will cause an increase of the global carbon dioxide (CO_2) emissions and, impose more and more challenging operational cost for operators. Communication energy efficiency represent indeed an alarming bottleneck in the telecommunication growth paradigm.

Recently, increasing maturity of mobile technology in combination with the growing amount of equipment deployed each year have woken up the need of innovating in the field of energy efficient communications [1]. Energy efficient enhancement in wireless communication can be achieved only if improvements are experienced in the whole communication chain for different operational load scenarios. Several investigations are on going on this research area, ranging from energy efficient cooling of base stations, to innovative energy efficient deployment strategies and frequency planning.

In this paper we focus on downlink energy efficient resource allocation techniques for OFDMA based systems. We consider LTE like systems which face a very challenging multi-user communication problem: many users in the same geographic area require high on-demand data rates in a finite bandwidth with a variety of heterogeneous services such as voice (VoIP), video, gaming, web browsing and others.

Classical scheduling algorithms such as the maximum channel to interference ratio (MCI) [2] or the proportional fair (PF) [3] were designed for OFDMA based systems to gain from space, time, frequency and multi-user diversity of a wireless system and attempt to instantaneously achieve an objective (such as the total sum throughput, maximum throughput fairness, or pre-set proportional rates for each user) regardless of QoS constraints of the active users in the system. MCI scheduler allocates resources to users with the highest momentary instantaneous capacity; PF scheduler tries to balance the resource allocation and serve momentary good users (not necessarily the best) while providing long term throughput fairness (equal data rates amongst all users). On the other hand, some scheduling algorithms are designed to support specific QoS constraints. For instance, Earliest Deadline First (EDF) [4] is designed to deal with real-time QoS constraints regardless to the momentary user's channel quality. Other schedulers are designed to cope with the coexistence of RT and NRT traffics (mixed traffic). One of the most famous scheduler which is designed to take into account both traffic delay constraints and instantaneous channel rate is the Modified Largest Weighted Deadline First (MLWDF) [5]. Its design objective is to maintain delay (and throughput) of each traffic smaller (and greater) that a predefined threshold value with a given probability. All these scheduling algorithms have been designed to maximize the system capacity while overcoming the mismatch between requested quality of service (QoS) and limited network resources. Alternatively, the scheduler could attempt to minimize the energy required per correctly delivered information bits, while meeting the QoS of admitted heterogeneous active users.

Information theorists have studied energy-efficient transmission for at least two decades [6] [7]. The work in [6] defines reliable communication under a finite energy constraint in terms of the capacity per unit energy, which is the maximum number of bits that can be transmitted per unit energy. This definition ensures that for any transmission rates below the capacity per unit energy, error probability decreases exponentially with the total energy. More recently, several researchers have designed solutions to trade spectral and energy efficiency. Single user energy efficiency optimization is a vivid investigation topic. In [8] the authors propose to trade off energy against transmission delay. Actually, Shannon theory indicates that it is desirable to transmit a packet over a longer period of time to save transmit energy. The authors consider a single user scenario with single type homogeneous traffic and, propose to scale the modulation down to conserve energy when its buffer is under loaded. When the buffer starts to fill up, the modulation order is increased to avoid long queuing times or buffer overflow. This proposal assumes that power consumption monotonically increases when transmission power increases. Under this assumption, as far as single user QoS constraints are met, power can be minimized by reducing the information rate at the expense of delay. Nevertheless, in case of a multi-user scenario, power consumption of base station can be rate-independent if the system is not heavily loaded. Actually, information rate depends on transmission power but also on the momentary link quality on frequency resources allocated to scheduled users. In [9], it is shown how, considering a single user's terminal energy consumption during uplink transmission, the lowest order modulation should always be used while accommodating the delay constraint to minimize energy consumption. In [9], it is recalled that the lowest order modulation should always be used while accommodating the delay constraint to minimize energy consumption. In [10], the authors propose for the single user case, an iterative algorithms to obtain optimum energy-efficient uplink link adaptation scheme, in which joint circuit and uplink transmit power consumptions are taken into account to maximize energy efficiency rather than throughput. In [11], the authors extend their work to uplink transmission in OFDMA systems to improve battery consumption at the mobiles. In [12], the authors analyze in a wireless sensor network, the best modulation strategy to minimize the total energy consumption (transmit and circuit energy) required to send a given number of bits when error-control codes are used.

Actually, as wireless is a shared medium, the overall system's energy efficiency is affected not only by the single user efficiency, but also by the combination of time and frequency allocation according to the momentary instances of the frequency selective channel between the downlink base station and the receiving selected user equipments. Hence, in order to achieve energy efficiency, a system approach is required. In our vision, the multi-user scheduler should indeed allocate time and frequency resources to minimize the transmission energy cost while meeting QoS requirements of all active users admitted by a base station.

The momentary system load plays an important role in the overall optimization design. Currently deployed base station are commonly designed so that they can accommodate the traffic demand at peak times. Nevertheless, cell traffic load notably varies during the day. Several researches pointed out that to save energy, base stations should perform a dynamic load and energy state arrangement, which balances extra load on a determined optimal set of base stations, thus maintaining minimum energy consumption. In this paper we do not propose novel dynamic sleep mode of base station components. We focus on resource allocation of active phases of base stations during downlink transmission.

The main concern of our work is to design an energy efficient multi-user scheduling algorithm that can be applied to heterogeneous traffic scenarios. Our goal is to trade momentary spectral efficiency with downlink transmission power while meeting the QoS constraints of active users admitted in the cell. To this end, we propose a scheduling algorithm which splits the resource allocation process into four steps. In a first step we identify which entities (packets) are rushing and which are not rushing. Then in step two we deal with urgencies: we assign resources only to entities that have an high probability of missing their QoS requirements regardless to their momentary link quality and their potential to save energy. Then, if any resources (here chunks) are still unscheduled, in a third step we allocate resources to users (non rushing) with highest momentary link quality, regardless to their QoS constraints. In the fourth step we perform energy efficient link adaptation to save downlink energy. We trade throughput (lowering the transmission spectral efficiency and allocating a larger number of chunks to UEs) with downlink power by limiting the power budget on each chunk. In this way we attempt to minimize downlink transmission power over a time window, which provides significant additional flexibility to the scheduling algorithm. In addition to throughput, both latency and spectral efficiency enter in the tradeoff. In an extreme case of latency tolerance or low load scenarios, the scheduler could simply just wait for the user to get close to the base station before transmitting or, allocates all frequency resource to a single packet transmission with significant lowered power. We call the proposed algorithm Green Adaptive Scheduler (GAS).

II. SYSTEM MODEL

In this work, the system model is mainly based on the 3GPP/LTE downlink specifications [13] [14], where both components of the cellular wireless network, i.e. base stations and mobile terminals, implement an OFDMA air interface. Using the terminology defined in [15], data is transmitted in frame of 10 ms equally divided in 10 sub-frames each made of two consecutive slots. The OFDM symbols are organized into a number of PRBs consisting of 12 contiguous sub-carriers for 7 consecutive OFDM symbols, i.e. a slot. With a bandwidth of 10 MHz, this leads to 50 PRBs available for data transmission. Each user is allocated one or several PRBs in two consecutive slots and, a transmission time interval (TTI) aligns on subframe and equals 1 ms. The network consists of 19 hexagonal three-sectored cells, where each BS transmits continuously and with maximum power. As a consequence, only the central cell is simulated, while others BSs are used for down-link interference generation only. Fast fading is generated using a Jakes model for modeling a 6-tap delay line based on the Typical Urban scenario [14], with a mobile speed equal to 3km/h. Flat fading is assumed for the neighboring cells. A link-to-system (L2S) interface is used in order to accurately model the physical layer at the system level. This L2S interface is based on EESM (Effective Exponential SINR Mapping) as proposed initially in [16].

In our investigation, we consider either Voice over IP traffic (VoIP) or near real time video sources (NRTV), both modeled here according to [14]. Since in our investigation we target to mimic realistic traffic scenarios, both traffic sources are not constrained to be full-queued. Furthermore, a limited number of control channels per TTI is considered, as the control channel capacity is always limited in commercial systems. In our study, this number, which corresponds to the number of maximum schedulable users in a TTI, is equal to 16, that is consistent with 3GPP/LTE system specifications [17].

In the central cell, the BS has a multiuser packet scheduler which determines the resource allocation, Adaptive Modulation and Coding (AMC) parameters and HARQ policy within the next sub-frame. We focus our investigation on two well known priority OFDMA based scheduling algorithms, which are maximum channel to interference ratio (MCI) per chunk [2] [18] and EDF (Earliest Deadline First) [4] applied to OFDMA. MCI allocates packets to users having the best instantaneous rate (over one or several PRBs). On the other hand, EDF scheduling rule allocates packets according to their remaining TTLs (Time To Live), thus granting priority to traffics with stringent QoS time constraints regardless to their momentary channel quality. With EDF, PRBs are first allocated to unscheduled packet with the smallest remaining TTL, while their spectral efficiency is maximized. At each first transmission attempt of a packet, the modulation and coding scheme (MCS) selection is based on the average EESM metric over the set of all PRBs allocated to one UE. We consider the set of MCS defined for the LTE standard (see table I. Afterwards, the scheduler sends downlink control messages that specify the resource allocation and the link adaptation parameters adopted in the next TTI while UEs send positive or negative acknowledgment (ACK/NACK) to inform the scheduler of correct/incorrect decoding of the received data. In our model either perfect or partial channel state information (CSI) can be considered for all links. Moreover, a feedback delay is included between the time when CSI is available at the destination and the time when the packet scheduler performs the resource allocation. More details on the adopted system model are summarized on Table II.

| MCS mode | Modulation | Coding Rate | Spectral efficiency |
|----------|------------|-------------|---------------------|
| 1 | QPSK | 1/3 | 2/3 |
| 2 | QPSK | 1/2 | 1 |
| 3 | QPSK | 2/3 | 4/3 |
| 4 | 16-QAM | 1/3 | 4/3 |
| 5 | QPSK | 3/4 | 3/2 |
| 6 | 16-QAM | 1/2 | 2 |
| 7 | 64-QAM | 1/3 | 2 |
| 8 | 16-QAM | 2/3 | 8/3 |
| 9 | 16-QAM | 3/4 | 3 |
| 10 | 64-QAM | 1/2 | 3 |
| 11 | 64-QAM | 2/3 | 4 |
| 12 | 64-QAM | 3/4 | 9/2 |

 TABLE I

 MODULATION AND CODING SCHEMES IN LTE.

III. OVERVIEW ON PACKET SCHEDULING ALGORITHMS FOR WIRELESS COMMUNICATIONS

Many researchers address the problem of defining an efficient and robust resource allocation strategy for multiple heterogeneous data users sharing the same resources over a wireless channel. Priority scheduler can deal with both allocation of time and frequency resources, in order to exploit multi-user diversity in both domains. This is often referred as time/frequency domain packet scheduling (TFDPS).

Two of these well known scheduling algorithms are investigated in our work: MCI and EDF. These priority

| Network | | | | |
|--|--|--|--|--|
| Parameter | Value | | | |
| | 2.0 GHz | | | |
| Carrier frequency Bandwidth | 10 MHz | | | |
| Inter-site distance | 500 m | | | |
| Minimum distance | 35 m | | | |
| TTI duration | 1 ms | | | |
| Cell layout | Hexagonal grid, 19 three-sectored cells | | | |
| Link to System interface | EESM | | | |
| Traffic model | VoIP, NRTV | | | |
| Nb of antennas (Tx, Rx) | (1,1) | | | |
| Access Technique | OFDMA | | | |
| Total Number of Sub-carriers | 600 | | | |
| | 12 | | | |
| Nb of Sub-carriers per PRB Total Nb of Chunks (PRB) | 50 | | | |
| | | | | |
| Propagation Channel | | | | |
| Parameter | Value | | | |
| Fast fading | Typical urban 6-tap model, 3 km/h | | | |
| Interference | White | | | |
| UE | | | | |
| Parameter | Value | | | |
| | | | | |
| Channel estimation | ideal | | | |
| Channel estimation | ideal ideal or partial | | | |
| CQI reporting | ideal or partial | | | |
| CQI reporting Turbo decoder | ideal or partial max Log-MAP (8 iterations) | | | |
| CQI reporting Turbo decoder Dynamic Res | ideal or partial max Log-MAP (8 iterations) ource Allocation | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter | ideal or partial max Log-MAP (8 iterations) ource Allocation Value | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC PER _{target} | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC PER _{target} CQI report | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC <i>PER</i> _{taryet} CQI report Packet Scheduling | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay EDF, MCI | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC PER _{Lasyget} CQI report Packet Scheduling Sub-carriers Allocation Strategy | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay EDF, MCI Chunk based allocation | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC <i>PER</i> _{target} CQI report Packet Scheduling | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay EDF, MCI | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC PER _{target} CQI report Packet Scheduling Sub-carriers Allocation Strategy Number of control channels per TTI | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay EDF, MCI Chunk based allocation | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC PER _{target} CQI report Packet Scheduling Sub-carriers Allocation Strategy Number of control channels per TTI | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay EDF, MCI Chunk based allocation 16 | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC <i>PERtarget</i> CQI report Packet Scheduling Sub-carriers Allocation Strategy Number of control channels per TTI H Parameter | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay EDF, MCI Chunk based allocation 16 ARQ Value | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC PER _{target} CQI report Packet Scheduling Sub-carriers Allocation Strategy Number of control channels per TTI H Parameter Stop and Wait | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay EDF, MCI Chunk based allocation 16 ARQ | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC <i>PERtarget</i> CQI report Packet Scheduling Sub-carriers Allocation Strategy Number of control channels per TTI H Parameter | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay EDF, MCI Chunk based allocation 16 ARQ Value synchronous adaptive 6 | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC PERtarget CQI report Packet Scheduling Sub-carriers Allocation Strategy Number of control channels per TTI H Parameter Stop and Wait Number of processes Retransmission Interval | ideal or partial max Log-MAP (8 iterations) Durce Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay EDF, MCI Chunk based allocation 16 ARQ Value synchronous adaptive 6 6 ms | | | |
| CQI reporting Turbo decoder Dynamic Res Parameter Nb of MCS AMC PERtarget CQI report Packet Scheduling Sub-carriers Allocation Strategy Number of control channels per TTI H Parameter Stop and Wait Number of processes | ideal or partial max Log-MAP (8 iterations) ource Allocation Value 12 (from QPSK 1/3 to 64-QAM 3/4) 10 % Each TTI, with 2ms delay EDF, MCI Chunk based allocation 16 ARQ Value synchronous adaptive 6 | | | |

 TABLE II

 MAIN SYSTEM MODEL PARAMETERS

scheduling algorithms have been proposed aiming at satisfying delay and throughput constraints of all active users or as many as possible users regardless to energy efficiency. While some scheduling algorithms take into account only the time constraints of the traffic flows (*eg.* EDF), others take into account the momentary channel state to optimize the overall cell throughput (*eg.* MCI). The key features and drawbacks of such schedulers are the following:

MCI: Its goal is to maximize the instantaneous system throughput regardless to any traffic QoS constraints. Even if maximum system throughput can be achieved with MCI, users whose momentary channels are not good for a relatively long period may starve and consequently release their connections. MCI is indeed inadequate for real-time traffic. Moreover, MCI never targets to trade spectral efficiency with energy efficiency even in low load scenarios for which system throughput do not need to be maximized to meet QoS of active UEs.

EDF: It allocates resources first to packets with smaller remaining TTLs. With OFDMA based transmission, allocation is decoupled: each packet is prioritized according to its remaining TTL (R_{TTL}) and then chunks are allocated to the ordered packets in order to maximize spectral efficiency. The drawback of this scheduler is that multiuser diversity is not exploited since any momentary channel state information is

taken into account in the scheduling rule. EDF never targets to trade spectral efficiency with energy efficiency even in low load scenarios for which transmission of packets can be delayed or slowed down to save energy.

IV. PROPOSED SCHEDULING ALGORITHM: GAS

Traditional packet scheduling algorithms are designed to increase the maximum system capacity, subject to QoS constraints and fairness. Nevertheless, most of the time, wireless systems are only moderately loaded. With state of the art priority scheduling algorithms, the exploitation of available time and frequency resources is typically not optimized from an energy perspective. The goal of our investigation is to design an energy efficient scheduling algorithms which meets the QoS constraint of an heterogeneous population of UEs. EDF-like schedulers do not profit much from time diversity as much as they should do. MCI like schedulers aim at maximizing the cell throughput regardless of the user QoS and the actual system load, and consequently they are totally insensitive to any time constraints of the data traffic. In not saturated system load scenarios, the exploitation of available time and frequency resources is typically not optimized from a energy perspective. Based on these observations, we propose to split the resource allocation process into four steps. First a Rushing Entity Classifier (REC) identifies rushing entities that must be treated with higher priority. Depending on the nature of the traffic, entities are UEs (NRT traffic) or packets (RT). This distinction is due the different definitions of OoS for different classes of traffic. Therefore, rushing entity classification is traffic-dependent. Second the proposed scheduler deals with urgencies: we schedule the transmission of rushing entities regardless to their momentary link quality. If any resources (here chunks) are still unscheduled, in a third step, GAS allocates resources to those users with better energy saving potential, regardless to their QoS constraints.

The proposed scheduling algorithm is summarized as follows:

Step 1: The REC classifies entities (packets or UEs) waiting to be scheduled as rushing or non-rushing. With RT traffic, packets are classified as rushing if $Th_{rush} \cdot TTL + \eta \ge R_{TTL}$. Where Th_{rush} is a threshold on the QoS deadline which depends on the traffic type, η is a constant which takes into account both retransmission interval and maximum allowed number of retransmissions. With NRT traffic, UEs and not packets are classified by the REC. Therefore, the i^{th} UE (UE_i) is classified as *rushing* if it has been under-served during TW_i . More precisely, every TTI the REC checks for each UE_i if $(TW_i - t_{now,i}) \leq (QoS_i - tx_{data,i})/R_{min}$. Where $t_{now,i}$ is the elapsed time since the beginning of TW_i , QoS_i the QoS requirements of the UE class of traffic, $tx_{data,i}$ the total data transmitted by user i during $(TW_i - t_{now,i})$ and R_{min} the minimum transmission rate of the system. Note that Th_{rush} , η and TW_i are scheduler design parameters.

Step 2: Resources (chunks) are allocated to *rushing entities* with an EDF-like scheduler which allocates *best* chunk(s) to entities with higher deadline priority. Deadline priority

metrics differ between RT and NRT traffics: while with RT traffic deadline priority depends on R_{TTL} , with NRT traffic it depends on the lack of data transmitted in TW_i . Again, chunks are selected in order to maximize the spectral efficiency.

Step 3: All unscheduled resources (chunks) are assigned to users which maximize spectral efficiency, regardless to any QoS constraints. Priority is granted indeed to UEs selected by standard MCI 'matrix-based chunk allocation' described in [19] with $\lambda_{k,n}(i) = R_{k,n}(i)$. Actually, the matrix contains the metrics $\lambda_{k,n}(i)$ of all possible user-chunk pairs (only for previously unscheduled chunks), where at time *i*, UE *k* has a metric for chunk *n* which is given by $\lambda_{k,n}(i) =$ $R_{k,n}(i)$. $R_{k,n}(i)$ is the instantaneous supportable rate for UE *k* at chunk *n*, depending on each UE's CQI. At each time *i*, given an unscheduled chunk *n*, the scheduled UE is $U_n(i) = \operatorname{argmax} \lambda_{k,n}(i)$. Indeed, AMC suggests for each UE_k assigned to chunk *n* the most spectral efficient MCS, $MCS_{k,n}^{*}$, which meets PER_{target} .

Step 4: We trade UE's throughput with downlink power by limiting the power budget on each *non-rushing* allocated chunk. A target low power MCS is fixed: $MCS_{low-power}$. Such $MCS_{low-power}$ can be for instance suggested by the admission control based on the actual traffic load of the cell. Then, for each UE_k assigned to chunk *n*, the scheduler scales down $MCS_{k,n}^*$ to low power MCS ($MCS_{low-power}$) only if $MCS_{k,n}^* > MCS_{low-power}$. Downlink transmission power is indeed reduced for chunks allocated to *non urgent* UEs.

V. SIMULATION RESULTS

In this section we assess the effectiveness of our proposed GAS algorithm (see section IV) comparing it to two scheduling algorithms often investigated in the literature: MCI, and EDF. Schedulers are compared in different traffic load scenarios in terms of the following power efficiency metric Γ_i :

$$\Gamma_i = \sum_{j=1}^{K_i^{UE}} \frac{N b_{i,j}^{chk} \cdot Pow_{i,j}^{dl}}{\rho_{i,j}} \tag{1}$$

where at TTI *i*, K_i^{UE} , $Nb_{i,j}^{chk}$, $Pow_{i,j}^{dl}$ and $\rho_{i,j}$ are respectively the number of active UEs, number of chunk allocated to user *j*, the downlink power for transmission on each chunk and the throughput of user *j*.

Schedulers are compared in two traffic scenarios:

Scenario A (*single traffic* scenario): unique traffic type in the cell for all UEs.

Scenario B (*mixed real-time traffic* scenario): coexistence of VoIP and NRTV traffic in the same cell.

Simulation results are given for the system and traffic models presented in section II. Results are averaged over 10 independent dynamic runs, where at the beginning of each run UEs are randomly uniformly located in the central cell. Each run simulates 50 seconds of network activity and at each TTI channel realizations are updated. We check if QoS is met based on the metrics defined in [13][14]. Real-time VoIP and NRTV UEs are satisfied if more than 95% of the UEs have a

residual BLER is below 2% and their respective transfer delay is below 50ms and 100ms.

In figure 1 we show our simulation results for scenarios A with single VoIP traffic and we mimic a number of active UEs in the cell which ranges between 20 to 540. Four scheduling algorithms are investigated under this scenario: EDF (reddashed curve), MCI (blue-dashed curve) and GAS(green-solid curve). Note that 540 satisfied VoIP UEs is the limit with any of the investigated schedulers. The highest system load (540 satisfied active UEs) is achieved with EDF and GAS only. MCI can satisfy up to 450 coexistent active UEs. Such maximum load gap between EDF and MCI is not surprising. Actually, since MCI allocates resources to UEs with better momentary link quality regardless of the user time QoS constraints, with the increasing number of real-time flows, many users may face momentary service starvation and consequently, exceed the maximum delivery delay (50 msec). This is not the case with EDF since it allocated best chunk(s) to entities with higher QoS deadline priority. From an energy efficiency point of view, we observe that EDF performs slightly better than MCI. Actually, UEs served with MCI experience an higher average transmission spectral efficiency. This causes an higher energy loss caused by the transmission of padding bits added when a UE has not enough data to entirely fill the allocated chunk(s). Our simulation results show how GAS algorithm permits to notably improve the schedulers' energy effectiveness. We observe that in low load scenarios (20 active UEs), the Γ reduction is up to 32.5%, and it decreases as expected for higher number of active UEs, having up to 31%, 29%, 26% and 21% of energy cost reduction with respectively 50,100, 200 and 350 (up to 550) active UEs.

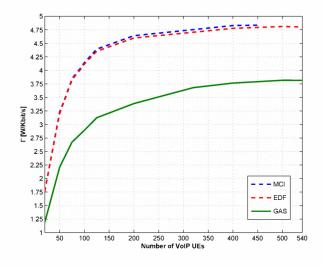


Fig. 1. Scenario A (single traffic): Γ performance evaluation with EDF, MCI and GAS schedulers under different load of satisfied VoIP UEs.

In figure 2 we show our simulation results for scenario B, when 75 NRTV UEs are active and the number of active VoIP UEs ranges between 0 and 250. This is a more challenging scenario since the system is heavily loaded and heterogeneous QoS constraints have to be taken into account in the scheduling rule. As in figure 1, we compare Γ performance of EDF (reddashed curve), MCI (blue dotted curve) and GAS (green-solid curve). While MCI cannot satisfy any additional VoIP UEs when 75 NRTV UEs are present in the system (therefore only a dot is represented on the figure), EDF and GAS can satisfy respectively up to 220 and 250 additional VoIP UEs. Simulation results show that comparing EDF and GAS, GAS improves Γ performance respectively of 18%, 16%, 10.5%, with respectively up to 50, 150 and 200 VoIP active UEs. With 250 VoIP UE only GAS can satisfy all active UEs since it is more flexible and effective in mixed traffic scenarios.

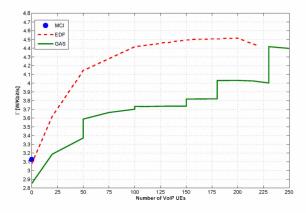


Fig. 2. Scenario B (mixed traffic): Γ performance evaluation with EDF, MCI and GAS schedulers with 75 UEs and a different load of satisfied VoIP UEs.

Eventually, in figure 3 we summarize the energy saving achieved with the proposed GAS scheduler for the different investigated load scenarios. We compare in terms of $\Delta\Gamma = \Gamma_{sota}/\Gamma_{gas}$, where Γ_{sota} and Γ_{gas} are respectively the power efficiency of state of the art schedulers (EDF, MCI) and GAS. We compare indeed, EDF with GAS (red-dashed curve) and MCI with GAS (blue-solid curve) in single VoIP scenario; we also compare EDF with GAS (green-dash-dotted curve) in real-time mixed traffic scenarios with 75 NRTV UEs. as shown on the figure the performance gain is, as expected, higher for higher traffic scenarios.

VI. CONCLUSIONS

In this paper we first investigated downlink power effectiveness of two reference scheduling algorithms MCI and EDF, in a systems that implements a realistic OFDMA air interface based on the 3GPP/LTE downlink specifications where no fullqueued traffic and unlimited number of control channel per TTI is assumed. We focus our investigation on single real-time and heterogeneous real-time traffic scenarios. We underlined that while EDF does not profit from multi-user diversity, MCI targets at maximizing the cell throughput regardless of the user's QoS constraints even in low load scenarios for which spectral efficiency can be trade with energy efficiency. Then, we come out with the definition of a novel scheduler, the GAS algorithm. GAS splits the resource allocation process in four steps: first, it identifies which entities (UE or packets) must be scheduled with high priority; second, it deals with rushing entities; third, remaining resources (if any) are assigned to

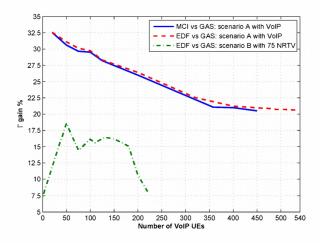


Fig. 3. Scenarios A and B: $\Delta\Gamma$ performance evaluation with EDF, MCI and GAS schedulers for different load scenarios.

users with highest momentary throughput. Fourth, downlink transmission power is reduced for chunks allocated in the third step. We evaluate the effectiveness of the proposed GAS algorithm comparing it with the above reference schedulers. Our simulations substantiate how GAS is a highly flexible and effective scheduler for a variety of traffic scenarios and it drives to notable energy cost reductions while meeting the Quality of Service (QoS) of UEs active admitted in the system.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement n°247733.

REFERENCES

- M. Gruber and O. Blume and D. Ferling and D. Zeller and M.A. Imran and E. Calvanese Strinati, "EARTH - Energy Aware Radio and Network Technologies", IEEE PIMRC workshop IOFC, Tokyo, Japan, September 2009.
- [2] A. Pokhariyal and T.E. Kolding and P.E Mogensen, "Performance of Downlink Frequency Domain Packet Scheduling for the UTRAN Long Term Evolution", in proceeding of IEEE PIMRC, Helsinki, Finland, September 2006.
- [3] K. Norlund and T. Ottosson and A. Brunstrom, "Fairness measures for best effort traffic in wireless networks", Proceedings of IEEE PIMRC, September 2004.
- [4] F. M. Chiussi and V. Sivaraman, "Achieving High Utilization in Guaranteed Services Networks using Early-Deadline-First Scheduling", 6th Int'l. Wksp. QoS, 18-20 May1998, pp. 209-217.
- [5] M. Andrews and K. Kumaran and K. Ramanan and A. Stolyar and P. Whiting and R. Vijayakumar, "Providing Quality of Service over a Shared Wireless Link", IEEE Communications Magazine, February 2001, pp.150-154.
- [6] R. G. Gallager,"Power limited channels: Coding, multiaccess, and spread spectrum," in Proc. Conf. Inform. Sci. and Syst., vol. 1, Mar. 1988.
- [7] S. Verdu, "On Channel Capacity per Unit Cost," IEEE Trans. Inf. Theory., vol. 36, pp. 1019-1030, Sept. 1990.
- [8] C. Schurgers and O. Aberthorne and M. B. Srivastava, "Modulation Scaling for Energy Aware Communication Systems", Proc. ISLPED, Huntington Beach, CA, August 6-7, 2001.
- [9] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, "An energy-effcient approach to power control and receiver design in wireless networks," IEEE Trans. Commun., vol. 5, pp. 3306-3315, Nov. 2006.
- [10] G. Miao and N. Himayat and Y. Li, "Energy-efficient transmission in frequencyselective channels," in Proc. IEEE Globecom 2008, pp. 1-5, Dec. 2008.

- [11] G.-W. Miao and N. Himayat and Y. (G.) Li and D. Bormann, "Energy efficient design in wireless OFDMA," Proc. IEEE 2008 International Conference on Communications, Beijing, China, May 2008.
- [12] Shuguang Cui and Andrea J. Goldsmith and Ahmad Bahai, "Energyconstrained modulation optimization for coded systems", in Proc. IEEE Globecom (2003).
- [13] 3GPP TSG RAN, "3GPP TR.25814, Physical Layer Aspects for Evolved UTRA (Release7)", v7.1.0 (2006-09).
- [14] 3GPP TSG-RAN1#48, Orange Labs, China Mobile, KPN, NTT Do-CoMo, Sprint, T-Mobile, Vodafone and Telecom Italia, "R1-070674, LTE physical layer framework for performance verification", 12th-16th February 2007, St Louis, USA.
- [15] 3GPP TSG-RAN, "3GPP TS36.211, Physical Channels and Modulation (Release 8)", v8.1.0 (2007-11).
- [16] K. Brueninghaus et al., "Link Performance Models for System Level Simulations of Broadband radio Access Systems", in proceedings of IEEE PIMRC 2005, Berlin, Germany, September 2005.
- [17] T. Henttonen et al., "Performance of VoIP with Mobility in UTRA Long Term Evolution", in proceedings of IEEE VTC spring 2008, Singapore, May 2008.
- [18] 3GPP TSG-RAN1#44bis, Motorola, "R1-060877, Frequency Domain Scheduling for E-UTRA", 27th-31st March 2006, Athens, Greece.
- [19] V. Ramachandran et al., "Frequency Selective OFDMA Scheduler with Limited Feedback", in proceedings of IEEE WCNC 2008, Las Vegas, USA, April 2008.