Uplink Scheduling in LTE and LTE-Advanced: Tutorial, Survey and Evaluation Framework

Najah Abu-Ali, *Member, IEEE,* Abd-Elhamid M. Taha, *Senior Member, IEEE,* Mohamed Salah, and Hossam Hassanein, *Senior Member, IEEE.*

***Abstract*—The choice of OFDM-based multi-carrier access techniques for LTE marked a fundamental and farsighted parting from preceding 3GPP networks. With OFDMA in the downlink and SC-FDMA in the uplink, LTE possesses a robust and adaptive multiple access scheme that facilitates many physical layer enhancements. Despite this flexibility, scheduling in LTE is a challenging functionality to design, especially in the uplink. Resource allocation in LTE is made complex, especially when considering its target packet-based services and mobility profiles, both current and emerging, in addition to the use of several physical layer enhancements. In this paper, we offer a tutorial on scheduling in LTE and its successor LTE-Advanced. We also survey representative schemes in the literature that have addressed the scheduling problem, and offer an evaluation methodology to be used as a basis for comparison between scheduling proposals in the literature.**

***Index Terms*—3GPP, LTE, LTE-Advanced, scheduling, SC-FDMA, carrier aggregation, coordinated multi-point transmission/reception, non-contiguous SC-FDMA, machine type communication.**

I. Introduction

THE INTRODUCTION of Long Term Evolution (LTE) came with a fundamental decision to change the multiple access technology for the network from W/CDMA to OFDM-based access. Orthogonal Frequency Division Multiple Access (OFDMA) was chosen for the LTE's downlink, while an OFDM variant, the Single-Carrier FDMA (SC-FDMA) was chosen for the uplink. Advantages of such multi-carrier access techniques include their robust communication and stable interference management. The techniques also facilitate dynamic frequency reuse techniques, e.g., soft frequency reuse. In addition, they allow exploiting multiuser diversity at granularities smaller than those possible in CDMA-based networks. Another advantage of multicarrier techniques is enhancing system throughput by mitigating the frequency-selective randomness, that is, frequency selective fading. This enhancement is achieved by modulating orthogonal subcarriers, and allows supporting different levels of user mobility and withstanding different communication conditions. More fundamentally, multi-carrier access techniques made viable numerous enhancements in LTE and its evolution, LTE-Advanced. These include advanced MIMO techniques, Carrier Aggregation (CA), Coordinated Multi-Point transmission/reception (CoMP) and relaying. But while OFDM substantially enhances physical layer capabilities, it has introduced considerable challenges when it comes to designing Radio Resource Management (RRM) functionalities such as packet scheduling. With LTE billed mainly as a packet-switched, IP-delivering and Quality of Service (QoS) maintaining network, LTE scheduling plays a crucial role as it manages the limited radio resources at LTE's access level in a way that maximizes the systems performance. Scheduling in LTE is located at the base station, termed evolved NodeB (eNB), and operates within LTE's MAC layer. The functionality is responsible for allocating shared radio resources among mobile User Equipments (UEs). The intelligence of the eNB packet scheduler is largely associated with its prompt awareness of network conditions such as wireless channel quality and the QoS experienced by the diverse Internet services running over the LTE interface. Scheduling LTE's uplink requires particular attention. In addition to traditional considerations for UE limited power budget, satisfying QoS requirements, and enhancing the throughput vs. fairness trade-off, SC-FDMAs advantage of low power requirements is largely realized when resource contiguity is enforced in the allocations made to a single UE. Such unique constraint has garnered great interest since 3GPP began deliberations on LTE in 2004, as will be showcased later in this work. In addition, since LTE was ratified in 2008, numerous enhancements have been introduced in succeeding releases. These enhancements include the possibility of non-contiguous allocations in SC-FDMA (Release 9). In the releases for LTE-Advanced (Release 10 and beyond), enhancements such as CA and CoMP were introduced. In CA, LTE nominal carriers can be aggregated into larger clusters to achieve higher data rates. Meanwhile, CoMP facilitates involving multiple eNBs in transmission both to and from UEs. LTE-Advanced also bear the introduction of certain services such as Machine Type Communications (MTC), which will substantially impact the uplink scheduling operation. Our interest in this work is threefold.

1) Offer a concise tutorial on scheduling in LTE and LTE-Advanced networks.
2) Survey representative scheduling proposals in the literature.
3) Propose an evaluation methodology for uplink scheduling proposals in LTE and LTE-Advanced.

Manuscript received August 3, 2012; revised March 1, 2013, July 13, 2013, and October 22, 2013.

N. Abu-Ali is with the College of Information Technology at UAE University, UAE (e-mail: najah@uaeu.ac.ae).

A-E M. Taha is with the College of Engineering at Alfaisal University, KSA (e-mail:ataha@alfaisal.edu).

M. Salah is with Alcatel-Lucent, Canada (e-mail: mohamed.salah@alcatel-lucent.com).

H. Hassanein is with the School of Computing at Queen's University, Canada (e-mail: hossam@cs.queensu.ca).

Digital Object Identifier 10.1109/SURV.2013.1127.00161

The tutorial begins with an overview of LTE, LTE-Advanced and their enhancements in Section II. It then describes the scheduling problem in both evolutions, where scheduling in LTE is described in Section III and Scheduling in LTE-Advanced in Section IV. In both sections, both general challenges and challenges specific to the different enhancements are discussed. A survey of the uplink schedulers is then made in Section V. The objective of the evaluation methodology proposed in Section VI is to unify evaluation approaches in conducting comparative performance evaluations of packet schedulers proposed for LTE and LTE-Advanced uplink transmission. While many previous proposals have been presented analytically and through simulation for downlink performance evaluation in downlink LTE or WiMAX, very few proposals have addressed performance evaluation in uplink LTE. For example, the work in [1] provides a survey for scheduling and interference mitigation algorithms in LTE, but does not offer any qualitative or quantitative evaluations. Similarly, the work in [2] surveys scheduling algorithms for generic wireless OFDMA systems (LTE's uplink is SC-FDMA) without evaluation. An analytical study of uplink scheduling schemes for LTE uplink was made in [3] using continuous-time Markov chains with states representing the change in the number of active users. The work analyzes only three algorithms without offering categorization, and it evaluates only the effect of the changing number of users, effectively limiting the applicability of the work for the packet-based architectures of 4G wireless networks. Finally, an evaluation is performed in [4]; it compares two proposed algorithms with three other schedulers using a Monte-Carlo simulation. The work does not offer categorization; rather, it focuses on aggregate cell throughput without consideration of a per-user throughput or for variations in the number and traffic of users. Results for a preliminary evaluation for some representative LTE schedulers are discussed in Section VII. Finally, Section VIII concludes.

II. Overview of LTE and LTE-Advanced

This section provides an overview of certain aspects in LTE and LTE-Advanced that pertain to the scheduling functionality, which will be discussed in the following section. The overview therefore spans network architecture, the air interface and frame structure, services and QoS, Hybrid ARQ, and physical layer enhancements in both LTE and LTE-Advanced. A brief guide on navigating through the relevant 3GPP specifications and standards is offered at the end of the section. Readers interested in a more detailed exposure can kindly refer to [5].

A. Network Architecture

The introduction of LTE, also called evolved Universal Terrestrial Radio Access (E-UTRA). (An LTE network is called an E-UTRA Network or E-UTRAN) marked two distinct differences in 3GPP network architectures. The first comprised placing substantial intelligence and independence at the Radio Access Network (RAN) level, in addition to a functional split between the data or user plane and the control plane. To achieve these differences, a deployed LTE or LTE-Advanced network, illustrated in Figure 1, comprises base stations (eNBs) that interconnect directly through an interface called the X2 interface, while eNBs connect to entities at LTE's core network, called Evolved Packet Core (EPC), like the Mobility Management Entities (MME) and Serving Gateways (S-GW) through an S1 interface. The direct X2 interconnection allows for substantially complex decisions, e.g., measurements, interference management and handover, to be executed directly at the access and without resorting to the entities at the core. Figure 2 elaborates on the functional split in LTE and LTE-Advanced, where the Packet Data Network Gateway (P-GW) is also shown. The highlighted functionalities in the eNB are ones that interact directly with the User Equipment (UE). On the user plane, the interaction involves the Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC), Medium Access Control (MAC) and the physical layer (PHY), while on the control plane it also involves the Radio Resource Control (RRC) and the Non-Access Stratum (NAS). The MAC objectives include mapping between logical and transport channels, multiplexing and scheduling MAC segments, relaying scheduling information, error correction (HARQ), and priority handling. The RLC performs correction (Automatic Repeat Control, or ARQ); concatenation, segmentation and reassembly for RLC segments, in addition to reordering and duplication detection. The PDCP mainly oversees ciphering and integrity protection, in addition to the transfer of control plane data. The RRC is the RAN component of the control plane and is responsible for the main control functionalities including broadcast of system information related to both the access and non-access planes, paging, establishing RRC connectivity between UE and the LTE network, security, mobility management, QoS and transfer of NAS messages. The NAS comprises all communications and signaling between the UE and the network core that are relayed by the eNBs. The UE corresponds to the core only through the MME. The NAS performs many tasks including bearer management; authentication; paging and mobility management when the UE is in the idle state; and security control. Functionalities of interest to this paper are the RRC, the RLC, and the MAC. We will rely on the descriptions of the PHY functionality in detailing certain aspects that pertain to the air interface and frame structure, in addition to certain constraints and enhancements that affect the scheduling operation.

B. Air Interface and Frame Structure

LTE and LTE-Advanced use OFDM as the PHY modulation method, employing OFDMA as the multiple access scheme for the downlink, and Single Carrier-FDMA (SC-FDMA) for the uplink. Both LTE and LTE-Advanced support TDD and FDD duplexing modes. In FDD, different frequency bands are utilized for the downlink and uplink transmissions, while in TDD the downlink and uplink share the same frequency bands but are separated in time. All transmissions are organized into radio frames of 10 ms each, with each frame further divided into ten equally sized subframes. In turn, a subframe is divided into two equally sized slots of 0.5 ms. For both downlink and uplink, and independent of the duplexing mode utilized, the base LTE radio resource is defined as a time-frequency

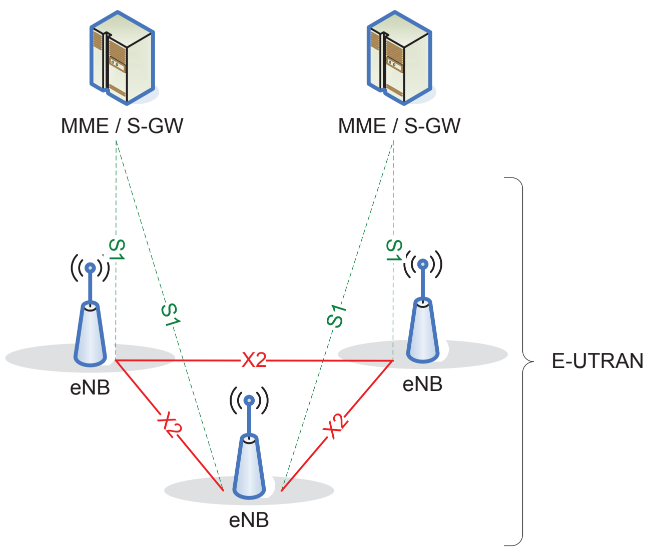


Fig. 1. Network architecture in LTE and LTE-Advanced.

resource block that spans 0.5 ms in time and 12 contiguous OFDMA/SC-FDMA subcarriers with a total bandwidth of 180 kHz. The resource block would use either six or seven OFDM symbols depending on whether a normal or extended (long) Cyclic Prefix (CP) is employed. The base LTE radio resource is also referred to as Physical Resource Block (PRB). LTE packet scheduling executes resource allocation decisions periodically once every 1 ms, which is defined as Transmission Time Interval (TTI). Hence, PRBs are always allocated in pairs that form a 180 kHz $\times$ 1 ms resource block. In the context of this paper and for all subsequent sections, the term PRB refers to the PRB pair just described.

C. Services and QoS

Achieving QoS for a certain application entails quantifying its requirements in terms of parameters that identify the target performance level. LTE's QoS framework is designed to provide an end-to-end QoS support. To achieve this, LTE provides QoS based on each flow requirements. LTE classifies flows into Guaranteed Bit Rate (GBR) and non-GBR flows. Flows in LTE are mapped into radio bearers which are the over-the-air connections. To accommodate end-to-end QoS, LTE differentiates between two types of radio bearers, S1 bearers and EPS bearers. An S1 bearer is a connection between an eNB and either the MME or the S-GW, while an EPS bearer is a connection between the EPS and the MME or S-GW, or the S-GW and the P-GW. There are two types of bearers in LTE, default bearers and dedicated bearer. The former, which is a non-GBR bearer that does not provide bit rate guarantees, is initiated and established at the startup time to carry all traffic. On the other hand, the latter can be either a GBR or a non-GBR bearer. If it is a GBR bearer, the UE can specify the guaranteed bit rate, packet delay and packet loss error rate. Each dedicated bearer is characterized by a Traffic Flow Template (TFT) with QoS parameters associated to it. An uplink TFT is used to map the UE uplink Service Data Flow (SDF) to specific QoS parameters, with the mapping carried out at both the eNB and the UE. Mapping for the downlink TFT is carried out at the S-GW or the P-GW. LTE groups bearers into classes. Each class is identified by a scalar number called the QoS Class Identifier (QCI), which in turn identifies a group of QoS parameters describing the packet forwarding treatment in terms of priority, tolerated delay, and packet error rate. Packet forwarding treatment is enforced by allocating radio resources for bearers through scheduling. Below are the definitions of these major quantitative parameters.

1) **Throughput.** Characterized through the Guaranteed Bit Rate, Maximum Bit Rate and Aggregate Maximum Bit Rate.
   a) **The Guaranteed Bit Rate (GBR).** Network resources allocated based on GBR are fixed and do not change after bearer establishment or modification. This is hence a guaranteed service data flow.
   b) **The Maximum Bit Rate (MBR).** This parameter limits the bit rate that can be expected to be provided to GBR bearer, and is enforced by a network shaper to restrict the traffic to its maximum bit rate agreement.
   c) **Aggregate Maximum Bit Rate (AMBR).** This parameter is used for non-GBR flows, and has two types, (Access Point Name) APN-AMBR and UE-AMBR. The APN-AMBR (Access Point Name-AMBR) is a subscription parameter stored at the Home Subscriber Server (HSS) per APN. The HSS defines a QCI for each Packet Data Network (PDN) (identifiable by an individual PDN identifier) and an APN-AMBR for each Allocation and Retention Priority (ARP). The APN-AMBR parameter refers to the maximum bit rate that can be achieved by all non-GBR bearers and all PDN connections of this APN. This parameter is enforced by P-GW in the downlink and by both UE and P-GW in the uplink. The UE-AMBR parameter, on the other hand, refers to the maximum bit rate allowed for all non-GBR bearer aggregates for the respective UE. The parameter is enforced in both downlink and uplink.
2) **Delay.** Specified by the packet delay budget. LTE defines nine categories for delay, with 50 ms being tightest and 300 ms being the slackest. The latter value is used for delay tolerant applications.
3) **Packet Loss.** Defined as the Packet Error Loss Rate, and is similar to the packet delay budget in having nine categories with 10^{-6} being best and 10^{-2} being the worst.
4) **Priority.** Specified by the Allocation/Retention Priority (ARP) parameter, which is used to indicate the priority of both allocation and retention of the service data flow. The ARP dictates whether a bearer establishment/modification request can be accepted or rejected in the event of conflicts in demand for network resources. At the time of exceptional network resources limitations, such as handover, ARP can be used by the eNB to drop

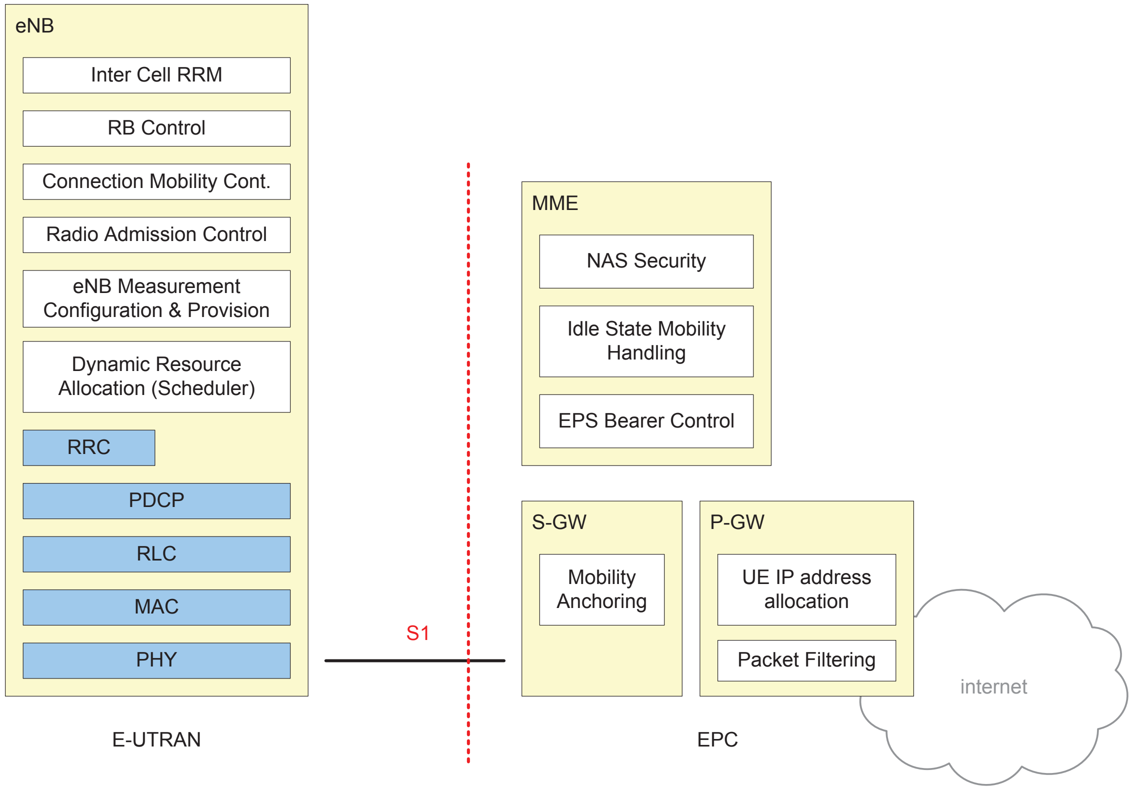


Fig. 2. Functional split in LTE and LTE-Advanced, showing eNB and the network core connected through the S1 interface.

a flow with a lower ARP to free up capacity. ARP, however, has no effect on the network treatment received by the flow once it is successfully established.

Note that GBR and MBR are defined per bearer while the AMBR parameters are defined per a group of bearers. All throughput parameters have two components, one for downlink and another for uplink. A default bearer is initiated and established at the startup time to carry all traffic. The default bearer is a non-GBR bearer, and does not provide bit rate guarantees. If a dedicated bearer is GBR, it can specify the guaranteed bit rate, packet delay and packet loss error rate. Each dedicated bearer is characterized by a TFT with QoS parameters associated to it. An uplink TFT is used to map the UE uplink traffic to specific QoS parameters, with the mapping carried out at both the eNB and the UE. Mapping for the downlink TFT is carried out at the S-GW or the P-GW. Table I gives an example of a traffic classification based on the QoS parameters defined in the LTE QoS framework. Each class is identified by a scalar number called the QoS Class Identifier (QCI), which identifies a group of QoS parameters describing the packet forwarding treatment in terms of priority, allowable delay, and packet error rate.

D. Hybrid ARQ

LTE provides two mechanisms of error detection and correction through re-transmission namely, the HARQ mechanism at the MAC layer and the ARQ at the RLC layer. The ARQ functions less frequently than the HARQ and handles errors not detected by the HARQ process. HARQ is designed to be simple and fast to improve the QoS performance. This improvement is achieved by reducing delay and increasing the system throughput through the fast retransmission. The feedback signal of HARQ is a one bit ACK/NACK and the HARQ can be sent at every TTI. HARQ is a stop-and-wait ARQ mechanism associated with the unicast transmission on the Uplink and Downlink Shared Channels (respectively, U-SCH and D-SCH). HARQ is not employed for broadcast and multicast traffic. To simplify the architecture, the HARQ functionality is terminated at the eNB, with the EPC, isolated from the HARQ procedures. For uplink transmission on the U-SCH, eNB decodes the transport block. If successfully decoded, the eNB sets the ACK bit in the synchronous feedback signal. The sender identifies the data transmission associated with this ACK signal from the Round Trip Time (RTT) and the timing of the feedback signal. Due to the synchronized feedback, no explicit numbering is required to identify the acknowledged data. Synchronous HARQ applied for uplink transmission is based on scheduling re/transmission of sub-frames at a predefined sequence of time instances. Subframes may be received out of order. Synchronous HARQ transmission is simplified by reducing the control signal overhead and the content of the feedback signal. Additionally, to expedite the HARQ operation, multiple HARQ processes can be concurrently employed for the uplink transmission.

E. Enhancements in LTE and LTE-Advanced

LTE is defined in 3GPP Release 8. In Release 9, enhancements were described for LTE that included allowing for non-

TABLE I
AN EXAMPLE OF QOS CLASSES IDENTIFIED BY THE QCI.

QCI	Resource Type	Priority	Packet Delay Budget	Packet Error Loss Rate	Example Services
1	GBR	2	100 ms	10^{-2}	Conversational Voice
2	GBR	4	150 ms	10^{-3}	Conversational Video (Live Streaming)
3	GBR	5	300 ms	10^{-6}	Non Conversational Video (Buffer and playback)
4	GBR	3	50 ms	10^{-3}	Real Time Gaming
5	Non- GBR	1	100 ms	10^{-6}	I MS Signaling
6	Non- GBR	7	150 ms	10^{-3}	Voice, Video, Interactive gaming
7	Non- GBR	6	300 ms	10^{-6}	Video (Buffer and playback)
8	Non- GBR	8	300 ms	10^{-6}	TCP Based
9	Non- GBR	9	300 ms	10^{-6}	

contiguous allocations in SC-FDMA (called non-contiguous or clustered SC-FDMA). LTE-Advanced, described in Release 10 and beyond, heavily relies on several technology advances to achieve its objective performance criteria, as set by the ITU-R open call in 2008. These advances included an expanded use of MIMO technologies in addition to the use of CA; in- and out-of-band small cells (Wi-Fi, femtocells); relaying; and a form of multi-cell MIMO called CoMP for both interference cancellation and transmission. LTE-Advanced also saw the introduction of new services including Location Based Services and Multimedia Broadcast and Multicast Services, and MTC. Our intent in this work is to focus on enhancements that substantially affect uplink scheduling. Accordingly, considerable attention has been given to clustered SC-FDMA , CA, CoMP and MTC. The effect of these enhancements will be discussed in detail in Section IV.

F. On Reading the 3GPP Standards

Descriptions provided in this work are derived from both 3GPP technical specifications and reports, in addition to interpretations found in industrial white papers and research papers. Understandably, navigating through 3GPPs standard space can pose a challenge to some readers. We hope the following may help in overcoming the challenge. LTE and its enhancements are respectively described in Release 8 and 9, while LTE-Advanced is described in Release 10 and beyond. For the latest in any release, the reader should consult http://www.3gpp.org/ftp/Specs/latest/. (A more detailed review of 3GPP releases can be found at http://www.3gpp.org/Releases) A good place to start in all access-relevant releases is TS 36.300 [6], which offers an overall description for both the access technology and network. Caution should be made in identifying the TS 36.300 for a specific release, e.g., carrier aggregation is discussed in 36.300 only after Release 10. Relevant to our topic herein, the following specifications and reports can be consulted.

- TS 36.201 describes PHY.
- TS 36.321 describes MAC.
- TS 36.322 describes RLC.
- TS 36.323 describes PDCP.
- TS 36.331 describes RRC.

An overview of enhancements to the LTE-Advanced physical layer, including CA, CoMP, relaying and support for release-heterogeneous deployments can be found in TR 36.814. The enhancements are further detailed in the following technical reports.

- TR 36.808 and TR 36.823: describe the CA operation[1].
- TR 36.819: describes the physical layer aspects of CoMP.
- TR 36.912: describes, among other things, how scheduling should be handled in CoMP.

Finally, MTC is discussed in TR 36.888

III. UPLINK SCHEDULING IN LTE

Scheduling for both the downlink and the uplink is part of the MAC layer at the eNB. 3GPP clearly states that downlink and uplink scheduling are two different entities, although it does allow integrated schedulers. Allocation of PRBs is made per UE and not per bearer or flow. The specific task of assigning the scheduled RBs to bearers is made by the UE by an uplink control function. Schedulers are engaged every TTI or at the end of periods spanning multiple TTIs. When a UE is scheduled over multiple TTIs further signaling may be required to specify information such as allocation time and allocation repetition factor. A UE requires the uplink to send both its user plane data (over Data Radio Bearers (DRBs)) and control plane data (over Signaling Radio Bearers (SRBs)). DRBs may carry either first-time transmissions or re-transmissions (HARQ/ARQ), while SRBs include the following (among other control signaling):

- **Scheduling Requests (SR).** Used to distinguish active users with data in their buffers from idle users. Each user with data pending in its buffer sends a one bit SR to the eNB informing of its need for an uplink grant to be scheduled.
- **Buffer Status Reports (BSR).** Informs the eNB the amount of data a UE needs to send. The report can be made either per group of (aggregated) radio bearers or per four groups.
- **Power Headroom Reports (PHR).** Informs the eNB about the available power at the UE for scheduling and RRM.
- **Sounding Reference Signal (SRS).** Introduced in uplink LTE as a wide-band reference signal, typically transmitted in the last SC-FDMA symbol of a 1 ms subframe, and is used to provide information on uplink channel quality. User data transmission is not allowed in this block, which results in an about 7% reduction in uplink capacity.
- **Channel Quality Indicator (CQI).** A measurement of the channel quality between the UE and the eNB.

[1]Several documents describe the CA, both inter- and intra-band, and in different bands. The reader is to consult http://www.3gpp.org/ftp/Specs/html-info/36-series.htm to distinguish operation in different settings.

- **Demodulation Reference Signal (DMRS).** Used for channel estimation when coherent detection and demodulation is needed. For the Physical Uplink Shared Channel (PUSCH), the DMRS has the same bandwidth as the uplink data transmission, and occupies the fourth SC-FDMA symbol for each uplink subframe. (In comparison, SRS has a potentially much larger transmission bandwidth, and is less often transmitted than the DMRS).

Channel dependent scheduling is possible in LTE, where the bit rate in an allocated PRB is determined by the Modulation and Coding Scheme (MCS). There are several ways in which the eNB can infer the channel condition. For example, CQI information transmitted by the UE can be directly mapped to particular MCS schemes. Alternatively, a UE can send an SRS, which can be used in estimating channel quality over all channels in a frame. Other methods involve the use of DMRS and estimating the Block Error Rate (BLER) through HARQ transmissions. Inevitably, the eNB decides upon an MCS scheme for each UE based on the inferred channel conditions. HARQ operation in the uplink is asynchronous, as opposed to the synchronous HARQ employed in the downlink. Accordingly, PRB assignments occur at predefined times between the UE and eNB (retransmission must retransmit after exactly 8 TTIs). If a grant is not provided for the UE for retransmission, then a non-adaptive HARQ retransmission would be attempted, without signaling. This entails no changes being made at the UE to attributes such as modulation order, code rate, and the resource allocation in the lifetime of a HARQ process. Constraining the uplink scheduler in terms of the total number of grants to be assigned per UE is the physical PRB availability, the quality of the radio link, and the transmit power budget. In order to maintain a low Peak-to-Average Power Ratio (PAPR) in SC-FDMA, PRB assignments made to a single UE must be contiguous. To elaborate, the transmit power applied by any given UE is equally distributed over the assigned PRBs. (Accordingly, a larger PRB assignments for a UE results in a lower transmit power per resource block.) Once the assignments have been decided, they are signaled by the eNB to the UEs on the Physical Downlink Control Channel (PDCCH) in the downlink frame, and which takes 1 to 3 symbols (8 to 51 bits). This naturally implies that the capacity of the PDCCH limits the number of UEs that can be scheduled in the uplink. The task of the uplink scheduler is therefore to decide upon the number of PRBs, their location, their MCS scheme, and the power with which each UE with traffic (be it data or control) is going to transmit. There are other operational aspects that may constrain the scheduling operation. For example, an out-of-synch UE would not be able to transmit. A UE employing Discontinuous Reception (DRX) in order to save power is also hard to schedule. Once the above is decided upon, the eNB relays this information to the UE through a scheduling grant. In turn, the UE decides upon which of its bearers gets served by how many PRBs in the grant that it is given. Once a UE receives its grant, it utilizes its uplink rate control function to assign the granted PRBs among its radio bearers. RRC controls this function by giving each bearer a priority and a prioritized bit rate (PBR). The UE serves its radio bearer(s) as follows.

1) First, all the radio bearer(s) are served in decreasing priority up to their PBR, then
2) If there are any remaining resources in the assigned grant, all the radio bearer(s) are served in decreasing priority until the grant is fully utilized.
3) Any remaining (unutilized) PRBs are padded.

In the case that all PBRs are set to 0, the first step is bypassed and the bearers are served in their strict priority order. Meanwhile, if more than one radio bearer has the same priority, the bearers will be served equally. Note that by limiting a UE's total grant, an eNB can ensure that the UE's AMBR is not exceeded. Finally, the eNB indicates on PDCCH whether the uplink grant is semi-persistent or non-persistent. A semi-persistent grant is one where the allocations may *persist* in following TTIs according to a periodicity dictated by the RRC. A non-persistent grant, on the other hand, is only made in single SR-grant cycles.

A. Enhancements in Release 9

In Release 9, a scheduling request prohibit timer was introduced. To elaborate on the necessity for a prohibit timer, further details need to be explained about the triggering and handling of a UE's scheduling request. According to Release 8, when new data arrives at an empty UE buffer, or data of higher priority than the data currently being served, a BSR is triggered so that the UE would report its newly changed buffer size. If the UE does not have sufficient uplink resources to transmit the BSR, an SR is triggered. SRs can be either sent on a Dedicated Scheduling Request (D-SR) channel , which requires UE-eNB synchronization, or through contending on the Random Access Channel (RACH) Scheduling Request (RA-SR). If a D-SR is made, a periodic resource will be assigned to allow in periods of either 5, 10, 20, 40 and 80 ms. If a grant is assigned by the eNB and received by the UE while the UE is awaiting a response to SR, the SR is considered canceled. During that time, the SR is considered pending. Suppose a fixed SR transmission opportunity of 5 ms in length. For an eNB to process an SR, it requires at least 4 ms. Accordingly, and with high probability, this automatically results in a needless transmission of at least another SR as the response to the first SR may not have been received, generating an unnecessary load on the Physical Uplink Control Channel (PUCCH) through which the SR signaling is carried. Release 9 therefore introduced an SR prohibit timer that prohibits the UE from sending another SR request until a response to the initial SR is received, or the timer runs out. Once the pending SR is received, the prohibit time is disabled.

B. Design Challenges

Except for defining two types of scheduling (semi-persistent and non-persistent), 3GPP does not specify how PRBs should be allocated, and the design of the scheduler is hence left open for vendor implementation. As with other systems, scheduling can be designed to take on one or more operational objectives, including:

- Maximizing spectral efficiency;
- Maximizing overall (aggregated) network throughput;

- Minimizing inter-cell interference; or
- Upholding specific fairness bounds/measures.

IV. Uplink Scheduling in LTE-Advanced

LTE-Advanced (defined in Releases 10 and beyond) relied upon several enabling technologies in order to achieve its operational objectives, including sustained 1 Gbps at no-to-low mobility levels, higher spectral efficiency inside the cell and at cell edges, and reduced deployment costs. LTE-Advanced also introduced the support of new services required to better the mobile user experience. In this section, we discuss three technologies that greatly impact uplink scheduling, namely, clustered SC-FDMA, CoMP, and CA. We also discuss the impact of a specific service which is MTC.

A. Clustered SC-FDMA

As noted above, maintaining a low PAPR in SC-FDMA mandated that allocations be made in uplink frame in a contiguous manner. This constraint considerably affects the performance of SC-FDMA compared to OFDMA. LTE-Advanced enhances the uplink multiple access by adopting Clustered SC-FDMA, which allows clustered (i.e., noncontiguous) groups of subcarriers to be allocated for transmission by a single UE in a single Component Carrier (CC). This allowance enables uplink frequency-selective scheduling while increasing the uplink spectral efficiency. It also enables simultaneously scheduling the PUSCH and the PUCCH together, which greatly reduces signaling latency. The introduction of Cluster SC-FDMA, however, introduces its own set of challenges. First, uplink schedulers designed for LTE may not be directly usable in LTE-Advanced. (The constraint now becomes non-contiguous allocation in a single CC, and is limited to two clusters so as to not adversely SC-FDMA's PAPR advantage.) Second, non-contiguous allocations directly result in increased PAPR, potentially leading to transmitter linearity issues. Third, the presence of multi-carrier signals increases opportunities for in-channel and adjacent-channel spur (spike) generation.

B. Coordinated Multi-Point Transmission

The use of OFDM-based multi-carrier access techniques makes LTE and LTE-Advanced networks inherently interference-limited. The main objective of utilizing CoMP in LTE-Advanced is to enhance the UE's throughput performance both at the cell edges and within the cell by mitigating inter-cell interference. This mitigation is achieved by coordinating transmissions and receptions over multiple cells, i.e., the cell serving the UE and its neighboring cells. In Release 11, little is noted on considerations to be made in uplink scheduling under CoMP. However, the following elaboration is offered on uplink CoMP reception [7]:

> Coordinated multi-point reception implies coordination among multiple, geographically separated points. Uplink CoMP reception can involve joint reception (JR) of the transmitted signal at multiple reception points and/or coordinated scheduling (CS) decisions among points to control interference and improve coverage:
>
> - **Joint Reception (JR).** PUSCH transmitted by the UE is received jointly at multiple points (all or some of the CoMP cooperating set) at a time, e.g., to improve the received signal quality
> - **Coordinated Scheduling and Beamforming (CS/CB).** user scheduling and precoding selection decisions are made with coordination among points corresponding to the CoMP cooperating set. Data is intended for one point only.

In the above, points mean a set of geographically co-located transmit antennas; sectors of the same site correspond to different points. To elaborate [8], Maximal Ratio Combining (MRC) is used in JR at multiple destination entities called Remote Radio Equipments (RREs). RREs can be connected to a central eNB. Full coordinated transmission or reception is achieved among RREs through a unified radio resource management at the central eNB. In turn, the central BS and the RREs are connected through the backhaul in a cell, e.g., optical fibers or microwave radio. Meanwhile, interference rejection combining is used where multiple UEs transmit the PUSCH simultaneously using the same RB. However, received weights are generated so that the received SINR or signal power after combining at the central eNB is maximized. The Minimum Mean Squared Error (MMSE) or Zero Forcing (ZF) algorithms are typically used to combine the received PUSCHs at multiple cell sites. Cell-edge user throughput is improved due to the increase in the received signal power. In CS/CB, only one UE transmits the PUSCH using an RB based on the coordinated scheduling among cells in CoMP reception. This arrangement directly improves the throughput for cell-edge users. As with regular scheduling, Release 11 does not specify the manner in which scheduling under CoMP should be approached. It does, however, elaborate on certain areas of possible enhancements [7]. In all these possibilities, listed below, coexistence with legacy UEs should be considered:

- The scheduler should allow for the possibility of coordination between different reception points or cells for receiving data and reference signals. (The manner in which this feature is permitted is not specified.)
- Certain enhancements to the PUCCH, the channel over which SRs and other control signals are sent, may be considered such as
  - Improving resource utilization efficiency;
  - Avoiding high inter-cell/point interference.
- Enhancements to the DMRS (applicable to both PUCCH and PUSCH) and SRS design, may be considered to
  - Increase capacity for the DMRS and SRS signals;
  - Improve the DMRS and SRS reception.
- The main role of power control in uplink LTE is to limit inter-cell interference while upholding minimum SINR requirements. 3GPP defines two types of power control: the Open-Loop fractional Power Control (OLPC), which depends on path loss distribution and the Closed-Loop fractional Power Control (CLPC) depends on measurements. Enhancements to both types may be considered, including for example:
  - Enhancement to support selection of intended reception point(s);

  - Path-loss determination and signaling that targets intended reception point(s).
- To ensure accurate reception of SRS at the coordinating points, further enhancements to the power control scheme for SRS may be considered.
- Enhancement for the uplink timing advance control to support efficient JR CoMP operation may be considered including possible enhancement on the Random Access Channel (RACH) transmission.

Clarifying the last point, the eNB advances the UE transmission timing so that uplink signals can be received time-aligned at the eNB. More specifically, in order to maintain uplink orthogonality between UEs in a cell, any timing misalignment should fall in the CP duration. The timing advance depends on the signal propagation delay, which depends on the distance between the eNB and the UE. With uplink CoMP, the UE signal is received by several eNBs and, in general, the distance from the UE to the eNBs is different. If the misalignment due to the different signal propagation delays is shorter than the CP duration, uplink orthogonality between UEs can be maintained even when a eNB receives signals from UEs of several cells. If the misalignment is larger than the CP duration, uplink orthogonality will be degraded and UEs signals will interfere with each other. This constraint imposes an upper limit on the potential distance between cooperating eNBs. Toward release 12, work has been initiated on what is now called "evolved-CoMP" where further enhancements are investigated to facilitate practical deployment. These enhancements include improving channel-state information feedback from the UE and broadening the practical applicability of CoMP solutions using relaxed backhauling requirements [9].

C. Designing Uplink Schedulers for CoMP

In scheduling the uplink in LTE-Advanced under CoMP, various operational aspects need to be taken into account. For example, synchronization of both cooperating and cooperatively served devices needs to be established in both time and frequency. At the same time, the scheduler needs to be designed to operate either under imperfect channel knowledge, or using accurate channel information that is obtainable given an acceptable level and delay of signaling. A reasonably constrained backhaul infrastructure should be assumed in designing practical schedulers. Backhaul latency can be categorized in one of three categories [10]:

- Minimal latency (in the order of μs) for eNB to RRE links
- Low latency (< 1 ms) associated with co-located cells or cells connected with fiber links and only limited number of routers in between
- Typical inter-cell latency associated with X2 interfaces.

X2 based backhaul links could be a bottleneck reducing a CoMP gain because of its low-capacity and high-latency characteristics. Uplink HARQ is based on synchronous re-transmissions, where a re-transmission can only be triggered in the subframe associated with the same HARQ process as the initial transmission [11]. The uplink re-transmission is triggered by a negative HARQ acknowledgement (NACK) or an uplink grant. In LTE FDD, these messages are transmitted 4 ms before the re-transmission or 4 ms after the initial transmission [12]. With uplink CoMP, the entire process of cooperative uplink reception hence needs to be completed 4 ms after the uplink transmission so that the serving BS can send a relevant HARQ feedback (or uplink grant) to the UE. This timing requirement is very challenging, and it is most likely not possible to be met if distant sites cooperate. Some of the noted enhancements are already underway. For example, to provide support of CoMP in LTE-Advanced the PDCCH requires considerable enhancement. 3GPP Release 11 is currently working on these enhancement of the PDCCH called enhanced PDCCH (EPDCCH), which improves the downlink control channel by supporting increased capacity for control signaling, additional carrier types, frequency domain inter-cell interference cancellation and beamforming and/or diversity [13].

D. Carrier Aggregation

Carrier Aggregation (CA), introduced in LTE-Advanced from Release 10, involves assigning more than one CC to the same UE in a single allocation. The feature enables achieving higher bit rates to users, in addition to increasing spectrum utilization. Under CA, A UE becomes associated with a single serving cell on the event of establishing or re-establishing an RRC connection called the primary serving cell (PCell). Depending on the network traffic load, the UE QoS requirements or other policy and deployment considerations, the UE may be associated with additional serving cells, called Secondary serving Cells (SCell), by establishing dedicated RRC connections. The UE is allocated a set of CCs and depending on its capability the set of CCs allocated may be contiguous or non-contiguous. CC allocated by the PCell are called PCC and those allocated by the SCell are called SCC. In downlink and uplink transmission and for better power management at the UE, the UE is typically receives on the PCC only. The reception on the SCC is enabled by the eNB when necessary (for example to meet QoS requirements) through signaling [14]. In either the downlink or the uplink, the semi-persistent scheduling defined for LTE-Advanced is limited to PCell only. This is logical since semi-persistent scheduling is meant for VoIP traffic, which has a low data rate requirements, and does not require resource aggregation over multiple CCs. Same-carrier scheduling is suitable when the UE is capable to receive PDCCH messages for each CC. In this case, control signaling is transmitted on the same CC as the corresponding data, where the eNB map uplink resource (PUSCH) and downlink resource (Physical Downlink Shared Channel, or PDSCH) and convey this information to the UE in a PDCCH message. In LTE-Advanced, to maintain backward compatibility, a single CC resource information is sent in a separate PDCCH message. Hence, a UE receives multiple separate PDCCH messages include downlink or uplink resources assignments per each scheduled CC. Cross-carrier scheduling becomes necessary when the UE is not configured to receive the PDCCH or not able to correctly receive the PDCCH on some CCs. In other words, cross-scheduling entails sending control signals to the terminal on CC and data on other(s). A straightforward example is the scenario faced when scheduling in heterogeneous networks, where interference level becomes unbearable for a subset of the CCs more than others. Specifically, in a heterogeneous network with, say low-power femtocells and high-power macrocells in an overlay, intercell interference becomes inevitable if both network levels are assigned the same carriers for operation, and traditional mitigation schemes such as fractional frequency reuse becomes ineffective. In such instances, cross-carrier scheduling offers a solution as the PDCCH can be transmitted over a clear carrier with tolerable (or eliminated) intercell interference. Another example, also in the context of heterogeneous deployment, can be in overlays employing both wide and narrow bandwidths, i.e., an LTE/2G overlay. Under CA, it becomes more efficient to send the PDCCH messages over the wide bandwidth CCs for two reasons: 1) while the narrow bandwidth CC has limited control channel resource, it has better frequency diversity; and 2) enabling cross-carrier scheduling simplifies the protection of the PDCCH transmission messages against the interference into only one CC. To enable control signaling using cross-carrier scheduling, a 3-bit Carrier Indicator Field (CIF) is included in the PDCCH messages to identify the CC that corresponds to the resource assignment contained in the PDCCH messages based. The PCC is normally numbered zero, while the SCCs are assigned a unique number. Scheduling and HARQ retransmissions are handled independently for each CC. Scheduling under CA proceeds as follows [15]

> When CA is configured, a UE may be scheduled over multiple serving cells simultaneously but at most one random access procedure shall be ongoing at any time. Cross-carrier scheduling with the CIF allows the PDCCH of a serving cell to schedule resources on another serving cell but with the following restrictions:
>
> - Cross-carrier scheduling does not apply to PCell i.e., PCell is always scheduled via its PDCCH;
> - When the PDCCH of an SCell is configured, cross-carrier scheduling does not apply to this SCell i.e., it is always scheduled via its PDCCH;
> - When the PDCCH of an SCell is not configured, cross-carrier scheduling applies and this SCell is always scheduled via the PDCCH of one other serving cell. A linking between uplink and downlink allows identifying the serving cell for which the downlink assignment or uplink grant applies when the CIF is not present:
>   - downlink assignment received on PCell corresponds to downlink transmission on PCell;
>   - uplink grant received on PCell corresponds to uplink transmission on PCell;
>   - downlink assignment received on SCell_n corresponds to downlink transmission on SCell_n;
>   - uplink grant received on SCell_n corresponds to uplink transmission on SCell_n. If SCell_n is not configured for uplink usage by the UE, the grant is ignored by the UE.

CA in Release 10 was limited to intra-band aggregation in the uplink. In Release 11 [16], the support expanded to inter-band aggregation. In addition, Release 11 describes multiple uplink timing advances, essentially enabling the use of non-co-located cells in CA. Finally, synchronization at the UE became achievable through the UE first synchronizing with the PCell, then seeking synchronization from SCells in neighboring sites. Accordingly, the PCell eNB will request a RACH for the synchronizing UE from the Scell after SCell activation through the PCell's PDCCH. 3GPP is moving toward expanding on CA band specification in Release 12, including definitions for new frequency bands for both intra- and inter-band aggregation scenarios. These include [17]:

- Five new intra-band scenarios will be defined:
  1) Band 1 (contiguous).
  2) Band 3 (non-contiguous), carried over from Release 11.
  3) Band 3 (contiguous).
  4) Band 4 (non-contiguous).
  5) Band 25 (non-contiguous).
- An additional eight inter-band scenarios will be defined
  1) Bands 3 and 5 with two uplink carriers.
  2) Bands 2 and 4.
  3) Bands 3 and 26.
  4) Bands 3 and 28.
  5) Bands 3 and 19.
  6) Bands 38 and 39.
  7) Bands 23 and 29.
  8) Bands 1 and 8.

Release 12 may also see the introduction of inter-site CA, which will allow a more efficient use of the fragmented spectrum from multiple cells in a heterogeneous setting.

E. Designing Uplink Schedulers for Carrier Aggregation

The choice to exploit CA in scheduling the uplink resources should be made with care [18]. For one thing, the capacity gain achieved in the uplink is less than that made in the downlink as it is dependent on the UE's limited power and the capability to utilize multi-carrier transmission. In fact, UE capabilities (intra-band CA, inter-band CA) need to be taken into account, and a sufficient number of CA-capable devices need to be operational before reasonable gains can be practically observed. User location would also affect the range of bands appropriate for the user, e.g., high-frequencies for users near the cell center and low frequencies for users at the cell-edge. At the same, only in scheduling a lower number of UEs using CA can a reasonable gain be achieved. A scheduler for CA needs to recognize that CCs resulting from CA are not backward compatible with LTE. This also entails somehow recognizing user capabilities prior to the scheduling process. Meanwhile, aggregations can be either contiguous or non-contiguous. However, an issue that emerges when simultaneously scheduling LTE and LTE-Advanced users is whether they should be scheduled jointly or independently. How can fairness be achieved, whether in static or dynamic allocations? More importantly, how can complexity be managed considering number and variety of users, and number, type, and availability of CCs? At the moment, only intra-band

CA is supported in the uplink, which means that the different carriers need to be part of the same frequency band, and have similar radio characteristics, at least in terms of path loss and shadowing. This limitation simplifies switching carriers on and off without the use of extensive measurements, but reduces the potential diversity. The limitation also means a simplified UE implementation. Only a single uplink timing advance value is supported for all CCs. Hence, the base station transceivers for different carriers should be at the same location to avoid different propagation delay. Consequently, the use of RREs, distributed antennas and repeater is limited by a distance that should be close enough to allow the signal to be received within the CP length for correct reception by a regular UE. Finally, a need exists in designing a new search space in PDCCH over which scheduling information can be passed to UEs, especially in the case of cross-carrier scheduling. In LTE, because PDCCH is shared by all UE, each UE monitors the control space to perform blind decoding in order to detect whether there is control information for itself or not. In LTE-Advanced, however, signaling in a certain PDCCH may specify allocations in a different CC. The need for a revised search space, however applies whether allocations are made over single or multiple CCs.

F. Machine Type Communications

Support for MTC in 3GPP began in Release 10. This support came in response to an increasing demand for wireless MTC. An inherent challenge in supporting MTC is that LTE, like its predecessor, was originally designed to carry communications between humans, or Human-to-Human communications. Especially LTE, which was designed to carry high data rates for broadband applications. Machines, however, send and receive small amounts of data and on varied connection generation rates. Such communication also entail redesign of certain protocols as the resulting control-to-data ratio increases. It also entails supporting a higher density of connections per cell that challenge the capacity of control channels such as PUCCH and PDCCH to relay scheduling requests and grants and measurements information. For example, only a maximum of 10 UE/MTC devices are allowed in a single subframe, which render the service of hundreds of MTC devices not feasible. Hence, simultaneous access to these resources is the solution for this problem. Further investigation is in progress. Furthermore, machines vary in their capabilities, e.g., energy, processing capabilities, transmission/reception capabilities, etc., in addition to their QoS requirements. As well, and relevant to this tutorial. MTC might pose a heavier demand on the uplink than on the downlink as most machines will be involved in sensing and measurement applications. For MTC (or Machine-to-Machine (M2M)) and Human-to-Human (H2H) communications to coexist, a resource allocation mechanism can either operate in an orthogonal manner where allocations for the both types are completely separated, or in a shared manner that maximizes allocation efficiency but does require interference management. Schedulers in LTE-Advanced will therefore need to manage five distinct types of communications[2]:

- The eNB-to-UE link (H2H).
- The eNB-to-MTC Device link (M2M).
- The eNB-to-MTC Gateway link (M2M relay or multihop).
- The MTC Gateway-to-MTC Device link (M2M relay or multihop).
- The MTC Device-to-MTC Device link (M2M).

Beyond Release 10, 3GPP expanded MTC support for low cost, low bandwidth devices by reducing their operating bandwidth [19]. For example, instead of operating at 20 MHz these devices can operate at 1.4MHz. This change, however, introduces the challenge of supporting these devices on the same carrier as other devices, including H2H devices. Solutions are currently deliberated to explore this possibility, including:

- **Using separate carriers.** A high bandwidth carrier could be fragmented into a set of narrower bandwidth carriers in order to support low bandwidth device. To support this mechanism, schedulers need to be designed to identify the allocation of such fragmented carriers to different classes of devices based on their capabilities. i.e., the allocated fragmented carrier should be less than or equal to the device capabilities. However, this method will segregate UE onto different carriers, which may degrade the performance of high capability devices, since the condition of operating on contiguous subcarriers in the uplink.
- **Using a virtual carrier.** PUSCH can be restricted to assigning only a restricted subset of subcarriers to the low bandwidth device forming a virtual carrier within the bandwidth of the conventional uplink carrier; this virtual carrier can be separately scheduled to low bandwidth devices. This method will require changes to the LTE standard by adding new logical channels and also affect the frame structure. However, it will allow the coexistence of the low cost devices with other LTE devices; i.e., network within the network.
- **Using carrier aggregation.** A high bandwidth Release 10 carrier can be fragmented into a set of lower bandwidth CCs, each of which is Release 8 compatible. One of these CCs could be used to support low bandwidth UEs while the CCs could be aggregated to form a higher bandwidth set in order to allow high peak rates to be provided to Release 10 devices in sub-frames where there is no MTC traffic. This scheme will allow Release 10 high performance UE to access the high bandwidth but high performance devices from release 8 and 9 will suffer the segregation effect of subcarriers due to the fact that they should operate on a contiguous subcarriers

V. A Survey of Uplink Schedulers for LTE and LTE-Advanced

In what follows, a survey of proposals for scheduling in LTE and LTE-Advanced is offered. Needless to say, the topic has amassed strong interest since first indications of the use of OFDM techniques in 4G networks, rendering it difficult to comprehensively account for all proposals. Accordingly, our intent in the following has been to select key representative

[2]Relay and multihop items are beyond the scope of this work.

TABLE II
COMPARISON TABLE FOR SURVEYED PROPOSALS. (EE: ENERGY EFFICIENT; SC: STANDARD-COMPLIANT; BC: BACKWARD COMPATIBLE).

	Basic or Enhanced	Allocation metric.	Optimized or Heuristic	QoS	EE	SC	BC
[4]	Basic	Attainable channel data rate, fairness.	Optimized.	BE	×	✓	✓
[21]	Basic	SNIR.	Heuristic.	BE	×	✓	✓
[22]	Basic	Fairness.	Optimized.	BE	×	✓	✓
[23]	Basic	Throughput maximization, fairness, queue stability.	Both.	BE	×	✓	✓
[24]	Basic	Generic.	Heuristic.	BE	×	✓	✓
[25]	Basic	Fairness.	Optimized.	BE	×	✓	✓
[26]	Basic	Fairness.	Optimized.	BE	×	✓	✓
[27]	Basic	Fairness.	Heuristic.	User data rate.	×	✓	✓
[28]	Basic	Fairness.	Heuristic.	User data rate.	×	✓	✓
[29]	Basic	Fairness.	Optimized.	Minimum and maximum user data rates.	×	✓	✓
[30]	Basic	Fairness.	Optimized.	User data rate, delay.	×	✓	✓
[31]	Basic	Attainable channel data rate.	Heuristic.	User data rate.	×	✓	✓
[32]	Basic	Fairness, SINR.	Both.	User data rate.	×	✓	✓
[33]	Basic	Fairness.	Optimized.	User data rate, delay.	×	✓	✓
[34]	Basic	Transmission power.	Sub optimal	Delay.	✓	✓	✓
[35]	Basic	Attainable channel data rate.	Optimized, simplified into greedy.	User data rate.	✓	✓	✓
[36]	Basic	Transmission time.	Heuristic.	BE	✓	✓	✓
[37]	Enhanced	Round robin.	Heuristic.	BE	×	✓	×
[38]	Enhanced	Reference Signal Received Power (RSRP).	Heuristic.	BE	×	×	×
[39]	Enhanced	Aggregate users data rate.	Optimized, simplified into greedy.	BE	×	✓	×
[40]	Enhanced	Channel quality, round robin.	Heuristic.	BE	✓	✓	✓
[41]	Enhanced	Channel quality.	Heuristic.	Delay.	×	✓	✓
[42]	Enhanced	Packet arrival rate.	Heuristic.	Jitter.	×	✓	✓
[43]	Enhanced	Probability of violating specific jitter.	Heuristic.	Jitter.	×	✓	✓
[44]	Enhanced	periodical(M/D/1) based on aggregated MTC data rate requirements.	Heuristic.	Packet delay budget, packet-drop rate.	×	✓	✓
[45]	Enhanced	Channel quality.	Heuristic.	BE	×	✓	✓

proposals with marks of a sufficiently unique contribution. With the focus of this work on uplink scheduling in LTE and LTE-Advanced networks, the reader is kindly referred to [20] for a survey on downlink schedulers. This survey spans four distinct groups of uplink schedulers. The first includes those aimed at basic LTE and LTE-Advanced networks, where no CoMP, CA and MTC are considered. The second includes proposals made for LTE-Advanced networks with CoMP; the third for CA; and the fourth for MTC. Efforts were made to include standard-compliant proposals, though some non-compliant proposals were included due to their novelty in other aspects.

Admittedly, there are several ways in which the surveyed proposals can be discussed, some of which are indicated in Table II. In the table, we have identified works targeting *basic* operation, i.e. regular point-to-multipoint mode without the use of enhancements or relay. We have also identified the metric driving the scheduling approach. As will be seen below, many proposal take into account a measure of fairness between bearers as a metric. On the other hand, throughput and attainable channel data rate are also accessible measures. Other metrics have also been employed such as transmission time, round robin and M/D/1 service. Proposals were also judged based on their optimality. To elaborate, several proposals utilized algorithms with proven optimality while others, acknowledging the complexity of optimal solutions, have proposed heuristics to facilitate practical implementation. The table further illustrates that not all schedulers take QoS into account, and that QoS prioritize different criteria. Most of the proposals are Best Effort (BE), while others have attempted to prioritize bearers based on data rates, delay, jitter or packet drop rate. Energy Efficiency (EE) was another sought criterion that was not addressed by many proposals. However, much of the proposals can be seen to be Standard Compliant (SC) and Backward Compatible (BC). To elaborate, standard compliancy refers to the fact that proposal was made per the 3GPP standards guidelines, while backward compatibility essentially refers to the viability of implementing the scheduler in both LTE and LTE-Advanced networks, with the latter possibly having enhancements. A careful consideration of Table II illuminates the fact a critical void persists in offering a comprehensive scheduler or set of schedulers for LTE and LTE-Advanced networks. Given the expanding set of factors involved (mobility; application requirements; device capabilities; PHY technologies; network capabilities; relays; CoMP; CA; MTC; etc.), all in addition to the particular requirements for scheduling OFDM-based access techniques, it becomes easier to grasp the scale of the scheduling problem at hand. The subsections ahead elaborate on the details of each of the four categories discussed above, with basic schedulers discussed firsts, followed by schedulers for CoMP, CA and MTC, respectively.

A. Basic LTE Schedulers

LTE uplink schedulers can be categorized into the following three categories: Best-Effort schedulers, QoS-Based Schedulers, and Power-Optimizing Schedulers.

1) Best-Effort Schedulers: Most of the proposals made for LTE uplink scheduling focus on maximizing performance metrics such as data throughput and fairness. Best-effort schedulers were designed in such a way that their main target is to maximize the utilization of the radio resources and/or the fairness of resource-sharing among UEs. Different metrics for fairness have been used, including min-max fairness and Proportional Fairness (PF). An example of a Best-Effort algorithm is demonstrated in [21]. The authors in [21] propose a set of three LTE uplink dynamic schedulers: 1) First Maximum Expansion (FME); 2) Recursive Maximum Expansion (RME); and 3) Minimum Area Difference (MAD). All three schedulers are proposed with the same PF utility function, but they differ mainly in the allocation scheme each function follows. In FME, the algorithm starts with finding the first UE-PRB pair with the maximum PF metric. Afterwards, the algorithm uses the PRB with the maximum metric as a starting point to expand the PRB allocation, starting with the first chosen UE, then continuing the PRB expansion process for other UEs. In the case of RME, the allocation algorithm starts in the same fashion as FME; that is, it finds the UE-PRB pair with the maximum metric and expands the allocation for the given UE until there are no more PRBs whose maximum metric belongs to the same UE. Once the algorithm finishes allocating PRBs to the first found UE, it then recursively performs the same procedure for the remaining UEs until all available PRBs are assigned. MAD algorithm, unlike the other two, is search-tree based. First, MAD scheduler derives a new per-PRB MAD metric for each UE, representing the difference between a UE's PF metric and the PRB's maximum metric among all UEs. Afterwards, the algorithm places the most likely UE-to-PRB allocation patterns into a tree. From here, MAD performs a Breadth First Search (BFS) algorithm which determines the desired UE-to-PRB allocation. Another work on PF-based schedulers is proposed in [22]. The authors introduce the concept of Fixed Transmission Bandwidth (FTB), where the scheduler groups the PRBs into equally sized Resource Chunks (RCs). The size of the RC depends on the scheduler's configuration. The authors propose two FTB-based scheduling algorithms, where one performs a linear search and the other a binary-tree search of possible UE-to-PRB allocation patterns. In [23], the authors propose a generic scheduling algorithm where scheduling policies can be easily integrated into the optimization problem after identifying the policy metric. As an example, the authors incorporated three policy metrics: proportional fairness, users queue length, and users data rate. A combinatorial optimization problem is defined and shown to be Strict NP (SNP) complex. A refined solution is then derived by deducing two heuristic greedy algorithms. An application of marginal utility was made in [4], where the authors propose a scheduling algorithm maximizing that of the UEs'. The utility is computed by taking the logarithm of the achievable transmission rate while catering to fairness. The proposed algorithm identifies a RB-UE pair with the highest marginal utility among all available RBs and UE. An exhaustive search is therefore needed to calculate all users marginal utility, which is a computationally expensive process. To overcome this complexity, the authors propose not to include all RBs; rather, the RBs are considered one at a time, identifying at for each RB the UE with the maximum marginal utility, until all RB are assigned. Both proposals are noted to not satisfy SC-FDMAs contiguity constraint. A different approach to the LTE uplink scheduler design is applied in [24], where the authors break down scheduling into two stages: first in frequency, then in time. For the first stage the authors propose a utility-driven scheduler where the utility is calculated for individual RBs in each UE's frame. The scheduler is generic by the design, and metrics as CQI, throughput or fairness can be utilized. In the second stage, incrementally expanded allocations are used to satisfy the contiguous allocation constraint. Accordingly, a UE is first allocated the RB with the highest utility, then an adjacent RB for which the same UE's utility is highest, and so on. This process is repeated until either the frame "edge" is reached, or the UE does not hold the highest priority in any adjacent RB. The proposal depends greatly on an exhaustive search, and may not applicable for time-constrained scheduling. Meanwhile, the allocations made are indifferent to individual user requirements. A UE can therefore be arbitrarily allocated more or less resources than it requires. Finally, unless fairness is included in computing the utility, it will not be applied. The authors in [25], [26] formulate the scheduling problem as a Nash bargaining solution to ensure proportional fairness among users while maximizing the total throughput of the network. This is done by minimizing the marginal utility of allocating the subcarriers. The proposed algorithm assumes limited feedback about the channel conditions from the users. The algorithm is claimed to work either in a distributed fashion, where it would be executed at each users device, or in a centralized manner where it would be executed at the BS. The authors analyze the complexity of the algorithm and prove that it can be considered within polynomial complexity. A drawback to this algorithm is that resources are allocated at the granularity of subcarrier.

2) Schedulers Optimizing QoS: Several works have investigated LTE uplink scheduling with QoS provisioning implemented as part of the schedulers' utility functions. An early contribution to QoS-based uplink scheduling is proposed in [27]. The authors propose Proportional Fairness with a Guaranteed Bit Rate (PFGBR) scheduling algorithm. The algorithm differentiates between UEs with QoS services from other UEs. The PFGBR algorithm assigns PF-based weights to UEs from both groups. In the case of UEs with QoS traffic, PFGBR uses a metric that combines PF with a term that is a function of Guaranteed Bit Rate (GBR), which increases their priority over non-QoS UEs. The work presented in [27] shows improved support for UEs with QoS requirements without starving UEs with BE traffic. Another QoS-based scheduler is investigated in [28], where the authors propose a GBR/PF LTE uplink scheduler combined with a QoS-aware Radio Access Control (RAC) algorithm. QoS provisioning is achieved by introducing a term that is a function of the UE's average throughput normalized by its GBR. The study shows that the proposed scheduler provides better support for QoS traffic streams, especially the ones with low GBR rate, such as VoIP services. In [29] a multiclass scheduler is proposed that is based on a generalized utility function and marginal utility. To support QoS and proportional fairness, the utility function accommodates different QoS measures, including required data rate and maximum permissible delay. The scheduler is modeled as a linear optimization problem, minimizing the opportunity cost function. The complexity and practicality of the algorithm is not investigated, and the resource allocation is at the granularity of subcarriers, which has two drawbacks. First, it violates the standard as the standard has the smallest resource unit for allocation is the PRB as optimized above. Second, optimizing the allocations at the subcarrier level increases the scheduler's complexity due to the large subcarrier space. Another work on QoS-Based scheduling was proposed in [30], where the authors proposed two QoS-based schedulers: Single-Carrier Scheduling Algorithm (SC-SA), and Multi-Carrier Scheduling Algorithm (MC-SA). The difference between the two schedulers is that SC-SA algorithm assigns each UE a maximum of one PRB if the number of schedulable UEs is larger than the number of available PRBs. MC-SA, on the other hand, assigns a UE either one or more PRBs based on the ratio between the UE's experienced throughput to its GBR. In [31], the authors propose two schedulers for the uplink and downlink transmission: TD and FD. The TD scheduler creates two candidate lists, one for GBR traffic, and the other for non-GBR. For GBR, bearers are classified into two sub-classes with two different priorities. The highest priority is enqueued in the list on top of the lowest priority. For non-GBR, the TD scheduler adds the entire non-GBR MAC-QoS class bearers into the candidate list. In each list and in each subclass the bearer candidates are prioritized based on their attained exponential average throughput weighted by the class weight. The FD scheduler starts with the GBR candidate list provided by the TD scheduler. The PRBs allocation is done iteratively based on the highest SINR. At the end of each iteration the achieved data rate of each bearer is calculated and checked if sufficient data is available in the bearer buffer to be served or if the sufficient guaranteed rate is achieved for that particular bearer. If the conditions are met, the bearer is scheduled and removed from the candidate list. Otherwise the bearer will remain as a candidate for the next iteration. The algorithm uses iterative search for allocating the best RB without constraining allocations to contiguous resource blocks for each user. Meanwhile, the work utilizes the required SINR as a QoS metric. In LTE, a user can only require a data rate, but not the SINR. Authors in [32] take an approach similar to that taken in [24]. Two algorithms are proposed. In the first, a gradient algorithm is used as the scheduling metric, and an integer-programming solution is offered. The gradient algorithm takes into account the channel conditions, proportional fairness and the data rate required by users. Noting the complexity of the first algorithm, the authors propose a heuristic algorithm based on the utility defined as the scheduling metric. The heuristic first assigns the RB with the highest utility to the UE, then assigns next RB to the UE with the next highest, and so on. To satisfy the contiguity constraint, the algorithm assigns RBs in between the two UE assignments to the first UE and continues this procedure until all RB are assigned. Unused RBs resulting from this procedure can be scattered in the frame, and may not have been assigned as allocations in this work are constrained by UEs data requirements. However, the can be allocated to pending HARQ retransmissions. (HARQ is assumed to be synchronous, non-adaptive). The design of the heuristic was noted to be highly dependent on the freshness and correctness of the CQI data, and is hence vulnerable to delays and errors in the channel quality measurements. After formulating the scheduling problem as an optimization of a utility function, the authors in [33] acknowledge the high complexity of an optimized solution. To facilitate practical implementation, the authors propose two algorithms: one greedy, the other achieving proportional fairness under delay constraints. In the greedy algorithm, the RB are assigned in a way that maximizes the marginal utility, satisfying maximum allowed delay and minimum throughput constraints for each user. Meanwhile, in the proportional fair algorithms the proportion between the current throughput to the total throughput is used as the marginal utility, and RBs are allocated to the users with the most critical delay requirement, given that the user has an above-threshold utility value. While the algorithm does consider delay, it does not offer operational bounds, especially as the work does not consider user data rate.

3) Schedulers Optimizing QoS and Power: The goal of the schedulers in this category is to reduce the power consumption of mobile UEs on uplink wireless transmissions. A scheduler in this category usually acquires some QoS aspects of the traffic flows transmitted on the LTE uplink, such as the packet delay budget or GBR requirements. In this case, the schedulers perform some algorithmic decisions to reduce the transmission power of a UE to a point where the UE still maintains its QoS requirements. The scheduling algorithm presented in [34] optimizes power consumption based on queue delay constraints. The concept of the scheduler here is to increase the experienced packet delay by a small amount while adhering to the UE's GBR and delay requirements. The work in [34] demonstrates that increasing the experienced packet delay can significantly reduce the amount of transmission power consumed on the uplink. Another proposal for Power-based LTE uplink scheduling is proposed in [35]. The scheduler design is based on binary integer programming. The scheduler starts with creating a matrix that represents all the allocation patterns possible with uplink PRBs that adhere to SC-FDMA's contiguity condition. The scheduler calculates the power allocation needed for every possible allocation pattern for each UE. Afterwards, the scheduler performs a greedy-based search algorithm to find the UE-PRB allocation pattern that minimizes the power expenditure on each UE while respecting their GBR requirements. In [36], a scheme is proposed that attempts to evaluate the energy consumption of a UE using FDMA against TDMA scheduling. The scheme investigates the efficiency of the manner in which scheduling is performed. That is, whether there is low transmission power and long transmission time, or high transmission power and short transmission time. They conclude that the TDMA scheme is more efficient than the FDMA scheme in reserving energy. In TDMA, all 48 RB in the 10 MHz channel are allocated for the user for specific time duration. Alternately, in FDMA, eight (8) RBs are allocated per user.

B. Schedulers for CoMP

Scheduling CoMP in the uplink is more challenging than in the downlink. At the time of writing this survey, not many proposal could be readily identified. It is expected, however, that more proposals will appear in the future. The authors in [37] present a round robin scheduler that support CoMP-JR for MU provisioning, with no considerations for supporting QoS guarantees and adaptive to medium conditions. Authors in [38] propose a scheduler that relies on a clustered operation, where a cluster comprises a group of CoMP cells. The authors further define measurement sets, which are cells within a cluster with a long term average Reference Signal Received Power (RSRP) received from serving cell and with a value less than a predefine threshold. This definition entails that frequency reuse of the RB within a cluster is allowed outside the measurement set, and that the number of cells in a measurement set can be less than or equal to the number of cells in a cluster. The scheduler works in two stages. In the first stage, all PRBs in the same cluster are considered available for the UE is within that cluster. An RB can be allocated to only one UE in a cluster. The scheduler allocates the same RB to UEs that are disjoint. This allows PRB reuse such that the frequency interference at the serving cell of the UE within the tolerable value. The work does not elaborate on how the interference threshold is defined, and does not offer means for guaranteeing QoS requirements. It also violates the 3GPP standard by locating the HARQ at the master cell of the cluster, and not the UEs serving cell. In [39], the authors designed a scheduler for the uplink CoMP-JR based on analyzing a combinatorial structure of the allocation problem in terms of user set partitions and permutations, and show that the optimal allocation under this setup is exponentially complex. They also propose a greedy algorithm where all users in the cell cluster are assigned cluster RB by pairing the users such that their instantaneous rate of users is maximized. The work does not consider QoS guarantees and, does not validate the claim for supporting a large number of cells and users.

C. Schedulers for Carrier Aggregation

Most work addressing scheduling for networks employing CA have focus the scheduling the downlink. An exception, however, can be found in [40], where the authors propose a scheduler that can support both LTE and LTE-Advanced UE. Release 8/9 users are allocated RB on the least loaded CC, with allocations made over only one CC. Meanwhile, allocations for LTE-Advanced users (Release 10 and beyond) are based on a newly defined metric that takes into account users path loss of the channel to decide if the UE is power limited. If so, the UE is assigned RBs over one CC; otherwise, the UE is assigned RBs over all available CC. This is to prevent power-limited UEs from causing underutilization in multiple CCs as their throughput will be decreased due to power limitation. At the same time, a UE will need to report CQI over each CC, which may lead to substantial overhead and degrade the performance in the uplink. Within each CC, resource allocation is made in a round robin manner. The work does not elaborate on handling QoS requirements.

D. Schedulers for Machine Type Communication

The uplink scheduler is very critical for MTC traffic since most of the traffic expected in an MTC network is in the uplink direction. The traffic is also expected to be bursty, with burst size and length directly proportional to the number of active machines, while the data itself per device is of a low rate. Literature addressing the uplink scheduling for MTC in LTE is scarce due to the problem's recency. In the following we survey most representative scheduler in this category: In [41], the authors propose two uplink scheduling algorithms that schedule MTC based on their delay tolerance. The algorithm first allocates UE traffic, then allocates the remaining RBs to MTC devices. In the first scheduler, RBs are ranked based on channel quality, then the RBs with the highest rank are assigned to machines with the least delay tolerance. Meanwhile, the second scheduler first rank devices based on their delay tolerance, then assigns the least delay-tolerant machine the RB with the best channel quality. The work does not elaborate on scalability, and whether serving devices based on their delay tolerance would allow for a sufficient number to be served each TTI. The authors in [42], [43] offer a clustering-based approach to scheduling machines in LTE-Advanced by clustering machines based on a their QoS requirements, with a scheduling period defined for each cluster and is adapted based on the clusters jitter. In adjusting the scheduling period, it is kept less than a certain pre-specified threshold for jitter. Co-existence with H2H communication is made possible by allowing the eNB to use the resources allocated to MTC devices with fixed probability. The work does not elaborate on certain aspects of the proposal such as whether scheduling H2H communications is made based on UE requirements, in addition to the granularity and formation of the clusters. The work in [44] builds on [42], [43] discussed above, by requiring periodical scheduling or semi-persistent scheduling for each group, and prioritizing allocations within each group based on delay tolerance. The authors employ an M/D/1 queuing system model that relates the scheduling period, average offered traffic load, and QoS requirements, in terms of packet delay budget and packet-dropped rate. They conclude by providing guidelines and recommendation for H2H and M2M mixed traffic scheduler design based on hybrid scheduler (full-dynamic/semi-persistent). The work does not elaborate on how UE/MTC resource estimation is made per group. The work also does not evaluate how the statistical QoS provisioning goodput per group was evaluated. It also does not discuss the granularity and formation of the clusters. In [45], a mixed scheduler for H2H and M2M communications and signaling is proposed. The scheduler allocates resources based on CQI at each TTI, and allocates H2H UEs before M2M devices, where residual RBs are assigned to MTC individually, not in groups, and in a manner that minimizes MTC connection rejection. The work does not provide QoS guarantees for MTC – only best effort and based on CQI.

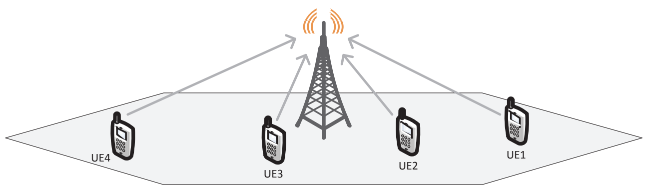


Fig. 3. System topology of LTE simulated LTE system

VI. An Evaluation Methodology

In what follows, we describe an evaluation methodology for uplink schedulers in LTE and LTE-Advanced. Figure 3 illustrates an LTE uplink system topology that can be assumed in simulation. The system represents the LTE uplink within a single-cell environment, with a cell coverage of 1km × 1km area. The cell's topology consists of one eNB at the center of the cell, with UEs located around the eNB in a spatially uniform pattern. Each UE within the system is equipped with a single, omnidirectional transmit antenna, while the eNB is equipped with two received antennas, resulting in a 1 × 2 Single-Input-Multiple-Output (SIMO) antenna configuration. No RAC needs to be implemented at the eNB, as the performance evaluation focuses on uplink scheduling. All UEs can be created at the beginning of each simulation run, and remain active for the entire simulation duration. The operating uplink bandwidth can be set to 5 MHz with an FDD configuration. The uplink bandwidth is hence divided into twenty-five PRBs as specified in the standard [46]. Twenty four PRBs constitute the PUSCH, which is used as a shared medium among UEs for data transmission. The remaining PRBs are assumed to be reserved for uplink control channels, which are chosen to be at the end of the spectrum to ensure contiguity of available PUSCH resources.

A. The Wireless Channel

The modeling of the wireless propagation channel is broken down into a macroscopic model and a microscopic model. The macroscopic channel model describes large-scale channel variations that depend on the UE's spatial position relative to the eNB, such as path loss, shadowing, and penetration loss. The macroscopic propagation model we chose here is to represent propagation path loss , PL, in the typical urban setup as found in [47], and it is expressed as

$$PL[\text{dB}] = I + \alpha \cdot 10\log_{10}(d[\text{km}]) + \chi_0 + P_{loss} \quad (1)$$

where I captures the free space propagation loss, which is equal to 128.1 dB. α is the path loss exponent, which is experimentally determined to be 3.76 in a typical urban environment. d is the distance between UE and eNB in km. χ_0 is a random variable that represents the shadowing effect in the typical urban environment, which follows a lognormal distribution with a mean of 0 dB and a standard deviation of 8 dB. P_{loss} is the penetration loss caused by signal penetration through obstacles, which is assumed to be a constant of 20 dB. The microscopic channel model mainly focuses on describing fast variations of the channel gain, which mainly entails multipath fading. Multi-path effect can either be slow fading or fast fading, depending on the UE mobility (i.e., the Doppler Effect). The microscopic channel model implemented the Tapped Delay Line (TDL) using a pedestrian, 12-tap Typical Urban (TU) Power Delay Profile (PDP) as described in [48]. Based on both the macroscopic and microscopic parts of the channel model, the CSI per SC-FDMA subcarrier (for each UE) is used to represent the path gain experienced by UE u at subcarrier k, $CSI_{u,k}$, and is determined as follows:

$$CSI_{u,k} = \frac{G_{UE} \cdot G_{eNB} \cdot (|H_{u,k,1}|^2 + |H_{u,k,2}|^2)}{PL_u \cdot \sigma_n^2 \mathcal{N}_{\mathcal{UE}} \Delta f} \quad (2)$$

where G_{UE} is the UE's antenna gain; G_{eNB} is the eNB's antenna gain; PL_u is the power loss experience by user u (equation (1)); σ_n^2 is the noise density per Hz; $\mathcal{N}_{\mathcal{UE}}$ is the noise figure of the receiver at eNB; and Δf is the subcarrier spacing. The terms $|H_{u,k,1}|^2$ and $|H_{u,k,2}|^2$ refer to the normalized multipath gains at eNB's receive antenna 1 and antenna 2, respectively. The summation of both gains is based on the transmit/receive diversity assumption stated earlier.

B. Link Adaptation

A Link Adaptation (LA) model can be used in to predict the appropriate MCS to use when transmitting the data on assigned resource blocks. Once the packet scheduler assigns UEs their corresponding PRBs, the effective SINR, γ_u, is determined for the assigned PRBs assuming a Minimum Mean Squared Error (MMSE) receiver [49],

$$\gamma_u = \left(\frac{1}{\frac{1}{N_u} \sum_{k=1}^{N_u} \frac{\gamma_{u,k}}{\gamma_{u,k}+1}} - 1 \right)^{-1} \quad (3)$$

where $\gamma_{u,k}$ is the SINR for UE u at subcarrier k, and N_u is the number of contiguous subcarriers assigned to UE u. $\gamma_{u,k}$ is determined per subcarrier k as follows

$$\gamma_{u,k} = \frac{P_u}{N_u} \cdot CSI_{u,k} \quad (4)$$

where P_u is the total transmit power assigned to UE u by the eNB.

C. Uplink Transmit Power

Determining the UE's total uplink transmit power is based on the Open Loop Power Control (OLPC) mechanism, where the LA process compensates for macroscopic effects experienced in the uplink channel between a UE and the eNB. Hence, the UE transmit power can be calculated according to [12]

$$P = \min\left(P_{t,MAX}, P_0 + 10\log_{10} N_{u,PRB} + \alpha PL_u\right)[dBm] \quad (5)$$

Whereas $P_{t,MAX}$ is the maximum UE transmission power set to 24 dBm as shown in Table III, $N_{u,PRB}$ is the number of PRBs allocated to UE u. PL_u is the path loss expressed in (1). P_0 and α are both cell-specific parameters, where P_0 is the base transmit power per PRB, and α is the fraction of path loss to be compensated for. The value of $P_{t,MAX}$ is set based on UE profile settings defined in Table 4.8 in [47]. For simplicity, UE's total uplink transmission power is assumed

TABLE III
SYSTEM SIMULATION PARAMETERS

Cellular Layout	Single-Cell with Omnidirectional Antenna
System Bandwidth	5 MHz
Carrier Frequency	2 GHz
Number of Resource Blocks	25
TTI Duration	1 ms
Path Loss Model	$128.1 + 37.6\log_{10}(d[\text{km}])$
Penetration Loss	20 dB
Shadowing	Lognormal: $\mu = 0$, $\sigma = 8$dB
Minimum Distance Between UE and Cell	90 m
Power Delay Profile	TU12 Profile, 12 taps
Channel Estimation	Ideal
MCS Settings	QPSK [1/6 1/4 1/3 1/2 2/3 3/4] 16QAM [1/2 2/3 3/4]
HARQ Process	OFF
eNB Antenna Gain	15 dBi
UE Antenna Gain	0 dBi
eNB Noise Figure	5 dB
Max. UE Transmit Power	24 dBm
Power Compensation	$P_0 = -58$ dBm, $\alpha = 0.6$
Frequency Re-use Factor	1
Simulation Time	10000 TTIs

to be allocated for data transmission only. Afterwards, the effective SINR calculated in (3) is mapped to pre-determined SINR ranges to ascertain the MCS to be used for data transmission. Based on the MCS selected and the number of PRBs assigned to the UE, the Transport Block (TB) size is calculated in bits. A TB is composed of the MAC packet data units, which includes the MAC header, RLC header, and the data payload, as well as a 24-bit Cyclic Redundancy Check (CRC) for error detection. Table III summarizes the simulation parameters discussed above.

D. *Traffic Model*

The traffic models developed for the LTE uplink simulator can be adopted from [50]. The QoS-based packet parameters for each traffic type are based on the QCI parameters illustrated in Table 6.1.7 in [51]. The Packet Delay Budget (PDB) values defined as part of the QCI parameters represent the maximum packet delay allowed between the UE and Packet Policy and Charging Enforcement Function (PCEF) in the network core. According to NOTE 1 in Table 6.1.7 in [51], the offset in packet delay between eNB and PCEF ranges between 10 ms up to 50 ms depending on how far away the PCEF is from the eNB to which the UE is associated. Accordingly, we chose to set the PDB offset to 50 ms so as to drive the access network to perform in tight delay scenarios where PCEF entity is furthest away from the access network. In addition, to better evaluate the performance of the scheduler and its distinction between the different traffic types, each UE can be assumed to carry a single traffic stream of a single traffic type.

1) VoIP Traffic: A possible VoIP traffic choice is the 12.2 kbps Adaptive Multi Rate (AMR) VoIP service with silence suppression [50]. In the active state, the VoIP source generates a 40 bytes VoIP packet once every 20 ms frame, with each packet consisting of 244 bits of payload data, and the remaining space consisting of the packet's overhead.

2) Video Streaming Traffic: Video streaming traffic is modeled as a low quality video stream running at a minimum guaranteed bit rate of 64 kbps [50]. When system capacity permits, the uplink video streaming traffic load can be allowed to increase per UE up to a maximum of 1024 kbps.

3) FTP Traffic: FTP traffic is generated with a CBR-based packet stream instead of the FTP traffic model described in the standard [50]. The CBR-based FTP traffic model helps create a fixed offered traffic load per UE to better map between the offered FTP traffic load per UE against its measured throughput. A constant packet size of 256 bytes with a traffic rate of 128 kbps per FTP connection can be chosen. MBR of the offered FTP traffic load per UE can be set to 1024 kbps. As a Best Effort traffic type, FTP has neither a GBR nor delay requirements associated with it. However, as some of the schedulers implemented here require a GBR value for their operation, such as PFGBR and MC-SA, a GBR of 10 kbps in this case can be assumed for each FTP connection. Table IV summarizes the simulation parameters of the traffic models discussed above.

TABLE IV
SUGGESTED PARAMETERS FOR TRAFFIC MODELS.

Traffic	Traffic Class	QCI#	PDB (ms)	GBR (kbps)	MBR (kbps)
VoIP	Conversational	1	50	12.2	64
Video Streaming	Streaming	2	100	64	512
FTP	Background	6	—	10	1024

E. *Performance Metrics*

The following metrics can be adopted to evaluate several aspects of the system performance under the operation of different uplink schedulers.

- **The cell's aggregate throughput.** Measured as
$$\bar{T}_{cell} = \frac{B}{t_{sim}} \quad (6)$$
Where B is the total number of bits successfully transmitted over the air interface from the UE to the eNB, t_{sim} is the total simulation time.
- **Intra-Class Fairness.** Represents the fairness among UEs of the same class. The Intra-class fairness is calculated using the min-max fairness index. Assuming UE i is the one with the maximum throughput, while UE j is the user with the minimum throughput, the inter-class fairness index can be calculated as:
$$F_{min-max} = \frac{\bar{T}_i}{\bar{T}_j} \quad (7)$$
where $\bar{T}_i$ is the throughput of UE i, and $\bar{T}_j$ is the throughput of UE j.
- **Inter-Class Fairness.** A measure of the fairness among UEs with different traffic classes. The measurement of the Inter-class fairness is performed using the well-known Jain's Index, which is calculated as follows:
$$F_{Jain} = \frac{\left|\sum_{i=1}^{N_{UE}} x_i\right|^2}{N_{UE}\sum_{i=1}^{N_{UE}} x_i^2} \quad (8)$$
where x_i represents the normalized average throughput of user i. To achieve the interclass fairness between different

QoS classes, the UE's average throughput is normalized with respect to the UE's MBR that is defined for each traffic type.

- **Packet Loss.** A measure of the percentage of packets of a certain traffic class dropped in regard to the transmission packet queue. This is due to exceeding their packet delay budget. The transmission buffer length in this experiment is assumed to be infinite in order to exclude packet drops due to packet buffer congestion.
- **Packet Delay.** The delay per packet is measured from the time the packet enters the RLC queue until the time it successfully arrives at the eNB. Packet delay is measured by collecting the delay stamps for all packets being sent within the entire simulation time, after which the experienced average delay of all UEs within a given traffic class is determined. The packet delay measurements, along with the measurements from the packet drops, can give an indication as to the ability of a scheduling algorithm to satisfy the QoS requirements of each traffic class with active transmissions.
- **TB Utilization.** TB utilization refers to the averaged percentage of TB size used for transmitting the data payload. TB in LTE refers to the PHY payload to be transmitted over the radio interface, which comprises the MAC packet plus a 24-bit CRC overhead. The TB utilization statistics are collected per UEs carrying traffic of the same QoS class. It is then averaged over the number of UEs belonging to that traffic class. TB utilization provides a measure on how well a scheduler predicts the needs of each UE in the system as well as its ability to minimize the resources wasted as a result from resource allocation mismanagement.

F. Evaluating Schedulers for CoMP

Environments for evaluating schedulers for CoMP must involve more than one cell/site. A site may include multiple cells that may cooperate. A scheduling proposal should be judged whether its clustering is static or dynamic. Compatibility with Release 8 also need to be investigated, which can be done with having both CoMP antennas as receiving and UE as transmitting. (In the case of enhancements, multi-user MIMO can be assumed). Unless synchronization between cells is being investigated, it should be assumed. At the same time, an acceptable backhaul delay should be accounted for. The traffic model should include burst traffic scenario. Meanwhile, the effectiveness of CoMP should be evaluated only for users at the cell edge. Lastly, either maximum ratio combining (joint scheduling) or interference rejection combining (coordinated scheduling) can be assumed.

G. Evaluating Schedulers for Carrier Aggregation

When evaluating schedulers proposed for uplink CA, the primary eNB should be the cell where the UE is always attached to. The UE should always be scheduled over that carrier. A user should be allocated CC(s) in secondary cells if there is a need to meet the UEs QoS requirements. If a user is not scheduled over Secondary CC(s), the UE will be disassociated by an expiring timer. Separate RRM blocks should be made to operate independently to guarantee backward compatibility with release 8. Whether the scheduler allocates multiple CCs for power limited UEs should also be investigated.

TABLE V

EXAMPLE COMPLEXITY LEVELS FOR SOME SCHEDULERS. (N_{UE} AND N_{PRB} ARE THE NUMBERS OF USER EQUIPMENT AND PHYSICAL RESOURCE BLOCKS, RESPECTIVELY.)

Scheduler	Complexity
RR	$O(1)$
RME	$O(N_{UE} \cdot N_{PRB})$
Greedy	$O(N_{UE} \cdot N_{PRB})$
GBR-ATB	$O(N_{UE} \cdot N_{PRB})$
PFGBR	$O(\log(N_{UE}) \cdot \log(N_{PRB}))$
MC-SA	$O(N_{PRB}^2)$
BMTP	$O(N_{UE} \cdot N_{PRB}^2)$

H. Evaluating Schedulers for MTC

When evaluating schedulers for MTC, the use of a dedicated and diverse model for machine traffic becomes important. The simulator will also need to be amended with MTCs signaling. Two types of scenarios can be evaluated: 1) per device scheduling; and 2) per group scheduling. In the case of the latter, the efficiency of the grouping algorithm relative to similarity in QoS requirements should be evaluated. Meanwhile, semi-persistence scheduling can be assumed, whether over every TTI or over multiple TTIs, depending on the granularity of the grouping algorithm. Both single-cell and multi-cell environments can be assumed in the evaluation environment, although single-cell must always be assumed when evaluating Release 8 compatibility. Other considerations for testing MTC environments include optimizing CQI reporting to minimize number of CQI reports and alleviate congestion [19].

I. Complexity Analysis

The complexity analysis of an OFDM-based scheduler is based mainly on the number of iterations a scheduling algorithm takes in searching for the final UE-to-PRB mapping. The search algorithm is the one that consumes most of the scheduler's total operation time in comparison to the computation of the utility-based metrics. Table V provides the complexity for some of the surveyed schedulers. Note that, especially when schedulers are designed for enhancements, further investigations need to be made into the overhead complexity, i.e., the information required for the scheduler to perform successful operation.

VII. EVALUATING UPLINK SCHEDULERS FOR LTE

An LTE uplink simulator is implemented using MATLAB to provide a unified, standard-compliant simulation setup. The simulator design focuses on having the necessary modules that fully capture the events occurring in the LTE uplink, and were implemented such that the simulator can be easily extendible to accommodate enhancements made in LTE-Advanced.

Representative uplink packet schedulers for LTE are simulated to provide a preliminary comparison between different scheduler design approaches. From the literature surveyed we

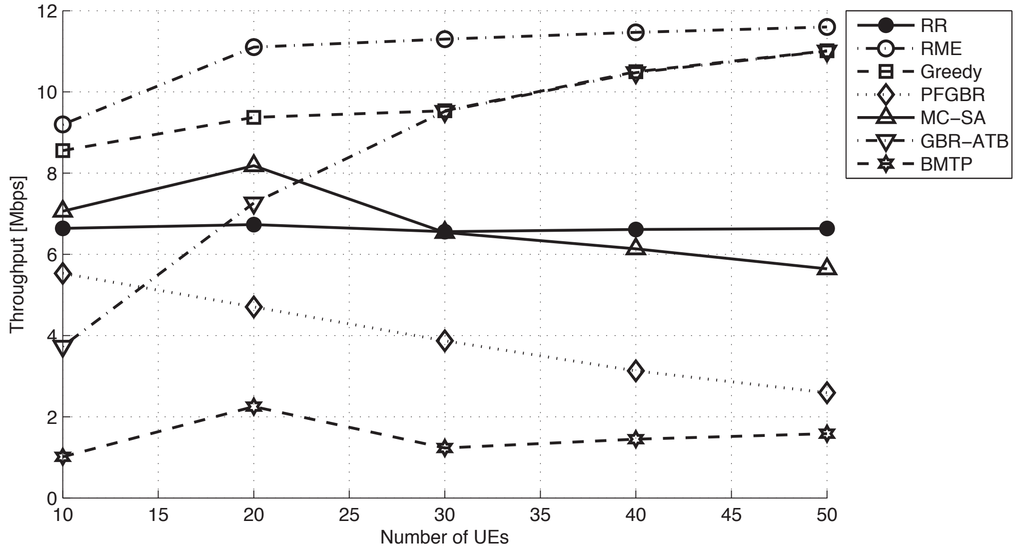


Fig. 4. Experiment 1:Aggregated Throughput

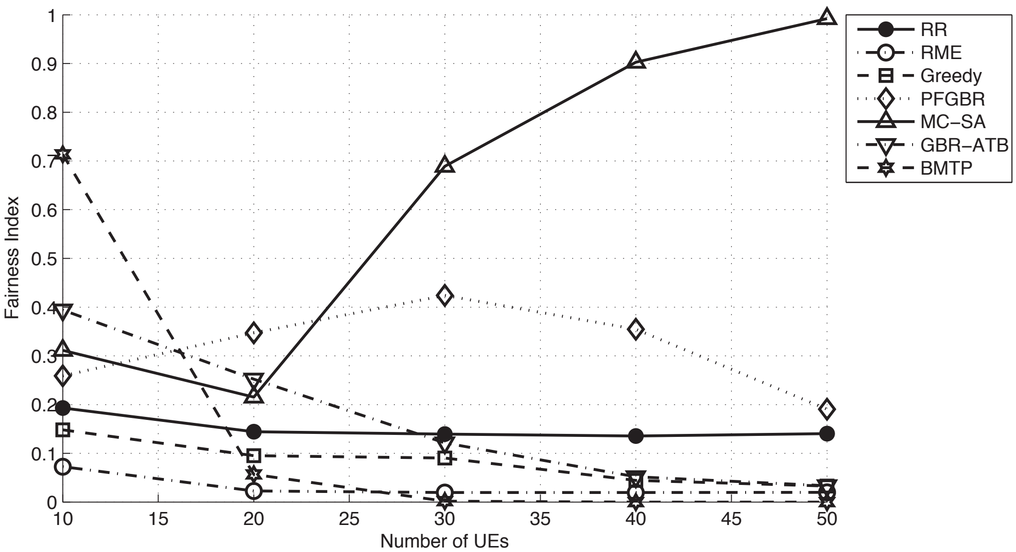


Fig. 5. Experiment 1:Aggregated Fairness

have chosen to simulate the Recursive Maximum Expansion (RME) [21] as an example of best effort schedulers; Proportional Fairness with Guaranteed Bit Rate (PFGBR) [27], Guaranteed Bit Rate with Adaptive Transmission Bandwidth (GBR-ATB) [28], and Multi-Carrier Scheduling Algorithm [30] as examples of schedulers optimizing QoS; and Block Allocation for Minimum Total Power (BMTP) [35] as an example of schedulers optimizing both QoS and power. A summary of these different approaches can be found in Table II. As baselines, we have also simulated two basic allocation approaches, namely Round Robin (RR) and greedy allocations (Greedy).

In what follows, we elaborate on the setup of a small set of experiments and showcase some of their results.

A. Experiment 1: Varying the Number of UEs Under Heavy Traffic

The experiment conducted here examines the effectiveness of PRB resource allocation of multiple UEs under heavy traffic load. Each UE generates a single 1 Mbps FTP stream, assuming no packet drops are taking place at the UE's transmission buffer. The use of a single traffic profile further assists us in determining the maximum attainable data throughput of the system under the packet scheduler's supervision. The results collected from this experiment function as a reference in evaluating the schedulers' performance with the presence of different traffic mixes. This applies to the experiments that are discussed in subsequent sections. Table VI lists some of the

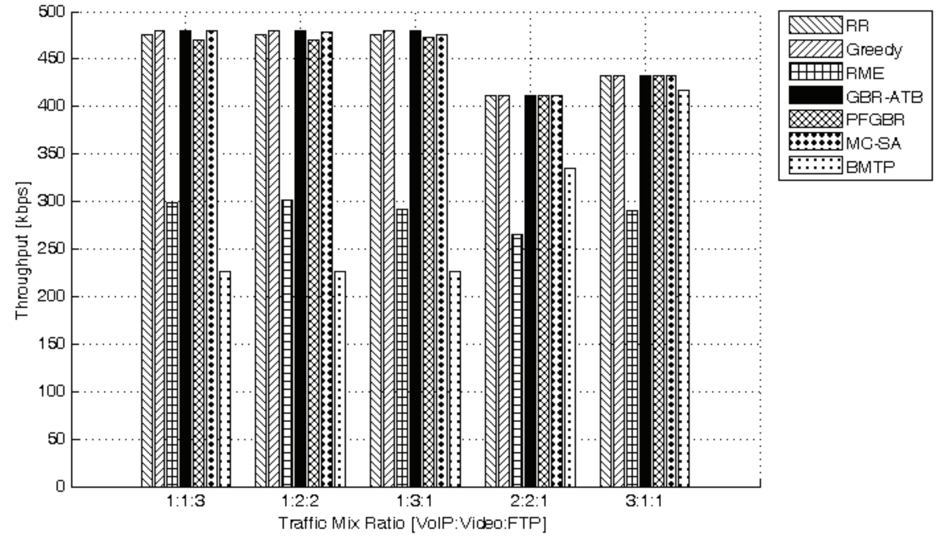


Fig. 6. Experiment 2: VoIP Aggregated Throughput

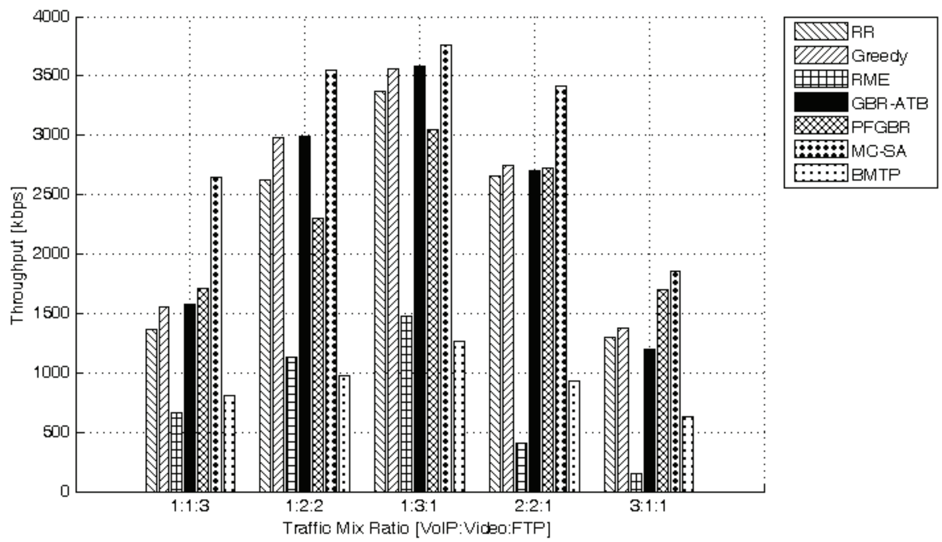


Fig. 7. Experiment 2: Video Aggregated Throughput

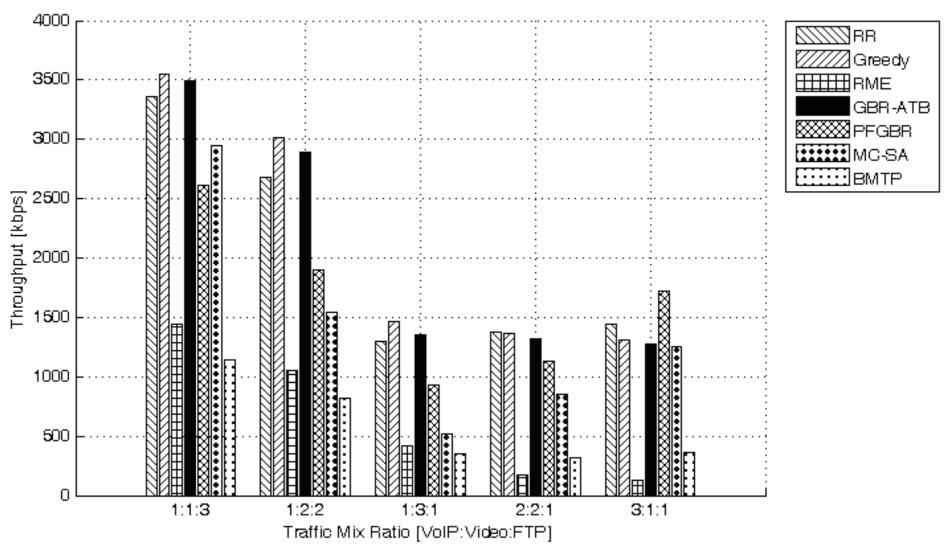


Fig. 8. Experiment 2: FTP Aggregated Throughput

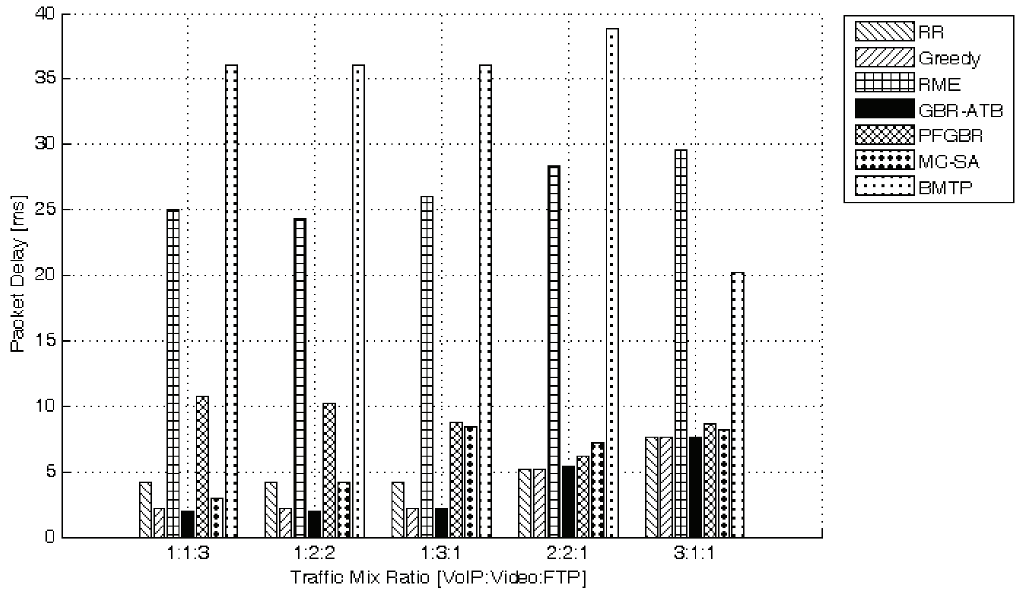


Fig. 9. Experiment 2: VoIP Average Delay

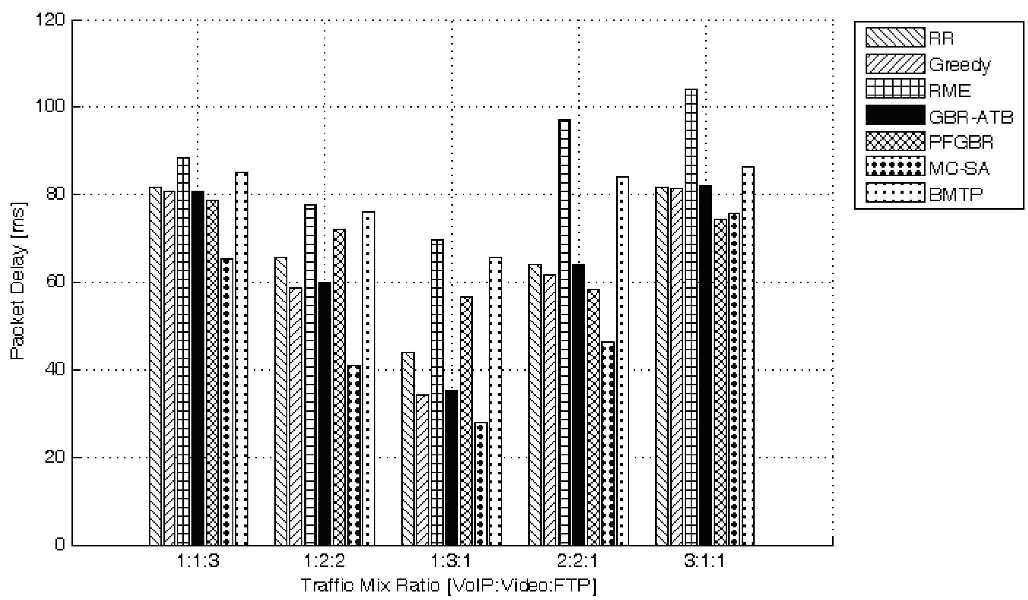


Fig. 10. Experiment 2: Video Average Delay

TABLE VI
SIMULATION PARAMETERS OF EXPERIMENT 1

Parameter	Description
Number of UEs	10, 20, 30, 40, 50
Traffic Profile	1 Mbps FTP stream per UE

experiment's configuration parameters such as the number of UEs and the traffic profile used.

Results obtained on the aggregate throughput and intraclass fairness are presented in Figures 4 and 5, respectively. In the figures, best-effort schedulers demonstrate higher throughput levels than that observed for QoS-optimizing and power and QoS-optimizing schedulers. Schedulers optimizing QoS, simulated under a single non-GBR FTP traffic load which neutralizes the QoS metrics, are effectively made PF-based schedulers. Certain QoS schedulers, such as PFGBR and MC-SA, do demonstrate better fairness levels than others. As for schedulers optimizing both QoS and power, the BMTP algorithm shows the lowest overall performance level. This low-performance can be attributed to the schedulers lowering the MCS level so that the threshold SINR to be met will be lower as well, which in turn lowers the uplink transmission power. This method of power optimization explains the relatively degraded throughput performance of the scheduler, as shown in Figure 4. Note that the power-saving approach chosen by

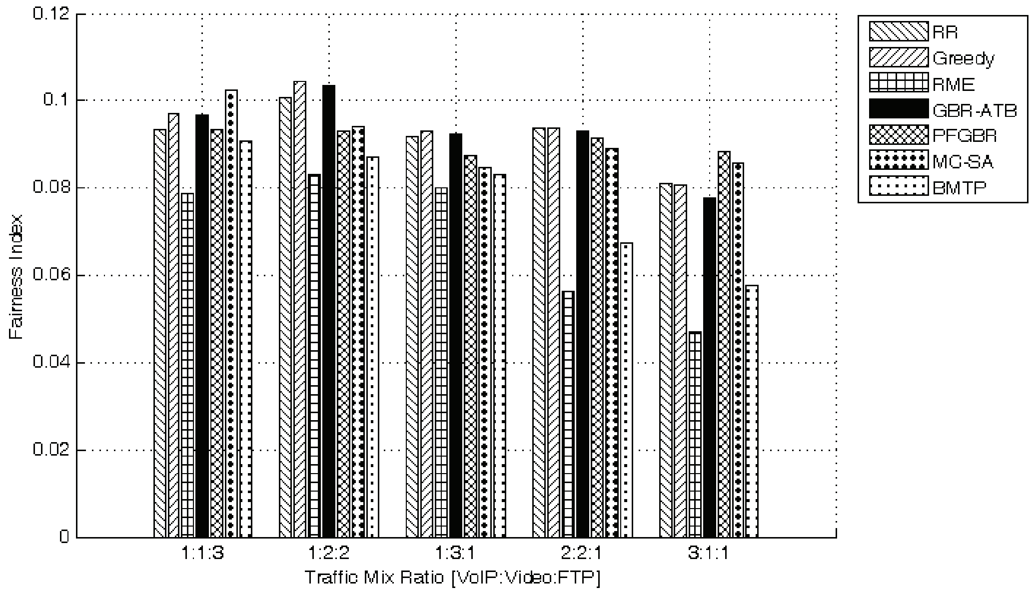


Fig. 11. Experiment 2: Inter-Class Fairness

TABLE VII
THE EFFECT OF TRAFFIC MIX RATIOS - PARAMETERS

Parameters	**Value**
Number of UEs	25
UE Ratios (VoIP : Video : FTP)	1:1:3, 1:2:2, 1:3:1, 2:2:1, 3:1:1

BMTP has the disadvantage of being very sensitive to UE's experienced channel condition, which is essentially a function of the UE's distance from the eNB. Such a disadvantage can have a detrimental effect on performance as the coverage area expands. Therefore, BMTP shares the same disadvantage as max-SINR schedulers introduced in earlier wireless systems where the scheduler favors UEs with better SINR, leaving UEs further away as victims for starvation.

B. Experiment 2: The Effect of Different Traffic Mixes on System Performance

The purpose of this experiment is to examine the schedulers' performance under different traffic mixes. The LTE uplink schedulers at the eNB do not distinguish between the different traffic types at a single UE when scheduling resources. Hence, we chose to provide each UE with one SDF carrying a single traffic stream in order to see the impact of the scheduler on the QoS experienced in traffic mix scenarios. The total traffic load for the entire experiment is set to 8145 kbps, of which 465 kbps is given to VoIP traffic, 3840 kbps to video streaming, and 3840 kbps to FTP traffic. The remaining parameters are listed in Table VII. The UE ratios presented in the table reflect the ratio of the number of UEs with a given traffic class to the number of UEs from the other traffic classes. Due to the limitation of the simulation time of our experiments to only 10 seconds, we fixate all VoIP streams at the active state to regulate the offered VoIP traffic load.

The results obtained for the cell's aggregate throughput, i.e., Figures 6 through 8, consistently show that the total system throughput of a given traffic class improves as the concentration of UEs belonging to that class increase within the system, given that the network traffic load is fixed. For UEs of a given traffic class, increasing the number of UEs decreases the contention of the traffic flows per UE, which decreases the per UE packet drops while increasing the probability of scheduling a UE of the same traffic class.

Compared to Experiment 1, best effort schedulers such as the RME suffer significant performance degradation in the presence of multiple traffic profiles. UEs with VoIP traffic have lower data rate requirements compared to UEs with video and FTP traffic. The PF metric used in the RME algorithm does not distinguish variations in data rates from one UE to another. Also, RME adopts a dynamic resource allocation method, which does not impose any restrictions on the number of PRBs that can be allocated to a single UE.

When looking at the average delay results of both VoIP and video traffic classes, i.e., Figures 9 and 10, Greedy's ability and that of other QoS-based schedulers to accommodate the QoS requirements of VoIP traffic fairs much better than others. Similar behaviors were found when the effects on packets drops was isolated.These observations comprise a strong indication that the modification from a classic PF utility function to a QoS-based utility function can significantly improve the performance of adaptive allocation schemes in accommodating traffic mixes.

Meanwhile, schedulers optimizing both QoS and power show poor performance with all three traffic classes simulated. This can be observed, for example, when observing the inter-class fairness shown in Figure 11. One main reason for such a poor performance level is the low efficiency of bandwidth usage for lowering the uplink transmission power, as mentioned in earlier discussions of Experiment 1. For example,

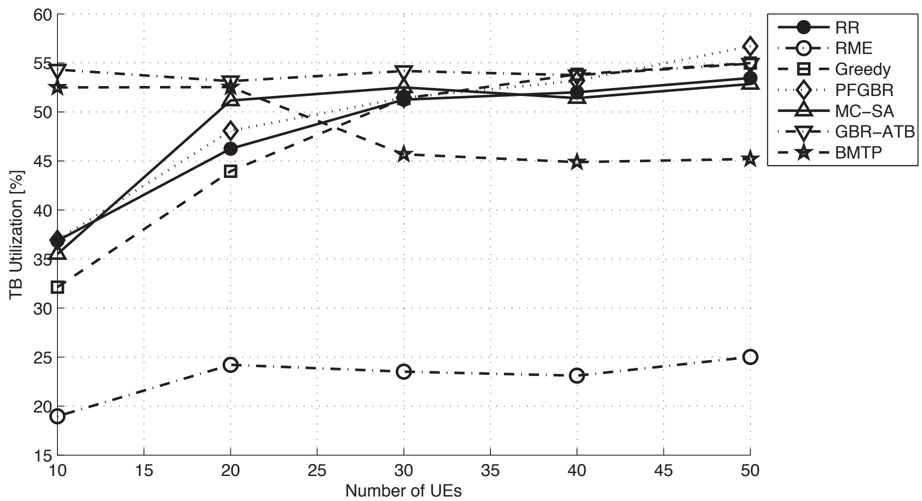


Fig. 12. Experiment 3-1: VoIP TB Utilization

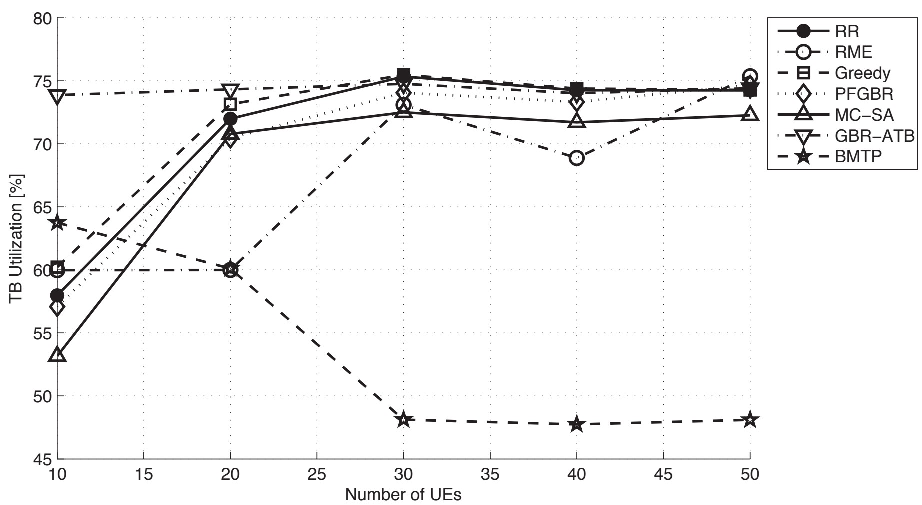


Fig. 13. Experiment 3-1: Video TB Utilization

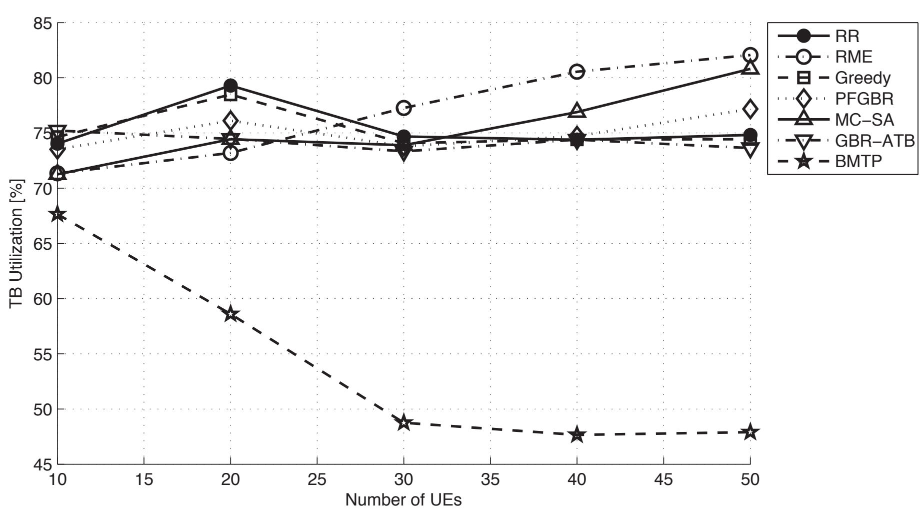


Fig. 14. Experiment 3-1: FTP TB Utilization

TABLE VIII
EFFECT OF NUMBER OF UES ON SYSTEM PERFORMANCE PARAMETERS.

Parameters	Value
Number of UEs	10, 20, 30, 40, 50
UEs Ratio (VoIP, Video, FTP)	2:2:1

allocating a VoIP UE three PRBs or more for transmitting a single VoIP packet causes UEs of other traffic classes to have higher starvation levels, eventually causing an overall lower satisfaction level for active upstream traffic.

C. Experiment 3: Effect of Varying Number of UEs Under Mixed Traffic Scenarios

The number of UEs within the cell and the amount of offered traffic load per UE affect the contention over the limited radio resources. Increasing the number of UEs leads to increasing the competition among UEs over the limited number of PRBs, which decreases the chances of transmission for each UE. Also, increasing the traffic load per UE leads to increasing the contention over the limited PRBs assigned to the UE. On the other hand, having a small number of UEs with a light load can lead to insufficient utilization of the assigned resources, resulting in a waste of transmission power and hence a waste of radio resources that could have been allocated elsewhere.

1) Varying UEs under Per-UE Fixed Uplink Traffic Load: The purpose of this experiment is to measure the utilization efficiency in terms of how well a UE of a certain QoS class can utilize its assigned TB. The simulation parameters for this experiment are shown in Table VIII. The traffic load per UE is fixed to show the effect of increase contention on the number of PRBs assigned per UE. The traffic rates of UEs with VoIP, video streaming, and FTP traffic streams are set to 14.4 kbps, 64 kbps, and 128 kbps, respectively. When constructing the TB for each transmission, a 6-byte overhead is assumed; 3 bytes are generated for the MAC and RLC headers, and 3 bytes are occupied by the CRC checksum.

A general trend observed in the TB utilization for all traffic classes, i.e., Figures 12 to 14, is that as long as there are low contentions on available radio resources, the utilization of assigned TB per UE also increase. Note that the TB utilization starts to saturate as the number of UEs in the system surpass the total number of available PRBs. Similar trends were observed when isolating the aggregate throughput. Specifically, it was observed that the aggregate throughput of MC-SA increases when the number of UEs changes from 20 to 30 UEs, which triggers the switch from scheduling method change executed when the number of UEs is smaller than the number of PRBs. UEs close to the cell edge and which experience relatively poor channel conditions will suffer from poor spectral efficiency due to using low MCS and hence smaller TB sizes. The effect of the overhead from the MAC header and the CRC checksum collectively reduces the amount of space available for the data payload further. On the other hand, UEs closer to the eNB with better channel conditions transmit at higher MCS, which reduces the TB overhead impact on the TB utilization. The variation of TB overhead just explained reduces the inter-class fairness among UEs of the same traffic class. An instance of this effect is shown in Figure 15 for the video traffic class.

TABLE IX
EFFECT OF NUMBER OF UES ON TB UTILIZATION WITH FIXED LOAD - PARAMETERS.

Parameters	Value
Number of UEs	10, 20, 30, 40, 50
UEs Ratio (VoIP, Video, FTP)	4:3:3

UEs with VoIP traffic show poor average TB utilization of assigned resources that never exceed 55%. This is due to VoIP UE's low traffic loads relative to video and FTP UEs. The low TB utilization is contributed mainly by UEs close to the eNB, where they get assigned MCS high enough that only a small portion of the TB is used for data payloads while the rest of the TB space is filled with padding. This is also observed with RR and Greedy schedulers where a UE is assigned only one PRB for 30 UEs and above. The results observed for VoIP traffic is a strong indication that supporting voice services over the SC-FDMA radio interface can be easily achieved since a VoIP packet can effortlessly fit into a single PRB if the UE is close enough to the eNB. Therefore, UEs close to eNB should experience very minimal delays and packet drops. On the contrary, UEs closer to the cell edge tend to transmit with a low MCS. As a result, TB sizes tend to get smaller to the point that the overhead size becomes more significant compared to the small space available for data payloads. This clearly indicates that packet drops within the system are mostly contributed from cell-edge UEs.

In other results, VoIP packet drops and packet delays do not seem to be affected by the number of VoIP UEs in the system. When increasing the number of VoIP UEs in the cell, the probability of placing UEs in close proximity of the eNB, that is midway between eNB and close to the cell edge, is the same. Therefore, as the total offered VoIP increases with the number of UEs, the percentage of VoIP packets that get successfully transmitted stays the same. More generally, packet delays and packet drops of both VoIP traffic and video traffic under the different schedulers show little variations as the number of UEs increases. Also, since the offered traffic load increases with the number of UEs, the overall delay and packet drops stay relatively the same for any number of UEs within the cell coverage.

2) Varying Number of UEs Under Fixed Total Uplink Traffic Load: In this experiment, we want to see the effect on system performance of varying the number of UEs while keeping the total traffic load fixed. Varying the number of UEs under fixed traffic illustrates how the total throughput of the system can vary with the traffic contention over the resources allocated per UE. The findings in this experiment are to be compared to the performance results obtained from Experiment 1 on the system settings that best suit the schedulers under study. The simulation parameters are listed in Table IX. The total network traffic load was fixed at 6400 kbps, where 256 kbps is reserved for VoIP, 3072 kbps for video streaming, and 3072 kbps for FTP traffic.

An interesting aspect to consider in this experiment is the performance of aggregate throughput from the different traffic

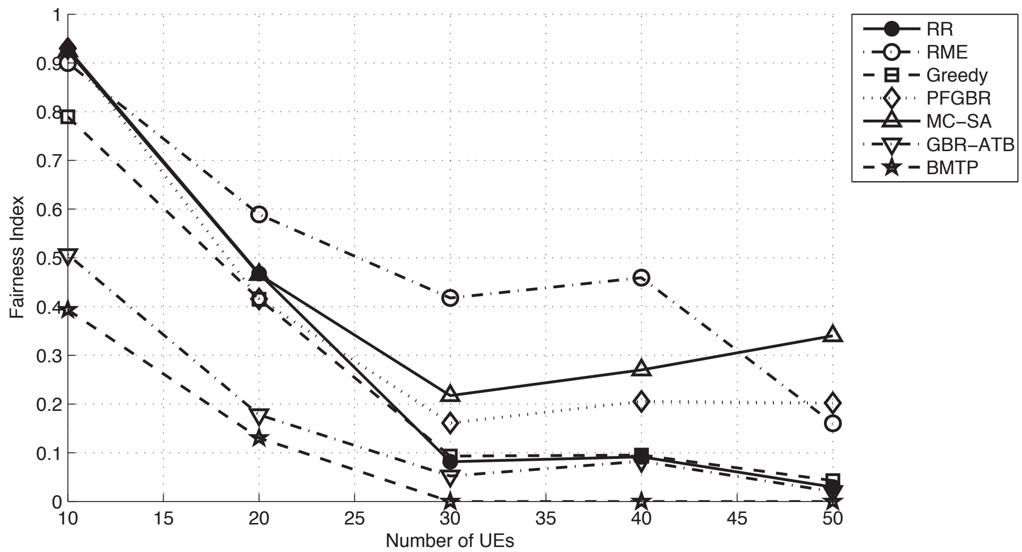


Fig. 15. Experiment 3-1: Video Min-Max Fairness

classes, shown in Figures 16 through 18. The system's total throughput from UEs of all three traffic classes constitutes about 30-50 % of the entire offered traffic load. Most of the loss in throughput comes from FTP traffic, followed by video streaming. UEs of the VoIP traffic class suffered the least performance degradation, as the experienced average packet loss did not exceed 20%, except in the case of PFGBR and BMTP.

VIII. Conclusion

Our intent in this work was to instigate a discussion on a more expanded treatment of scheduling in LTE and LTE-Advanced. A tutorial on uplink scheduling in both evolutions was offered, followed by a survey of proposals made in the literature. As things stand, there are several outstanding issues that need to be addressed. Schedulers with considerations for clustered (non-contiguous) SC-FDMA, carrier aggregation and coordinated multi-point transmission/reception, are receiving increased attention. Meanwhile, only few proposals addressed the growing device heterogeneity (UE/MTC) that is expected to happen in this decade. And while the focus of this work was on uplink scheduling, approaching scheduling in the downlink and uplink as a hybrid problem is certainly a possibility. A considerable portion of this work elaborated on how uplink scheduling proposals should be evaluated, and discussed an example setup that investigated the performance of some representative proposals. As LTE deployments continue to be rolled out, validating scheduler performance becomes essential for both vendors and network operators. 3GPP continues to capitalize on the choice of OFDM-based techniques for both its downlink and uplink access, and enhancements introduced promise to deliver substantial gains in network performance. However, as the number and type of devices increases, and as the sophistication of the physical layer enhancements rises, the design of both real- and non-real time scheduling becomes more challenging, especially as designers try to circumvent the increasing complexity of the functionality.

List of Acronyms

3GPP	Third Generation Partnership Project
AMBR	Aggregate Maximum Bit Rate
ARP	Allocation/Retention Priority
ARQ	Automatic Repeat Request
BSR	Buffer Status Reports
CC	Component Carrier
CLPC	Closed Loop Power Control
CoMP	Coordinated Multipoint Transmission/Reception
CP	Cycle Prefix
CQI	Channel Quality Indicator
CS/CB	Coordinated Scheduling and Beamforming, a CoMP variant
DFT-S-OFDM	Discrete Fourier Transform Spread OFDM, or Clustered SC-FDMA
DMRS	Demodulation Reference Signal
DRX	Discontinuous Reception
D-SR	Dedicated Scheduling Request
eNB	evolved NodeB
EPS	Evolved Packet System
FDD	Frequency Division Duplexing
GBR	Guaranteed Bit Rate
H2H	Human-to-Human
HARQ	Hybrid Automatic Repeat Request
ITU-R	International Telecommunications Union-Radiocommunication Sector
JR	Joint Reception, a CoMP variant
LTE	Long Term Evolution
LTE-Advanced	Long Term Evolution Advanced
M2M	Machine-to-Machine
MAC	Medium Access Control
MBR	Maximum Bit Rate
MCS	Modulation and Coding Scheme
MME	Mobility Management Entity
MTC	Machine Type Communication

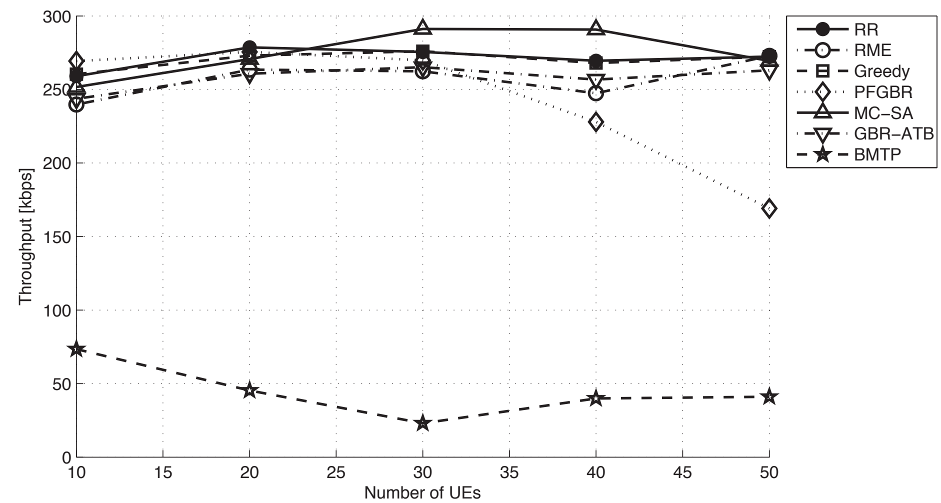


Fig. 16. Experiment 3-2: VoIP Aggregated Throughput

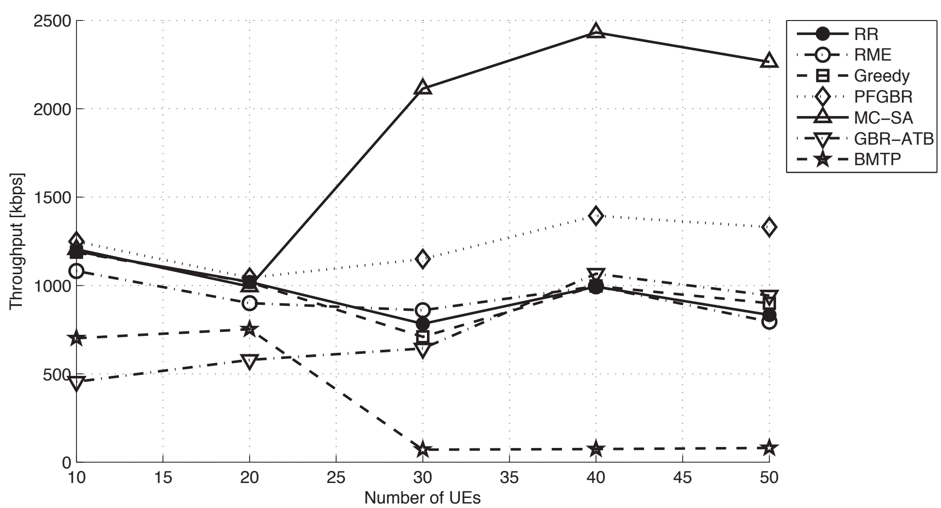


Fig. 17. Experiment 3-2: Video Aggregated Throughput

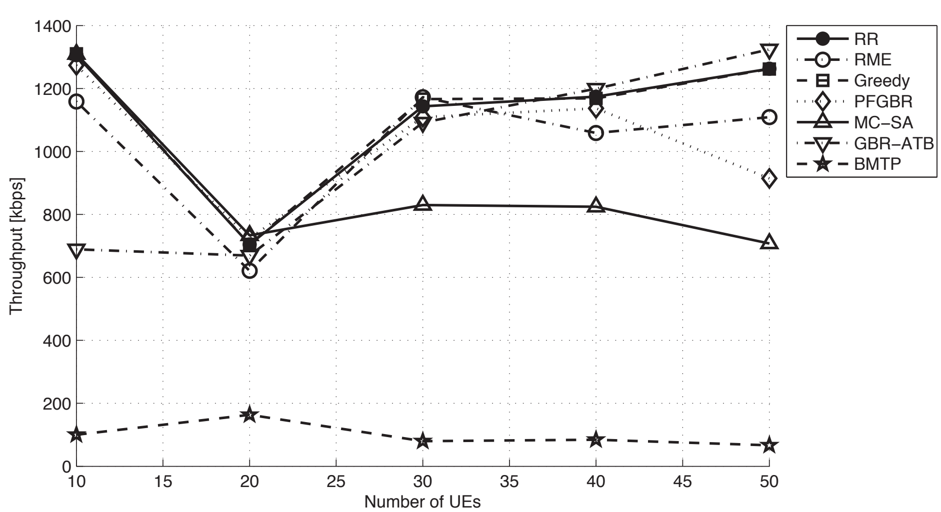


Fig. 18. Experiment 3-2: FTP Aggregated Throughput

NAS	Non-Access Stratum
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OLPC	Open Loop Power Control
PAPR	Peak to Average Power Ratio
PCell	Primary Cell in carrier aggregation
PDCCH	Physical Downlink Control Channel
PDCP	Packet Data Convergence Protocol
PF	Proportional Fairness
P-GW	Packet Data Network Gateway
PHR	Power Headroom Report
PHY	Physical Layer
PRB	Physical Resource Block
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Shared Channel
QCI	QoS Class Identifier
QoS	Quality of Service
RAC	Radio Access Control
RACH	Random Access Channel
RAN	Radio Access Network
RA-SR	RACH Scheduling Request
RLC	Radio Link Control
RRC	Radio Resource Control
RRE	Remote Radio Equipment
RTT	Round Trip Time
S1	Interface between eNBs and network core
SCell	Secondary Cell in carrier aggregation
SC-FDMA	Single Carrier Frequency Division Multiple Access
SDF	Service Data Flow
S-GW	Serving Gateway
SR	Scheduling Requests
SRS	Sounding Reference Signal
TB	Transport Block
TDD	Time Division Duplexing
TFT	Traffic Flow Template
TTI	Transmission Time Interval
UE	User Equipment
U-SCH	Uplink Shared Channel
X2	Interface between eNBs

ACKNOWLEDGMENTS

This work was made possible by a National Priorities Research Program (NPRP) grant from the Qatar National Research Fund (Member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

The authors would like to thank Hatem Abou-Zeid and Mohammad Tauhidul Islam for their feedback on later drafts of this paper. The authors would also like to thank the editor and the reviewers for their valuable input on the manuscript.

REFERENCES

[1] R. Kwan and C. Leung, "A survey of scheduling and interference mitigation in LTE," *JECE*, vol. 2010, pp. 1:1–1:10, Jan. 2010.

[2] E. Yaacoub and Z. Dawy, "A survey on uplink resource allocation in OFDMA wireless networks," *IEEE Commun. Surveys Tutorials*, vol. 14, pp. 322 –337, quarter 2012.

[3] D. Dimitrova, J. van den Berg, G. Heijenk, and R. Litjens, "LTE uplink scheduling: Flow level analysis," in *Multiple Access Communications* (C. Sacchi, B. Bellalta, A. Vinel, C. Schlegel, F. Granelli, and Y. Zhang, eds.), vol. 6886 of *Lecture Notes in Computer Science*, pp. 181–192, Springer Berlin / Heidelberg, 2011.

[4] E. Yaacoub, H. Al-Asadi, and Z. Dawy, "Low complexity scheduling algorithms for the LTE uplink," in *IEEE Symp. Comput. Commun.*, pp. 266 –270, July 2009.

[5] N. Ali, A.-E. Taha, and H. Hassanein, *LTE, LTE-Advanced and WiMAX: Towards IMT-Advanced Networks*. ITPro collection, Wiley, 2011.

[6] http://www.3gpp.org/ftp/Specs/html-info/36300.htm.

[7] 3GPP, "Coordinated multi-point operation for LTE physical layer aspects (Release 11)," TR 36.819, 3rd Generation Partnership Project (3GPP), Dec. 2011.

[8] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, "Coordinated multipoint transmission/reception techniques for LTE-Advanced," *IEEE Wireless Commun.*, vol. 17, pp. 26 –34, June 2010.

[9] "LTE, Release 12 and beyond," 2013. Nokia white paper, available from http://goo.gl/fbynK; accessed 20 February 2013.

[10] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 10)," TR 36.814, 3rd Generation Partnership Project (3GPP), Mar. 2010.

[11] P. Marsch and G. Fettweis, *Coordinated Multi-Point in Mobile Communications: From Theory to Practice*. Coordinated Multi-point in Mobile Communications: From Theory to Practice, Cambridge University Press, 2011.

[12] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (Release 11)," TS 36.213, 3rd Generation Partnership Project (3GPP), Dec. 2012.

[13] S. Ye, S. Wong, and C. Worrall, "Enhanced physical downlink control channel in LTE-Advanced Release 11," *IEEE Commun. Mag.*, vol. 51, pp. 82 –89, Feb. 2013.

[14] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 10)," TS 36.300, 3rd Generation Partnership Project (3GPP), Jan. 2013.

[15] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 11)," TS 36.300, 3rd Generation Partnership Project (3GPP).

[16] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Carrier Aggregation Enhancements; UE and BS radio transmission and reception (Release 11)," TR 36.823, 3rd Generation Partnership Project (3GPP), Sept. 2012.

[17] "3GPP LTE standards update: Releases 11, 12, and beyond," 2012. Agilent white paper, available from http://goo.gl/BWR0W; accessed 20 February 2013.

[18] "LTE-Advanced, the advanced LTE toolbox for more efficient delivery of better user experience," 2011. Nokia white paper, available from http://goo.gl/gbhuf; accessed 20 February 2013.

[19] 3GPP, "Study on provision of low-cost Machine-Type Communications (MTC) User Equipments (UEs) based on LTE (Release 11)," TR 36.888, 3rd Generation Partnership Project (3GPP), June 2012.

[20] F. Capozzi, G. Piro, L. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in lte cellular networks: Key design issues and a survey," *IEEE Commun. Surveys Tutorials*, vol. 15, no. 2, pp. 678–700, 2013.

[21] L. Ruiz de Temino, G. Berardinelli, S. Frattasi, and P. Mogensen, "Channel-Aware Scheduling Algorithms for SC-FDMA in LTE Uplink," in *IEEE 19th Int. Symp. Personal, Indoor and Mobile Radio Commun.*, pp. 1 –6, 15-18 2008.

[22] F. Calabrese, P. Michaelsen, C. Rosa, M. Anas, C. Castellanos, D. Villa, K. Pedersen, and P. Mogensen, "Search-Tree Based Uplink Channel Aware Packet Scheduling for UTRAN LTE," in *IEEE Veh. Technol. Conf.*, pp. 1949 –1953, 11-14 2008.

[23] H. Yang, F. Ren, C. Lin, and J. Zhang, "Frequency-domain packet scheduling for 3GPP LTE uplink," in *INFOCOM*, pp. 1 –9, Mar. 2010.

[24] F. Calabrese, C. Rosa, M. Anas, P. Michaelsen, K. Pedersen, and P. Mogensen, "Adaptive Transmission Bandwidth Based Packet Scheduling for LTE Uplink," in *IEEE 68th Veh. Technol. Conf.*, pp. 1 –5, 21-24 2008.

[25] E. Yaacoub and Z. Dawy, "A game theoretical formulation for proportional fairness in LTE uplink scheduling," in *IEEE Wireless Commun. Netw. Conf.*, pp. 1 –5, Apr. 2009.

[26] E. Yaacoub and Z. Dawy, "Achieving the Nash bargaining solution in OFDMA uplink using distributed scheduling with limited feedback," *Int. J. Electronics and Commun.*, vol. 65, no. 4, pp. 320 – 330, 2011.

[27] P. Wen, M. You, S. Wu, and J. Liu, "Scheduling for streaming application over wideband cellular network in mixed service scenarios," in *IEEE Military Commun. Conf.*, pp. 1 –7, Oct. 2007.

[28] M. Anas, C. Rosa, F. Calabrese, K. Pedersen, and P. Mogensen, "Combined Admission Control and Scheduling for QoS Differentiation in LTE Uplink," in *IEEE 68th Veh. Technol. Conf., 2008*, pp. 1 –5, Sept. 2008.

[29] R. Ruby and V. Leung, "Towards QoS assurance with revenue maximization of LTE uplink scheduling," in *2011 9th Annu. Commun. Netw. and Services Research Conf. (CNSR)*, pp. 202 –209, May 2011.

[30] O. Delgado and B. Jaumard, "Scheduling and Resource Allocation for Multiclass Services in LTE Uplink Systems," in *IEEE 6th Int. Conf. Wireless and Mobile Computing, Netw. and Commun. (WiMob)*, pp. 355 –360, Oct. 2010.

[31] Y. Zaki, T. Weerawardane, C. Gorg, and A. Timm-Giel, "Multi-QoS-aware fair scheduling for LTE," in *IEEE 73rd Veh. Technol. Conf. (VTC Spring)*, pp. 1 –5, May 2011.

[32] M. Al-Rawi, R. Jantti, J. Torsner, and M. Sagfors, "Opportunistic Uplilnk Scheduling for 3G LTE Systems," in *The 4th Int. Conf. on Innovations in Information Technology*, pp. 705 –709, 18-20 2007.

[33] O. Delgado and B. Jaumard, "Scheduling and Resource Allocation in LTE Uplink with a Delay Requirement," in *8th Annu. Commun. Networks and Services Research Conf. (CNSR)*, pp. 268 –275, May 2010.

[34] Z. Li, C. Yin, and G. Yue, "Delay-Bounded Power-Efficient Packet Scheduling for Uplink Systems of LTE," in *The 5th Int. Conf. on Wireless Communications, Networking and Mobile Computing*, pp. 1 –4, Sept. 2009.

[35] F. Sokmen and T. Girici, "Uplink Resource Allocation Algorithms for Single-Carrier FDMA Systems," in *Wireless Conf. (EW), 2010 European*, pp. 339 –345, Apr. 2010.

[36] M. Lauridsen, A. Jensen, and P. Mogensen, "Reducing LTE Uplink Transmission Energy by Allocating Resources," in *2011 IEEE Veh. Technol. Conf. (VTC Fall)*, pp. 1 –5, Sept. 2011.

[37] D. Jiang, Q. Wang, J. Liu, G. Liu, and C. Cui, "Uplink coordinated multi-point reception for LTE-Advanced systems," in *5th Int. Conf. on Wireless Communications, Networking and Mobile Computing*, pp. 1 –4, Sept. 2009.

[38] S. Brueck, L. Zhao, J. Giese, and M. Amin, "Centralized scheduling for joint transmission coordinated multi-point in LTE-Advanced," in *Int. ITG Workshop on Smart Antennas (WSA)*, pp. 177 –184, Feb. 2010.

[39] S. Wesemann, W. Rave, and G. Fettweis, "Combinatorial advantages due to joint resource allocation in uplink CoMP systems," in *The 8th Int. Symp. Wireless Commun. Systems (ISWCS)*, pp. 740 –744, Nov. 2011.

[40] H. Wang, C. Rosa, and K. Pedersen, "Uplink Component Carrier Selection for LTE-Advanced Systems with Carrier Aggregation," in *IEEE Int. Conf. Commun. (ICC)*, pp. 1 –5, June 2011.

[41] A. Lioumpas and A. Alexiou, "Uplink scheduling for Machine-to-Machine communications in LTE-based cellular systems," in *IEEE GLOBECOM Workshops*, pp. 353 –357, Dec. 2011.

[42] S.-Y. Lien, K.-C. Chen, and Y. Lin, "Toward ubiquitous massive accesses in 3GPP machine-to-machine communications," *IEEE Commun. Mag.*, vol. 49, pp. 66 –74, Apr. 2011.

[43] S.-Y. Lien and K.-C. Chen, "Massive access management for QoS guarantees in 3GPP machine-to-machine communications," *IEEE Commun. Lett.*, vol. 15, pp. 311 –313, Mar. 2011.

[44] A. Gotsis, A. Lioumpas, and A. Alexiou, "M2M Scheduling over LTE: Challenges and New Perspectives," *IEEE Veh. Technol. Mag.*, vol. 7, pp. 34 –39, Sept. 2012.

[45] S. Y. Shin and D. Triwicaksono, "Radio resource control scheme for machine-to-machine communication in LTE infrastructure," in *Int. Conf. on ICT Convergence (ICTC)*, pp. 1 –6, Oct. 2012.

[46] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation," TR 36.211, 3rd Generation Partnership Project (3GPP), Mar. 2010.

[47] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) System Scenarios," TR 36.942, 3rd Generation Partnership Project (3GPP), Apr. 2010.

[48] 3GPP, "Radio transmission and reception," TR 45.005, 3rd Generation Partnership Project (3GPP), June 2010.

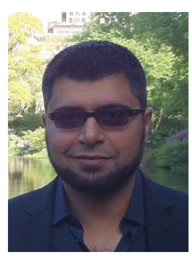
[49] F. Calabrese, *Adaptive Radio Resource Management for Uplink Wireless Networks*. Phd Thesis, Polytechnic University, June 2006.

[50] Orange, China Mobile, KPN, NTT DoCoMo, Sprint, T-Mobile, Vodafone, Telecom Italia, "LTE Physical Layer Framework for Performance Verification," TS R1-070674, 3rd Generation Partnership Project (3GPP), Feb. 2007.

[51] 3GPP, "Policy and Charging Control Architecture," TR 23.203, 3rd Generation Partnership Project (3GPP), Mar. 2010.

Najah Abu Ail is currently an associate professor in the Faculty of Information Technology at the University of the United Arab Emirates. She got her PhD in computer networks from the Department of Electrical Engineering from Queen's University in Kingston, Canada. Her research interests include analytical and measurement based performance evaluation of wired and wireless computer networks, radio resource management in broadband wireless networks and design analytical and reliable schemes to provide for quality of service in wireless sensors networks. She published her work in several prestigious journals and conferences and delivered several tutorials in her fields of research at high-ranked conferences. Dr. Abu Ali is a PI/co-PI of several major research projects that have been funded by several agencies such as QNRF and NRF / United Arab Emirates University. She also co-authored a book on the architecture and operation of 4G and beyond technologies.

Abd-Elhamid M. Taha is currently an assistant Professor in Electrical Engineering at Alfaisal University, his B.Sc. (Honours) and M.Sc. in Electrical Engineering were earned at Kuwait University in 1999 and 2002, and his Ph.D. in Electrical and Computer Engineering was earned Queens University, Canada in 2007. Dr. Taha's general research interest is in the area of computer networks and communications. His particular focus, however, has been on radio resource management in wireless networks. Recent themes in this direction include the design of resource schedulers with reduced complexity, and enabling machine-to-machine communications. Other currently active areas of interest include simplified localization in massive wireless sensing networks, mobile security in the Internet of Things (IoT), and modeling in networked cyber-physical systems. He has written and lectured extensively on broadband wireless networks, focusing on radio resource management techniques. He is also the co-author of the book LTE, LTE-Advanced and WiMAX: Toward IMT-Advanced Networks by Wiley & Sons and a presenter for several tutorials at key IEEE Communications Society events. His service record includes organizing and service on the editorial and technical program committees of many esteemed publication events and venues, as well as advising and reviewing activities for funding agencies, technical book publishers and research journals and conferences. Dr. Taha's currently a Senior Member of the IEEE and the IEEE Communications Society, as well as a member of the ACM.



Mohamed Salah received his Masters of Applied Science in Electrical Engineering from Queens University in 2011. From 2008 to 2011, he worked on LTE E-UTRAN towards his Masters thesis, with the focus on packet scheduling in LTE MAC layer at the eNodeB. He afterwards joined Alcatel-Lucent as a software Development Engineer at the research and development site in Canada. Ever since he has been working as part of the 5620 SAM design team to support SNMP and Netconf-based network management features for Alcatel-Lucents MME product.

Hossam Hassanein is a leading authority in the areas of broadband, wireless and mobile networks architecture, protocols, control and performance evaluation. His record spans more than 500 publications in journals, conferences and book chapters, in addition to numerous keynotes and plenary talks in flagship venues. Dr. Hassanein has received several recognition and best papers awards at top international conferences. He is the founder and director of the Telecommunications Research (TR) Lab at Queen's University School of Computing, with extensive international academic and industrial collaborations. Dr. Hassanein is a senior member of the IEEE, and is the past chair of the IEEE Communication Society Technical Committee on Ad hoc and Sensor Networks (TC AHSN). He is an IEEE Communications Society Distinguished Speaker (Distinguished Lecturer 2008-2010).