Dynamic UL LTE Scheduler for Supporting Smart Grid Applications

Ahmed Hassebo[1]* and Sumeet Chaudhary[1]

[1]*Department of electrical and computer engineering, Purdue University Northwest, Westville, IN, USA*

Abstract: Cellular dependent Machine-to-Machine (M2M) infrastructure is one of the key internet of things (IoT) empowering advances with immense market potential for cell specialist organizations sending 4G long term evolution (LTE) systems. The motivation for this paper is to investigate whether current commercial 4G LTE systems can possibly bolster a portion of the evolving strategic IoT applications. To accomplish this goal, we propose and devise a basic hybrid LTE uplink (UL) scheduling algorithm that uses a common LTE's dynamic scheduling for supporting human to human (H2H) applications just as M2M applications and Semi-Persistent Scheduling (SPS) for strategic IoT services that consistently require steady radio resources sharing regularly. The simulation results show that present public 4G LTE networks can possibly sufficiently bolster a portion of the developing strategic smart grid applications including Phasor Measurement Units (PMUs), with a constraint latency as low as 20 ms and extraordinary reliability.

Keywords: M2M communications, H2H communications, IoT, 4G LTE, SPS, and PMU.

1. INTRODUCTION

The recent breakthroughs in the IoT applications, involving e-health, smart transportation systems (STS), smart grid, and smart houses, have made this technology an integral part of our diurnal routine. Soon, IoT will facilitate the remote interconnection and management of billions of smart devices, actuators, and sensors via the Internet. Cellular M2M communications are considered one of the essential IoT driving solutions with a promising massive market for cellular service provider implementing the LTE networks. The universal mobile connection will be delivered to the anticipated tens of billions of Internet-based devices through the fourth generation (4G) and the emerging fifth generation (5G) cellular technologies.

The cost effectiveness and easy availability of the commercial LTE cellular networks are being exploited by numerous critical industries to deliver connectivity to their users, sensors, and smart M2M devices [1]. Most of the emerging IoT applications have strict constraints in terms of reliability and restrictive end-to-end (E2E) delay. The delay bound specified for each application refers to the device-to-device latencies caused due to the resultant delay from the application-level processing time and communication latency [2]. Latency, reliability, and availability are required by IoT application has distinctive performance constraints. Typically, the most dominant traffic in M2M applications is in the uplink (UL) direction (much higher than that in the downlink (DL) direction).

Thus, efficient LTE UL scheduling algorithms at the base station ("Evolved NodeB (eNB)" per 3GPP standards) are more critical for M2M applications. Originally, LTE was not proposed for IoT applications, as the traffic produced by M2M devices (running IoT applications) have different characteristics than those generated by the conventional H2H-based (e.g., voice/video and data) communications [3]. Additionally, due to the massive deployment of M2M devices and limited available radio resources, the effective radio resources management (RRM) and the UL scheduling establishes a severe task in the implementation of LTE for M2M communications.

Existing LTE quality of service (QoS) standard and UL scheduling algorithms cannot conform to such a comprehensive deviating performance constraints of these IoT M2M applications [4]. Though 4G LTE networks can support very low Packet Loss Ratio (PLR) at the physical layer, such reliability, comes without regard for increased latency from tens to hundreds of ms due to the aggressive use of retransmission mechanisms. Current 4G LTE technologies may satisfy a single performance metric of these mission critical applications, but not the

E-mail address: ahassebo@pnw.edu, schaudhaery@pnw.edu

simultaneous support of extreme-high reliability and low latency as well as high data rates.

Several QoS-aware UL LTE scheduling algorithms for supporting M2M services, as well as H2H services, as reported in [4-17]. However, urgent IoT applications cannot be supported by the majority of these algorithms, as they are not latency-aware [1]. Additionally, these algorithms are basic and do not entirely agree with LTE's signaling and QoS standards. For instance, a common practice is an assumption that the time domain UL scheduler located at the eNB prioritizes user equipment (UEs)/M2M devices connection requests based on the head-of-line (HOL) packet waiting time at the UE/device transmission buffer. However, as will be detailed below, LTE standard cannot support a mechanism that could allow the UEs/devices to communicate with the eNB uplink scheduler about the waiting time of uplink packets residing in their transmission buffers.

A wide range of urgent applications, for example, driverless vehicles, smart grid management, real-time operation, remote surgery, etc. have been enabled by the recent development of the Ultra-Reliable Low-Latency Communication (URLLC) concept [5]. URLLC and 5G new radio (NR) technologies possibly will grow into a prospective business fact.

Thus, deploying feasible time sensitive IoT applications will have to be postponed until URLLC and 5G NR solutions turn into achievable business. Because IoT applications, specifically mission critical, will have a significant impact on the welfare of all humanity, the immediate or near-term deployments of these applications is of utmost importance. It is the purpose of this paper to explore whether current commercial 4G LTE cellular networks potentially provide for some of the emerging time sensitive IoT services. In this paper, smart grid is nominated as an explanatory IoT example as one of the most challenging IoT services. Smart grids are essential to support delay sensitive services that have rigorous constraints in terms of E2E latency and reliability (e.g., real-time system protection). At the same time, support a substantial number of connected M2M devices with relaxed latency and reliability constraints (e.g., smart meters (SM)), as depicted in figure 1.
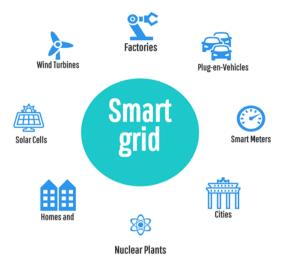

Figure 1. Smart grid applications.

PMUs are devices deployed throughout the power grid (mainly within Substations) that provide synchronized measurements for the peak value and the phase of sinusoidal voltage and current waveform. These synchronized measurements provide accurate system state measurements in real-time and are expected to be massively deployed for real-time wide-area monitoring and control (WAMC) of the next-generation power grid. The information generated within the WAMC is used for mission-critical smart grid applications including state estimation, control, and protection of the power grid.

To accomplish our objective, we propose and devise a simple hybrid LTE UL scheduling algorithm that employs a conventional LTE's dynamic scheduling for supporting H2H services as well as M2M applications and SPS for mission critical IoT applications. These applications have a necessary constraint for constant radio resource allotment regularly (like supporting voice in LTE networks). Precisely, we introduce a comprehensive LTE UL performance analysis that entirely corresponds to 4G LTE signaling and QoS standards to realistically evaluate the viability of commercial 4G LTE networks to provision such a various set of developing smart grid applications as well as conventional H2H applications.

The simulation results have shown that existing commercial 4G LTE systems can boost a few of the contemporary delay sensitive smart grid applications, comprising PMUs, with Packet Delay Budget (PDB) constraint as low as 20 ms. If achieving $PLR < 10^{-6}$ is the criteria for extreme-high reliability, then commercial 4G LTE systems cannot instantaneously support both extreme-high reliability and low latency.
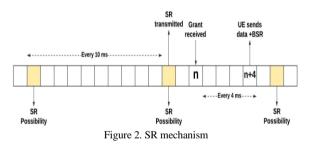
2. OVERVIEW OF LTE SIGNALING MECHANISM AND QOS MODEL

A. Signaling Mechanism

LTE standards identified two MAC layer signaling messages, Buffer Status Report (BSR), and Scheduling Request (SR), to demand resources from the eNB. At the initiation of the scheduling process, UE/device transmits SR (during its SR occasion) on the Physical Uplink Control Channel (PUCCH) to inform the eNB that UE/device has data to transmit. Each UE is assigned a specific offset within an SR period. Subsequently, the UE must expect for its specific offset subframe (i.e., TTI) to transmit its SR [19-21]. The offset has already assigned by the eNB during the RRC connection setup [20].

LTE standard (Release 8) identifies five distinct SR periods of 5, 10, 20, 40, or 80 ms [20]. Consequently, shorter periods of 1 ms and 2 ms were introduced in Release 9. In this paper, an SR period of 10 ms was assumed so that SR offsets are within the 0-9 ms range. Note that the SR does not include any information about the UE/device buffer status. As a result, the scheduler at the eNB does not have a detailed knowledge of buffer content. Therefore, the eNB must allocate the initial resources (uplink grant) blindly. In this work, it is assumed that, for the initial UL grant, the scheduler allocates a static size of bytes for each UE/device. Successively, these are translated to 1, 2, 3 RBs in the frequency domain, varying based on the status of UL channel.

Figure 2. SR mechanism

On the other hand, the BSR enables the UE/device to communicate with the eNB and pass information about the volume of data pending in the transmission buffer along with their priority. As the UE/device may have relatively few radio bearers QoS class identifier (QCIs) in its buffer, a considerable signaling overhead might be required to keep the eNB informed of the status of such massive radio bearers (logical channels). LTE standard instituted the notion of a Logical Channel Group (LCG) to decrease the signaling overhead. This methodology allocates a group of logical channels (with similar QoS constraints) to one out of four groups, with distinct priorities. The radio bearers' allocation to an LCG is launched during RRC configuration. An LCG is a group of logical channels identified by a distinctive 2-bit LCG ID. Therefore, the eNB has been informed by UE/device

about the volume of the data pending transmission per-LCG. The eNB replies with per-LCG grant to UE/device.

RRC constructs two BSR Timers: Periodic BSR-Timer and retransmit-BSR-Timer (RETX_BSR_TIMER). There are three distinct kinds of BSR created by the UE/device including: Regular BSR, Periodic BSR, and Padding BSR.

a) Regular BSR is generated when the UL transmission queue was empty and new packets arrive in the next TTI. Additionally, as a new high priority data enters UL transmission queue given that the data waiting in the buffer has a lower priority. Finally, when the UE/device transmits a BSR but never receives a grant and the RETX_ BSR_TIMER expires. When the BSR is transmitted, the timer is started, and the timer is terminated when receiving a grant. This time spans from 320 ms up to 10.24 seconds. [6].

b) Periodic BSR is initiated every n TTIs. Every UE/device holds a periodic BSR timer. A new BSR is triggered when the timer terminates. The time ranges from 5 ms up to 2.56 seconds and designed by RRC.

c) Padding BSR, in each cycle, if the eNB allocate resources to the UE/device beyond the aggregate data volume in its transmission buffer, the unexploited space is termed as "padding". If this padding space is sufficient to host a BSR then the UE can transmit a padding BSR.
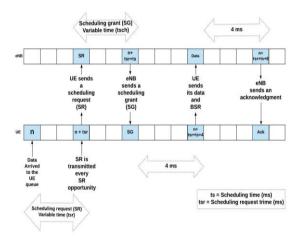

Figure 3. BSR mechanism

Note that if either a Regular or Periodic BSR is initiated, it will be transmitted at the first TTI. As the BSR is advantageous than the UE data, the eNB allocates the UL grant along with the BSR to the UE. Based on the data structure of the BSR, Figures 4-a and 4-b respectively show two forms of the BSR.

1. Short BSR: is a 1-byte Mac Control Element (CE) where the amount of data in UL transmission buffer can be reported by the UE/device for only one specific LCG.

2. Long BSR: is 3-bytes MAC CE where the amount of data in UL transmission buffer can be reported by the UE/device for all four LCGs along with their priority.
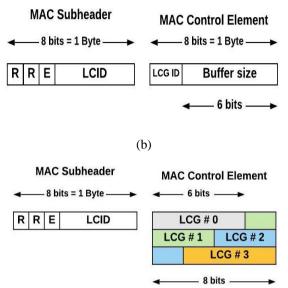
(a)

(b)

Figure 4. BSR format. (a) Short BSR and (b) Long BSR

B. QoS Model

The QoS model depends on the logical notion of an "EPS bearer," where "bearer" indicates a reasonable IP transmission route between the UE and the mobile core network (CN) with explicit QoS parameters (delay, PLR, etc.). A unique QCI is assigned to each bearer by the network and is composed of a radio bearer and a mobility tunnel. Bearers are comprised of two categories: guaranteed bit- rate (GBR) and non- guaranteed bit- rate (non-GBR) bearers. A GBR bearer is distinguished by a guaranteed bit rate (GBR) and maximum bit rate (MBR). In case of numerous non-GBR bearers in the same UE, UE splits an Aggregate Maximum Bit Rate (AMBR). As shown in Table I, the 3GPP specifications identify nine standard QCIs, each QCI is characterized by bearer type (GBR versus non-GBR), priority, packet delay, and PLR.

TABLE I.   QCI TABLE

QCI	Bearer Type	Priority	PDB (ms)	PLR	Example
1	GBR	2	100	10^{-2}	VoIP call
2		4	150	10^{-3}	Video call
3		3	50	10^{-6}	RT Gaming
4		5	300	10^{-3}	Vid stream
5	Non-GBR	1	100	10^{-6}	IMS Signal
6		6	300	10^{-3}	Video
7		7	100	10^{-6}	Gaming
8		8	300	10^{-6}	Buffer Streaming
9		9	300	10^{-6}	

3. THE SYSTEM MODEL

In this paper, we considered A 20 MHz LTE type-I system. As shown in Figure 5, a 2.5 km radius single cell base station is communicating with numerous fixed smart devices (experience a time invariant channel) and mobile UEs concurrently. These devices/UEs are placed at random locations within the cell coverage area. In the simulation environment, M2M devices are conceptual devices, that signify any measurement, controlling, and regulation function(s) for any IoT application including strategic services. To implement the simulation environment, we assumed that within the cell every UE/device has its own channel conditions, and the eNB is completely aware of the status of UEs' channels.
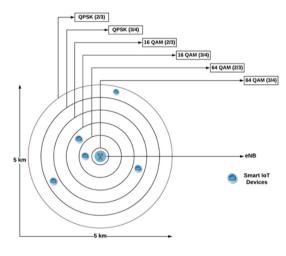

Figure 5. The modeled system

As shown in Table II, four different smart grid applications have been used to model M2M applications. APP1, APP2, and APP3, are delay sensitive services with strict PDB and extraordinary reliability constraints. On the other hand, APP4 models an application with comfortable latency and reliability constraints, e.g., SMs. Time sensitive APP1 models PMUs, which cause static length messages at regular intervals (constant bit rate) to be generated. Note that the PDB value specified in table II is the time taken for the packet, which has entered the device transmit buffer to be transmitted to the eNB.

TABLE II.   M2M TRAFFIC CHARACTERISTICS

M2M Apps	Packet Size (Byte)	Inter arrival time (ms)	# devices	PDB (ms)	Uplink Load (Mbps)
APP 1	100 Fixed	20 Fixed	600	20	24
App 2	A mean 100 Exp.	20 Exp.	600	40	24
App 3	A mean 100 Exp.	40 Exp.	600	60	12
App 4	125 Fixed	(100 - 500) Unif.	600	500	1.2 - 6

PMUs continuously generate 100 bytes static length packets (incorporating IP and UDP overhead) at regular interval of 10 ms, 16.6 ms, 20 ms and 100 ms [2]. These times are defined by the sampling rate at which the measurements are planned. These intervals are distinct based on the constraints of the control applications and the frequency of the power cycle. At present, 100 Hz, 60 Hz, 50 Hz, and 10 Hz sampling periods are utilized [2]. This paper assumes a 50 Hz for both the power cycle and phasor sampling period. In the literature, PMUs latency constraints reported, which fluctuates between 8 ms and 100 ms [7] [8]. This is synchronous with the real-time control system requirements.

Mission critical APP2 and APP3 developed as Event-Driven applications, which implies just when an event occurs in the monitored environment, data is transmitted to the server and are provisioned using typical dynamic scheduling. As shown in Table II, packet sizes as well as inter-arrival times are exponentially distributed. The lowest priority App4 simulates Time-Driven application (SMs), where SM devices transmit data to the server consistently. The transmission interval of each Time-Driven device was uniformly distributed between 50 and 500 ms. All devices transmit their payloads (incorporating the 28 bytes IP/UDP header) in an unsynchronized fashion. Typical H2H applications are modeled using the traffic parameters shown in table III.

TABLE III.   H2H TRAFFIC CHARACTERISTICS

H2H Apps	Inter arrival time dist.	# Users	Data Rate (Kbps)	Uplink Load (Mbps)
Voice	Two states Markov	30	12.2	0.366
Video	Truncated Pareto	30	64	1.92
BE	Self-Similar	30	400	12

Uplink LTE utilizes a single carrier frequency division multiple access (SC-FDMA), eNB grants UE/device an adjacent quantity of resource blocks (RBs). A RB comprises 12 adjacent subcarriers (180 kHz) in the frequency domain and 1 TTI in the time domain. For the dynamic scheduling, every TTI, eNB computed the resource provision for the UEs/devices and then transmitted UL resource grants to the UEs/devices. These grants involve an adjacent set of RBs allotted to the UE/device accompanied by the modulation and coding scheme (MCS) as depicted in table IV.

TABLE IV.   MCS ZONES

Modulation Rate	Coding Rate	SNR (dB)	RB Rate (kbps)	RB Rate (Bytes)
64 - QAM	(3/4)	22	756	94.5
64 - QAM	(2/3)	14.1	672	84
16 - QAM	(3/4)	10.3	504	63
16 - QAM	(2/3)	5.9	448	56
QPSK	(3/4)	4.3	252	31.5
QPSK	(2/3)	6.7	224	28

The simulation parameters utilized in this paper are summarized in table V.

TABLE V.   SIMULATION PARAMETERS

Simulation Parameter	Value
System Bandwidth	20 MHz
Number RBs	100 RB
No. of M2M devices	2400
No. of H2H users	30
Number of MCS-Zones	6 zones
Modulation Schemes	64 - QAM, 16 - QAM and QPSK
Coding Schemes	(3/4) and (2/3)
Channel Model	FGN Multipath Fading model
Pathloss Model	$L(d) = 128.7 + 10\log(d)$
Carrier Frequency	2 GHz
Flows per user/device	3 flows per H2H user 1 flow per M2M device

4. PROPOSED HYBRID UL SCHEDULING ALGORITHM

Given that the emphasis is on commercially deployed 4G LTE systems, the proposed UL scheduling algorithm must fully conform to 4G LTE signaling and QoS standards. Thus, improvements established beyond Release 8 standards are not considered in this work. The hybrid-scheduling algorithm utilizes a conventional LTE's dynamic scheduling for supporting H2H services as well as M2M applications (App2, APP3, and APP4) and Semi-Persistent Scheduling (SPS) for delay sensitive IoT applications which always necessitate persistent radio resource distribution continuously (M2M APP1/PMUs).

A. Semi-Persistent Scheduling

The UE/device is conventionally scheduled dynamically on a per sub-frame basis, with the control information signaled on the Physical Downlink Control Channel (PDCCH). In this case, the UE/device is addressed using the cell radio network temporary identifier (C-RNTI). On the other hand, the UE/device may also receive a semi-persistent grant/allocation where the UE/device is addressed using the SPS-RNTI. In this case, the scheduling control information is signaled once via the PDCCH. The eNB preconfigures UE/device with an SPS-RNTI and a periodicity. Once preconfigured, and the UE/device receives an allocation using the SPS-RNTI (instead of the typical C-RNTI), then this allocation would be recurred according to the preconfigured periodicity [6]. This same configuration is used until modified or released. Thus, the UE/device is not required to request resources each subframe, saving a substantial control plane overhead.

The semi-persistent scheduling implies that the eNB can change the resource allocation type or location if required, for instance, for link adaptation. If a dynamic grant is received in the TTI marked for SPS data, the UE/device utilizes the radio resource specified by dynamic scheduling at that TTI and does not utilize the radio resource constructed by SPS, as the dynamic grant takes precedence [6].

It is assumed that each PMU device is configured with SPS-RNTI. Depending on the device location, eNB assigns a fixed number of RBs (e.g., 1, 2, 3 or 4) which is/are equivalent to the static PMU 100 bytes packet, and periodicity of 20 TTIs. Because there are 600 PMU devices, therefore, we divided them into 20 groups, each group of 30 devices. These devices were configured with a certain offset within the 20 TTIs period as it must wait for its specific offset TTI to transmit its buffer data. Since the PMU devices are fixed, the RB allocations and MCSs remain static for the current SPS configuration. In case of changes of the radio link condition, a new allocation (SPS configuration) will have to be sent on PDCCH. The normal dynamic schedule shall be used to schedule any HARQ re-transmission [6].

B. Dynamic Scheduling

The classic approach of mapping M2M applications to a newly introduced set of QCIs (radio bearers) necessitates introducing new LCGs, which requires modifications/ changes to the current LTE signaling mechanisms and QoS standards. Thus, we do not pursue this approach. The eNB RRM module is a critical component of the scheduling process since it performs the bearer control function that configures parameters that are specific to the uplink bearers. The RRM bearer control function manages the UE/device queue length via the PDCP discard timer, which is configured based on the PDB associated with each Application. Therefore, if there are any packets delayed beyond the allowed PDB limits, while waiting to be scheduled, are dropped.

The dynamic scheduling algorithm works as follows and as summarized in figure 7.

1) Allocate every M2M or H2H connection request to one and only one of the nine QCIs (radio bearer).

2) RRM groups the data and signaling bearers having common QoS constraints into to a maximum of four LCGs per UE/device.

3) Assume the following combination as shown in figure 6

    a) Data radio bearers (DRBs) with QCI 5 for IMS signaling and QCI 3 for time sensitive M2M APP1, will be allotted for LCG 0.

    b) DRB with QCI 1 for voice call and QCI 2 for M2M APP2, will be allotted for LCG 1.

    c) GBR DRB with QCI 4 for video and QCI 7 for M2M app3, will be allotted for LCG 2.

    d) Non-GBR DRB with QCI 6 for best effort traffic and QCI 9 for M2M APP4, will be allotted for LCG 3.
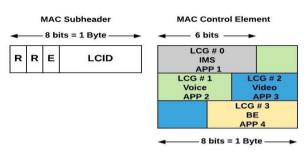

Figure 6. Mapping of QCI bearers to LCG

4) A classic UL scheduler orders UEs/devices flows based on the priority of the QCI allotted to a given application. The problem with this approach is that the UE/device reports to the eNB the volume of data pending transmission per group of QCIs. This kind of grouping permits the scheduling rules to be applied per LCG rather than per bearer/QCI. However, in the proposed algorithm, because RRM configures the Priority, PDB, and Prioritized Bit Rate (PBR) is done per uplink bearer, not per UE/device. UE/device utilizes these parameters to distribute the received uplink grant from eNB among bearers within LCG.

5) The PBR is allocated in proportion to the GBR rates. The principals of a token bucket algorithm is used to calculate the number of tokens credited to a given bearer, where every bearer is credited a number of tokens equivalent to PBR. Within a LCG, RRM allocates priority to the bearer as per the QCI priority. The received grant is allocated to the bearer with highest priority (M2M APP1) until all tokens are consumed, followed by another bearer in priority (M2M APP2) until tokens of all bearers in LCG 0 are served. Same steps are repeated within LCG 1, LCG 2, and finally within the lowest priority LCG 3, until either all resources are allocated, or bearers are served.
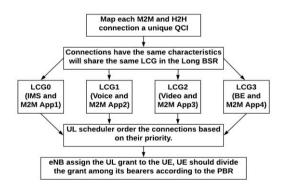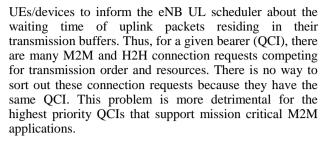

Figure 7. The proposed dynamic scheduling algorithm

The essential shortcoming of this simple approach, which strictly follow the 4G LTE signaling and QoS standards, is that the eNB RRM identifies the radio bearers comprised in the group and their priorities but does not have status of an individual bearer. LTE standard does not provide a mechanism that enables the

UEs/devices to inform the eNB UL scheduler about the waiting time of uplink packets residing in their transmission buffers. Thus, for a given bearer (QCI), there are many M2M and H2H connection requests competing for transmission order and resources. There is no way to sort out these connection requests because they have the same QCI. This problem is more detrimental for the highest priority QCIs that support mission critical M2M applications.

This problem can be addressed if the UE/device transmits to eNB a second BSR that contains HOL packet delay per bearer, as has been reported in [9], but once again the LTE signaling and QoS standards must be modified. The dynamic scheduling in this work addresses this problem by multiplying each UE/device connection request that have the same priority (QCI) by the following metric: (R_i/H_i), where (R_i) is the present data rate to be assigned to (UE_i or $device_i$) this cycle and (H_i) is the average assigned rate that is already granted to (UE_i or $device_i$) over the past 100 cycles. Multiplying by this metric only enhances fairness among M2M devices and UEs but does not address the critical timing problem.
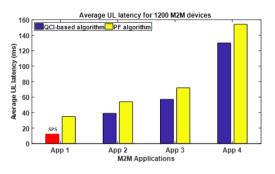
5. SIMULATION RESULTS

In the simulation, we have utilized two significant performance metrics: (1) the PLR (2) the average UL latency. To enumerate the communication link reliability measured between the communication source and destination, PLR is the most frequently utilized parameter in communication systems [8]. In this paper, the PLR will be specified for a given M2M application as every M2M application has its own distinctive performance constraints. PLR is defined as follows:

$$\text{PLR} = \frac{\text{\# of packets generated by the Source} - \text{\# of received packets with T<PDB}}{\text{\# of packets generated by the Source}} \quad (1)$$

The UL latency is identified as the time difference between the time when the packet entered the device transmission buffer to the time when the packet was transmitted to the eNB (i.e., it is just from the device to the eNB and does not include processing delay; not an E2E latency). The simulation results will be compared with a typical reference model, e.g., proportional fairness (PF).

Figure 8 shows the average UL latency (300 devices per M2M application for a total of 1200 M2M devices) of the hybrid (dynamic QCI-based scheduling and SPS algorithms) model versus PF model. As anticipated, the UL latencies for both dynamic and SPS was found to be lower than those of the PF for all M2M applications. It can also be realized from Figure 8 that the UL latency for 300 PMUs (APP1) is about 50% of its 20 ms PDB. However, the UL latency for each of the other 2 mission critical applications (APP2 and APP3) is almost within the PDB range of 40 ms and 60 ms, respectively.

Figure 8. Comparison between QCI-based and SPS-based scheduling and PF

Figure 9 shows the average UL latency for 300 devices (same as Figure 8 but on a different scale) and 600 devices per M2M application for a total of 1200 and 2400 M2M devices, respectively. As the number of devices increase up to 600 per application, none of the strategic smart grid applications can meet its own PDB requirement. The UL latency for 600 PMUs (APP1) is almost twice of that of the allowed 20 ms PDB.
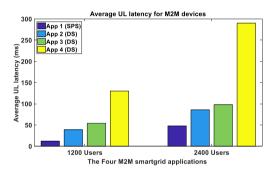

Figure 9. Average UL latency for each of the four M2M applications

Figure 10 depicts the average UL latency against the total quantity of M2M devices for all the four applications. As can be indicated from the Figure, the UL latency of the PMUs (APP1) is within the 20 ms PDB if the total number of devices does not exceed 1400 (350 PMU devices). The UL latency for each of the other two-time critical applications (APP2 and APP3) can meet the required 40 ms and 60 ms PDBs but for lower number of 1200 devices. As the number of total devices increase above 1500, the UL latencies for all applications increase rapidly and their performances are no longer satisfactory.

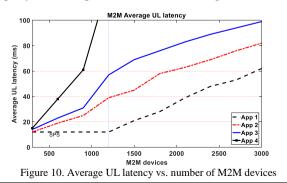Figure 10. Average UL latency vs. number of M2M devices

Figure 11 illustrates the PLR against the total amount of M2M devices for all the four services. As you can see from the Figure, the PLR of the PMUs can be as low as 10^{-6} but only if the total of devices does not exceed 800 devices (200 PMU devices). This is a significant result explicitly validates that strategic APP1 (PMUs) may satisfy just a distinct performance metric (meets PDB latency constraint) for 350 PMU devices (see Figure 9). To simultaneously support both high-reliability (10^{-6} PLR) and low latency, the number of supported devices drops to 200 PMUs. Same trend is applicable to both APP2 and APP3.
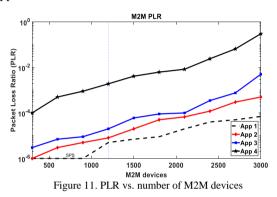

Figure 11. PLR vs. number of M2M devices

Figure 12 validates the average UL latency against the number of H2H UEs utilizing three H2H conventional applications. It can be perceived from the figure that the corresponding PDB for a total number of 30 H2H users and only 1200 M2M devices can hardly be met by the voice and video services. Figure 13 shows the average UL latency versus the PMU packet size for 300 and 600 PMU devices. As can be seen from the Figure, for a total of 300 PMU devices, PMUs cannot meet its 20 ms PDB as the packet size exceeds 140 bytes. As the number of PMU devices increase up to 600, the maximum packet size that cam meet the 20 ms PDB is about 60 bytes.
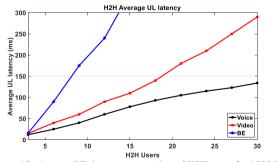

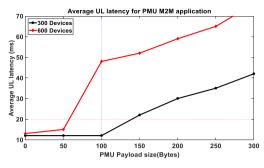Figure 12. Average UL latency vs. number of H2H users for 1200 M2M devices

Figure 13. Average UL latency vs. PMU packet size

Figure 14 displays the average UL latency for H2H Voice and Video applications [10] for the 30 UEs versus the total number of M2M devices. As can be seen for the Figure, both applications cannot meet their corresponding PDBs if the total numbers of M2M devices exceed 1200, in total agreement with the results of Figure 11.
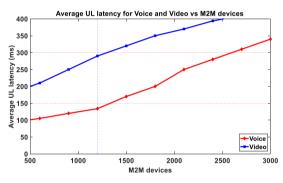

Figure 14. Average UL latency for Voice and Video versus number of M2M devices

6. CONCLUSION

This work has proposed and devised a simple hybrid UL LTE scheduling algorithm. This scheduling algorithm exploits a conventional LTE's dynamic scheduling for supporting H2H applications as well as M2M applications. It also utilizes a SPS to support time sensitive IoT applications that continuously necessitate a regular constant radio resource allotment. The simulation results designate that existing commercial 4G LTE networks potentially have to sufficiently provide some of the evolving strategic smart grid services including PMUs which necessitates 20 ms latency. The number of supported devices that can adequately meet such latency (PDB) requirements is about 300 devices per each of the four supported M2M applications along with few tens of H2H users [11].

To simultaneously meet high reliability (PLR = 10^{-6}) and low latency, the total number of supported PMU devices drops to almost one-half. If achieving PLR $< 10^{-6}$ is the criteria for extreme-high reliability, then commercial 4G LTE networks cannot concurrently support both extreme-high reliability and low latency.

In this work, the PDB is assumed to be 20 ms, which considers only the communication link delay within the Radio Access Network (RAN), i.e., device-to-eNB [12]. Reasonably speaking, considering the application level processing latency together with latency within the mobile core, the E2E delay would be in the limit of about 50-100 ms. Thus, a more accurate approximation is to claim that current commercial 4G LTE networks theoretically have to effectively support some of the developing time sensitive smart grid applications (or any other similar IoT services) including PMUs, with realistic E2E latency constraints in the range of 50-100 ms and extreme reliability.

The practical results of this work can be utilized by industrial corporations and utilities, which are planning to utilize commercial 4G LTE networks for supporting their private IoT services, as preliminary strategies to ensure that LTE can support the performance necessities of these applications.

REFERENCES

[1] C. Beard, "High Reliability 4G and 5G Cellular Wireless Services for Smart Cities," in *IEEE smart cities ISCW-KC-2016*, Kansas city.

[2] F. B. a. K. Basu, "Performance Evaluation of Time-Critical Smart Grid Applications," in *Proceedings of the Eleventh International networking conference (INC)*, Frankfurt, 2016.

[3] S. Y. A. a. H. A.-R. Ali Othman, "Dynamic Switching of Scheduling Algorithm for Uplink LTE System," *International Journal of Computing and Network Technology,* vol. 6, no. 3, Sept 2018.

[4] A. M. T. M. S. a. H. H. N. Abu-Ali, "Uplink Scheduling in LTE and LTE-Advanced:Tutorial, Survey and Evaluation Framework," *IEEE COMMUNICATIONS SURVEYS & TUTORIALS,* vol. 16, no. 3, pp. 1239-1265, 2014.

[5] "5G Ultra-reliable low-latency communications," *IEEE Network Magazine (Special Issue),* vol. 32, no. 2, 2018.

[6] S. Ahmadi, LTE-Advanced A Practical Systems Approach to Understanding the3GPP LTE Releases 10 and 11 Radio access technologies, Elsevier, 2014.

[7] I. 61850-5, "Communication networks and systems in substation-Part 5: Communication requirements for functions and device models," International Electrotechnical Commission, Geneva - Switzerland, 2003.

[8] S. a. r. n. f. s. e. d. (SUNSEED), "D3.1 Traffic modelling, communication requirements … for real-time smart grid control. Version2.0," 2015.

[9] N. Afrin, J. Brown and J. Y. Khan, "A packet age based LTE uplink packet scheduler for M2M traffic," in *7th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Carrara, VIC, Australia, 2013.

[10] S. Y. A. A. A.-O. a. H. A.-R. Ali Othman, "Comparative Performanceof Subcarrier Schedulers in Uplink LTE-A under High Users' Mobility," *International Journal of Computing and Digital Systems,* vol. 4, no. 4, 2015.

[11] A. Y. Hassebo, Commercial 4G LTE Cellular Networks for Supporting Emerging Mission-Critical IoT Applications, New York: The City College of New York, ProQuest Dissertations Publishing, 2018. 13425690., 2018.

[12] A. R. a. M. A. A. ". I. U. 2. N. A. Hassebo, "A Hybrid LTE Uplink Scheduling Approach for Supporting Mission-Critical Smart Grid applications," in *2018 IEEE 9th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York, 2018.

[13] 3. T. V. 36.211, "Evolved Universal Terresterial Radio Access (E-UTRA): Physical channel and modulation Release 8," 2009-12.

[14] J. B. a. J. Y. Khan, "Performance analysis of an LTE TDD based smart grid communications network for uplink biased traffic," in *IEEE Globecom Workshops*, Anaheim, CA, USA, 2012.

[15] J. B. a. J. Y. Khan, "Performance comparison of LTE FDD and TDD based Smart Grid communications networks for uplink biased traffic," in *IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, Tainan, Taiwan, 2012.

[16] A. E. Mostafa and a. Y. Gadallah, "A statistical priority-based scheduling metric for M2M communications in LTE networks," *IEEE Access,* vol. 5, pp. 8106 - 8117, 2017.

[17] A. S. Alexiou and a. A. Lioumpas, "Uplink scheduling for Machine-to-Machine communications in LTE-based cellular systems," in *IEEE GLOBECOM Workshops (GC Wkshps)*, Houston, TX, USA, 2011.

[18] N. Afrin, J. Brown and a. J. Y. Khan, "Performance analysis of an enhanced delay sensitive LTE uplink scheduler for M2M traffic," in *Australasian Telecommunication Networks and Applications Conference (ATNAC)*, Christchurch, New Zealand, 2013.

[19] A. E. a. Y. Gadallah, "BAT: A Balanced Alternating Technique for M2M Uplink Scheduling over LTE," in *81st Vehicular Technology Conference (VTC Spring)*, Glasgow, UK, 2015.

[20] K. W. a. Y. T. Y. Hsu, "Enhanced cooperative access class barring and traffic adaptive radio resource management for M2M communications over LTE-A," in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, Kaohsiung, Taiwan, 2013.

[21] S. Lien, K. Chen and a. Y. Lin, "Toward ubiquitous massive accesses in 3GPP machine-to-machine communications," *IEEE Communications Magazine,* vol. 49, no. 4, pp. 66 - 74, April 2011.

[22] I. A. a. S. Venkatesan, "A QoE preserving M2M-aware hybrid scheduler for LTE uplink," in *International Conference on Selected Topics in Mobile and Wireless Networking (MoWNeT)*, Montreal, QC, Canada, 2013.

[23] N. R. A. C. K. V. S. a. B. R. T. M. Giluka, "Class based dynamic priority scheduling for uplink to support M2M communications in LTE," in *IEEE World Forum on Internet of Things (WF-IoT)*, Seoul, South Korea, 2014.

[24] T. d. Andrade, C. Astudillo and a. N. d. Fonseca, "Impact of M2M Traffic on Human-type Communication Users on the LTE Uplink Channel," in *7th LATINCOM*, Arequipa, Peru, 2015.

[25] Y. H. C. X. a. L. H. S. Zhenqi, "Research on uplink scheduling algorithm of massive M2M and H2H services in LTE," in *IET International Conference on Information and Communications Technologies (IETICT 2013)*, Beijing, China, 2013.

[26] J. Brown and a. J. Y. Khan, "Predictive resource allocation in the LTE uplink for event based M2M applications," in *IEEE International Conference on Communications Workshops (ICC)*, Budapest, Hungary, 2013.

[27] A. A. a. G. Miao, "Lifetime-aware scheduling and power control for M2M communications in LTE networks," in *Proceeding IEEE Vehicular Technology Conference (VTC-Spring)*, Glasgow, UK, 2015.

[28]   S. R. Y. J. N. H. a. M. Z. W. Y. B. Chen, "'an energy- efficient scheduling algorithm for real-time machine-to-machine (M2M) data reporting," in *Proceedings of IEEE Global Communications Conference (GLOBECOM),*, Austin, TX, USA, 2014.

[29]   A. A. a. A. Aghvami, "On Radio Resource Allocation in LTE Networks with Machine-to-Machine Communications," in *IEEE 77th Vehicular Technology Conference (VTC Spring)*, Dresden, Germany, 2013.

[30]   D. F. a. T. S. L. Wang, "A Bandwidth Bargain Model based on Adaptive Weight Fair Queuing," 2006.

Ahmed Hassebo is a visiting Assistant professor at Electrical and Computer Engineering department in Purdue University Northwest. Dr. Hassebo has been awarded MPhil and PhD degrees of EE from The City College of The City University of New York in 2016 and 2019, respectively. He also awarded the BSc of EE from Al-Azhar University, Egypt. His research interests including 4G/5G dynamic bandwidth allocation algorithms, and IoT mission critical applications.

Sumeet Chaudhary is a Visiting Assistant Professor in the Department of Electrical and Computer Engineering. He has earned his PhD. in Electrical Engineering from the University of Cincinnati. His dissertation involved developing a lightweight Electromagnetic Motor. The approach was to use composite manufacturing technology to reduce "touch labor" and produce a device which is simpler to manufacture than using conventional metal lamination techniques.

Appendix

In this appendix, we would like to elaborate in explaining the resource block (RB) assignment method. As given in step # 5 in table.

The eNB transforms the incoming data from every single flow to the proper instant data rate using the following equation.

$$R_i^{new} = \left(1 \ e^{-\frac{t_i^k}{N}} \right) \frac{l_i^k}{t_i^k} \ e^{-\frac{t_i^k}{N}} R_i^{old} \qquad (2)$$

R_i^{new} is the instant data rate for i^{th} flow l_i^k and t_i^k is the length of time of kth packet of ith flow. N is a constant between 100 and 500 ms, we use N = 300 ms.

After the instant data rate was obtained, the expected number of RBs per flow can be calculated using the following method depending on the MCS zone:

$$Req.RBs = \sum_{i=1}^{72} floor \left(\frac{Instantenous\ data\ rate}{MCS\ Zone\ Rate} \right) + 1 \qquad (3)$$

$$MCS\ Rate = \frac{\#\ OFDM\ sym}{TTI} * \frac{\#\ Subcarriers}{RB} * \frac{\#\ bits}{Sym} * \#\ RB \qquad (4)$$

$OFDM\ sym$ = 14 OFDM symbols per TTI (i.e., subframe), where TTI = 1 ms.
$Subcarriers$ = 12 Subcarriers per RB.
$bits$ = Number of bits per symbol
Depends on the modulation scheme