Uplink Contention Based Multiple Access for 5G Cellular IoT

Dr Debby Lin, Dr Gilles Charbit, Dr I-Kang Fu
MediaTek Inc.
Email: {Debby.Lin, Gilles.Charbit, IK.Fu}@mediatek.com

***Abstract*—This paper introduces novel UL access methods for 5G cellular IoT where devices can directly send a scheduler grant or a small data packet in the contention-based (CB) manner. The novelty of this work is in the combination of medium access control (MAC) contention access with physical layer optimization with new numerology applied to single carrier frequency division multiple access (SC-FDMA) waveforms to remove the need for pre-synchronization of devices on the uplink (UL). An analysis of the new procedures compared with the current LTE random access procedure is provided. The simulation results show that over 50% latency reduction and about 500% capacity increment can be achieved by contention based scheduling request (CB SR) compare to the LTE UL grant via random access. Further, it is shown that CB data is a viable method to transmit small data packets compare to LTE scheduling request via random access followed by scheduled data transmission.**

Keywords—Internet of Things, Long Term Evolution, Contention-Based Access, Random Access, SC-FDMA

I. Introduction

Internet of thing (IoT) communications with massive number of connections is widely expected to be one of the key research challenges for the 5th generation (5G) mobile cellular network. IoT air interface will need to support both high connection density for massive IoT devices and low radio latency for reliable communication applications [1, 2]. In addition, the network coverage is critical to connect various IoT devices in the field and thus it is more reasonable to support the IoT communications over lower frequency bands (e.g. <6 GHz). In 3GPP standards, the Rel-13 LTE MTC (machine type communication) feature will enable a 1.4 MHz compatible carrier which could be overlaid within 20 MHz LTE signal without interference [3]. This creates the room for physical layer optimization based on the machine communication requirements in order to achieve the better coverage and the lower power consumption.

II. IoT Data Reporting via LTE Random Access Channel

A. Background

In LTE, devices can obtain UL grant through either dedicated scheduling request (SR) or random access (RA) procedure. SR allows multiple users multiplexing together with different cyclic shifts and orthogonal sequences to occupy the dedicated resource elements on physical uplink control channel (PUCCH). At most 36 devices can be multiplexed together for SR in one physical resource block (PRB) on PUCCH. Assume one out of six PRB is allocated for SR per PUCCH subframe for machine communications which gives about 16% SR overhead for UL grant and further assume the SR periodicity is 80 ms (the largest specified SR periodicity in LTE), a cell can support up to 36*80=2880 connections at most. In practice, there may need to be a compromise for PUCCH detection performance, with 18 or 12 SR opportunities per PUCCH frame instead of 36. This may result in smaller number of machines supported per cell. This way doesn't scale up well to massive IoT connections. For example, assuming up to 40 devices in a home in the future (the 3GPP assumption in GERAN release-13 Cellular IoT study Item) and about 6000 houses per km^2 in urban London, there could be about 240,000 devices per km^2 or about 190,000 devices per cell in a typical cell size of 500 meters [4, 5]. There may be even larger number of devices per km^2, 10^6 – 10^7 devices per km^2, being considered in 5G. To support massive IoT, either very large PUCCH resource should be configured for providing SR opportunities for UL grant requests, or the periodicity of SR should be very large. This could result in the scheduling delays of several minutes, which is too long for machines that need to be triggered to send some data (e.g., burglar alarm, health monitoring, and temperature reading.).

A 4-step Random Access procedure is designed in LTE for initial cell access and UL grant request once if SR is not configured, which is illustrated in Figure 1. User equipment (UE) sends a random access preamble with embedded 1-bit indication of required message size needed for further signaling. The eNB assigns a UE identity (ID), provides an UL grant, and UL transmission timing in the random access response. Further signaling exchanges in step 3 and 4 are carried out to prepare the connection for UL data transmission and to resolve the contention. The minimum latency for the RACH procedure is at least 15 msec excluding the step 1 waiting time. From LTE RACH procedure, large signaling latency and signaling overhead (at least 4 steps handshaking) are needed in order to get a small packet data transmitted, which is not effective and efficient for the massive IoT UL data transmissions.

978-1-4799-8091-8/15/$31.00 ©2015 IEEE

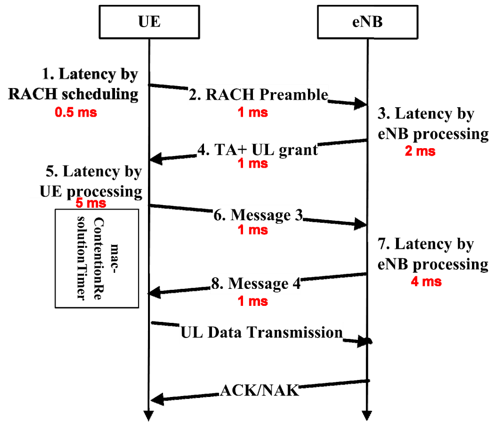

Figure 1: Small data reporting via Random Access procedure

When supporting cellular IoT applications, the LTE network implementation could configure more resources for RACH and allow IoT devices to report small data through message 3 (MSG3). This could be quite useful for the IoT applications such as sensor reporting, where each of the massive IoT devices only need to report small amount of data infrequently. By avoiding the unnecessary PDCCH decoding and PUSCH resource allocation, the IoT devices could quickly transmit the small data through RACH and then back to power saving mode. In the following analysis, this is treated as the baseline mechanism for the performance benchmark.

B. Proposed Contention Based (CB) UL Access Mechanism

The concept of data RACH procedure outlined in [7] considered a random access preamble sent with data. Further research in [8] investigated the impact by pilot structure and power allocation. UL contention-based access for LTE was proposed in [9] to reduce the latency compared to scheduled access, where an UL grant shared by a group of UEs is used to transmit data packets. The contention happens naturally when multiple UEs use the same radio resources simultaneously. In scheduled access, such contention does not occur as eNB allocates dedicated resource for each device transmission on the UL via a UE-specific UL grant. In this paper, contention-based UL access is proposed with optimized physical layer structure design with new numerology applied to SC-FDMA waveform without the need for pre-synchronization of UEs via UL timing alignment procedure specified in LTE.

The basic concept of the proposed solution is to simplify the UL data transmission procedure by (a) include a small signaling payload (e.g. UE ID) as a scheduling request (SR) for UL resource grant or (b) directly include a small data payload in the first step of the random access procedure as illustrated in Figure 2a and 2b respectively. Option (a) is suitable for LTE revolution with further modification on RACH, while option (b) is more suitable for new system design (e.g., 5G) by defining a new physical channel. In option (b) UE can directly send small data packet in contention-based manner, then eNB will confirm the UL packet data reception within the contention resolution message with an ACK/NACK for the transmitted UL packet. Option (a) and (b) are abbreviated as CB SR and CB data in the following description in this paper.

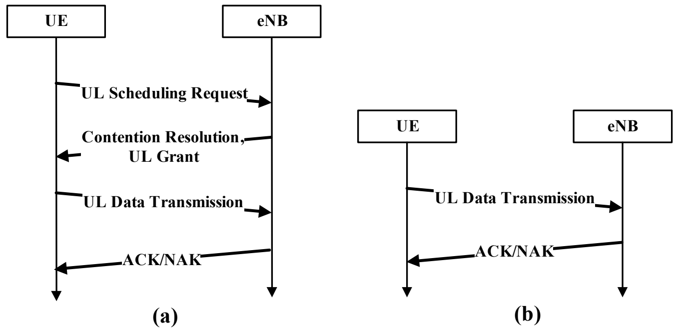

Figure 2: Data reporting via optimized Random Access procedure.

Essential to simplifying the legacy RACH procedure is to remove the need of the timing advance (TA) adjustment to pre-synchronize devices on the UL. The simplified procedure has the added benefits to reduce the complexity and the battery drain in low-cost IoT devices. Different physical layer structure design is necessary to support this procedure. As an example of the numerology design, the SC-FDMA waveform format can be used at the UE transmitter side with a subcarrier spacing of 3.75 kHz instead of the 15 kHz used in the current LTE system in UL [10]. The symbol length is 1/3.75 kHz=266.67 us. A cyclic prefix (CP) of 66.67 us can be added, which is much longer than that 4.7us CP used in LTE. With this waveform numerology, a 1-msec subframe can accommodate 3 SC-FDMA symbols including CP. The CP of 66.67 us tolerates time misalignment in the cell radius up to 10 km (i.e. assuming velocity of light $3 \cdot 10^8$ m/s, 2 way propagation distance is $3{\cdot}10^8 * 66.67{\cdot}10^{-6} /2 = 10$ km). Hence, the eNB doesn't need to send the timing advance command to the UE. By this way, the simplified 2-step contention-based UL access procedure can be achieved, where the random access preamble and the random access response messages can be omitted.

III. Simulation Setup

A. Physical abstraction modeling

A Physical abstraction model shown in Figure 3 is derived from link level simulations (LLS) of a SC-FDMA transceiver system with ultra narrow band waveform. The waveform numerology used was described in the previous section. At the UE, an 8-bit CRC is added to a 16-bit UE ID. The data payload passes a tail biting convolutional encoding (TBCC) with generator polynomial G=(133, 171, 165) and code rate 1/3. TBCC coding has lower complexity than turbo coding, and is known to outperform or have similar performance as Turbo coding for small data payload less than 250 bits [12]. After that, the QPSK modulated signals go through fast Fourier transform (FFT), resource element (RE) mapping and inverse FFT to generate the SC-FDMA waveform. A CP is then added to preserve the circular convolution property before transmission through a channel. At the eNB receiver, the CP is firstly removed and the received signal further passes FFT, RE demapping and IFFT processing, QPSK demodulation, and TBCC decoding. If CRC check is successful, the 16-bit UE ID and/or the small data payload may be retrieved (this may include a Buffer Status Report to allow further data transmission on scheduled access). The block error rate (BLER) simulation curves as a function of signal to noise ratio (SNR) obtained from LLS are used as the input to the system level simulator (SLS) which is further described below.

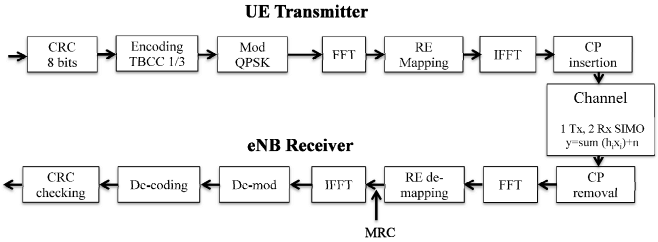

Figure 3. SC-FDMA Transceiver chain

B. *System level modeling*

A typical hexagonal cell layout with 21 cells in the SLS as shown in Figure 4 is used to simulate the number of machines transmitting a small data on CB UL access. All the devices are uniformly distributed in the cells. The devices with UL data to be transmitted will randomly choose a resource slot for the contention. Given a UE i in the UE set **I**, where I is the set composed the indices of those UEs that have UL data to transmit. The signal-to-interference-plus-noise ratio (SINR) of UE i can be expressed as

$$\mathrm{SINR}_i = \frac{\mathrm{P}_{i,\mathrm{R}(i)}\mathrm{L}_{i,\mathrm{B}(i)}}{\sum_{\substack{k\in \mathrm{I} \\ k\neq i}} \mathrm{P}_{k,\mathrm{R}(i)}\mathrm{L}_{k,\mathrm{B}(i)} + \mathrm{N}_0}$$

where B(i) is the serving eNB of UE i, R(i) is the resource slot chosen by the UE i for the contention, $\mathrm{P}_{i,j}$ is transmitted power of UE i chosen resource slot j, $\mathrm{L}_{i,j}$ is the large scale fading from UE i to the base station j and N_0 is the noise power. Note that $\mathrm{P}_{i,j}$ is determined by reaching a target SNR, SNR_0, at the serving eNB of UE i and is capped by the maximum transmit power P_{max} = 23 dBm. Here $\mathrm{P}_{i,j}$ =0 if $j \neq \mathrm{R}(i)$.

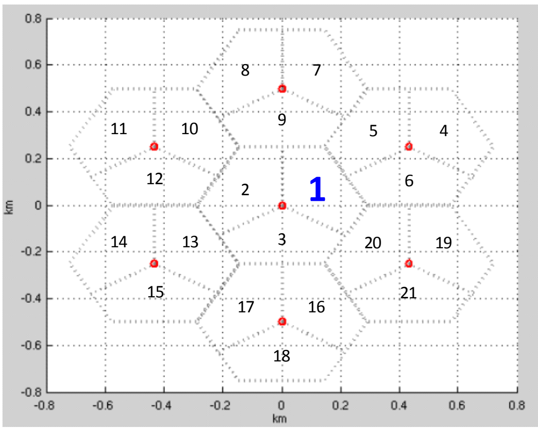

Figure 4. 21-cell SLS simulator layout.

C. *Contention-based UL resource configuration*

A traffic model with uniform distribution of UL data transmissions over a period 60 seconds was used [5]. The packet arrival distribution for CB scheduling request (SR) are illustrated in Figure 5(a) and 5(b), respectively. In Figure 5(a), 16-bit UE ID plus 8-bit CRC are encoded and modulated into 72 bits (36 QPSK symbols). With narrower subcarrier spacing 3.75 kHz, 48 subcarriers can be created in one PRB (180 kHz) and 3 OFDM symbols can be arranged in one millisecond. Assume 20% out of 6 PRBs resources is used for the ID contention. Here we arrange 3 PRBs and 3 msec frequency-time resources to create 36 contention slots and the periodicity of contention resource is 7.5 msec. Thus, 36 QPSK symbols occupy 4 subcarriers and 9 OFDM symbols (3 msec). In Figure 5(b), small data packet with 192 info bits including 16-bit UE ID, 168-bit data, and 8-bit CRC are encoded into 576 bits (288 QPSK symbols). Assume 80% out of 6 PRBs resources is used for the data contention. 6 PRBs and 12 msec frequency-time resources is used to create 36 contention slots and the periodicity of contention resource is 15 msec. Packet size with 168 bits of data (21 bytes) would be sufficient for most periodical smart reader reports, which are typically 4 bytes, with some special reports being larger sizes – e.g. 18 bytes for system failure [11].

When a device awakes from DRX state, it has to wait until there are CB slots available for signalling or data payload transmissions. If the transmission fails, device has to wait until the next available CB slots after some processing delay in the UE and eNB. These transmission opportunities are illustrated by arrows for the 1st transmission and the 2nd transmission (if the first transmission fails) in the figures.

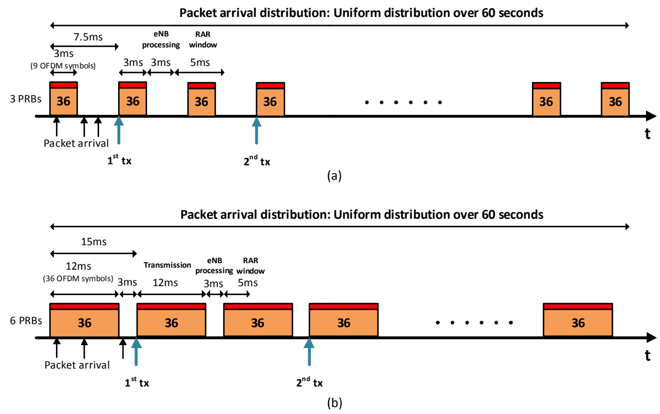

Figure 5. Example of packet arrival distribution for (a) CB SR and (b) small data packet

IV. SIMULATION RESULTS

The SLS simulation results for the proposed CB SR (i.e. Figure 2(a)) and the SR via Rel-8 LTE RACH are shown in Figure 6. In LTE RACH scenario 1, the LTE eNB is assumed not to be able to decode a RA preamble transmitted by multiple devices. In LTE RACH scenario 2, it is assumed that a practical receiver may be able to detect the stronger received RA preamble even when the collision occurs. However, since all the UEs transmitting the same preamble receive the same random access response (RAR) and then transmit their MSG3 in the same allocated UL resource, thus the transmission and the re-transmission of MSG3 may collide all the way until the maximum number of MSG3 transmission is reached. This is the worst case.

A similar methodology was used for the latency determination of proposed 2-step CB SR procedure. CB SR with packet size of 24 bits before encoding is shown to outperform both SR with LTE RACH scenarios. Figure 6 shows that around 160,000 devices transmit a CB SR successfully within a 30 ms latency as compared to about 40,000 and 10,000 for LTE RACH scenario 1 and 2, respectively. On the other hand, to accommodate 160,000 devices in the cell, the access latency of CB SR is 30 ms while the latency of LTE RACH scenario 1 and 2 are 45 ms and greater than 90 ms, respectively. Assuming for LTE RACH scenario 1, CB SR provides a fourfold increment in signalling capacity for SR with a latency of 30 ms. On the other hand, with 160,000 devices in the cell, CB SR provides a 33% latency reduction. SLS simulations showed that further latency reduction of over 50% can be achieved for smaller number of devices.

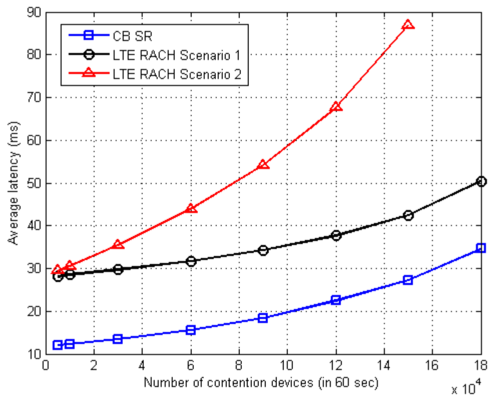

Figure 6. Latency performance improvement by proposed CB SR

The simulation results for the proposed CB data (i.e. Figure 2(b)) with 192-bit packet size are shown in Figure 7 to Figure 9. Three cases are simulated. The maximum number of transmission (1st transmission and re-transmissions) was set to 4 and 10. The target SNR was set to 5 dB and 8 dB. The tolerable packet failure was set to 10%. The capacities of case 1, 2 and 3 are shown to be 53000, 61000, and 68000 devices per cell, respectively, with the best performance obtained for a maximum number of transmissions set to 10 and the target SNR of 8 dB. The average number of transmissions to get the packet decoded successfully is shown in Figure 8. With the achievable capacitates under the tolerable packet failure rate 10%, the average transmission numbers of case 1 (53000 devices), 2 (61000 devices) and 3 (68000 devices) are 1.85, 1.84 and 3.48 transmissions, respectively. Figure 9 illustrates the average latency of three cases.

Sending similar packet data size in LTE would require devices to obtain a scheduling grant via RA procedure and then transmit data using scheduled access. The SR via RA procedure requires 1 DL PRBs per device for message 4, 1 UL PRB per device for message 3, and 2 PRBs for transmission of 576 coded bits. Assuming we treat DL PRB for message 4 as an equivalent UL PRB, the LTE benchmark may require 4 PRBs total for SR via RA procedure plus scheduled data. This is in addition to the shared resources for transmission of RA preamble and RAR. Assuming 80% of 6 PRBs available for data, the LTE benchmark can support a maximum of 0.8*6*1000*60/4=72,000 devices in 60 seconds.

On contention access, a small data packet of 576 bits requires 2 PRB for data assuming QPSK modulation and 1 PRB for contention resolution or 3 PRBs in total. Assuming 80% of 6 PRBs available for data, CB data can support 0.8*6*1000*60/3= 96,000 CB opportunities. The actual number of devices per 60 seconds supported depends on contention collision and detection performance. The results in Figure 7 and 9 would suggest that a capacity of 40,000 devices can be supported with a latency of 40 ms and a Packet Error Rate (PER) of ~3% or lower.

For the fair comparison, capacity of LTE benchmark, capacity at similar 40 ms latency should be considered. Latency for benchmark is latency of SR via random access and an additional scheduling delay needed in UE to decode message 4 (including UL grant) and prepare transport block for UL transmission. Some re-transmission delay should also be taken into account, as CB data simulations in SLS setup include re-transmissions. Assuming scenario 2 in Figure 6 and including scheduling delay of ~10 ms, approximately 40,000 devices could be supported at latency of 40 ms. This implies capacity of CB data is similar to that of LTE benchmark, but has no scheduler latency. The scheduling latency could become significant in case more machines make an SR via RA procedure than can be immediately served by scheduler to transmit data packet on scheduled resources at any given time.

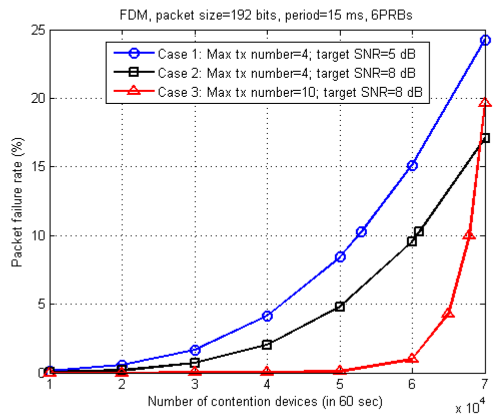

Figure 7. Packet failure rate performance of proposed CB data access channel

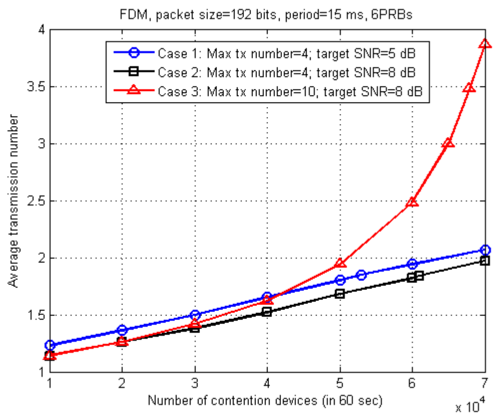

Figure 8. Average transmission number of proposed CB data access channel

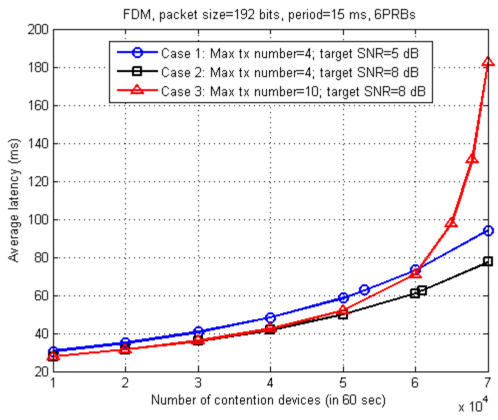

Figure 9. Average latency of proposed CB data access channel

The SLS simulations would suggest that packet size of 192 bits before coding (576 bits after 1/3 coding) is the breaking point beyond which CB UL access does not outperform the LTE benchmark. It is believed that this is due to a 2 dB BLER performance loss of TBCC decoder when packet size is increased from 24 bits (information part before 1/3 coding) to 192 bits as shown in Figure 10. Further improvement of CB UL access could be achieved by improving the performance of eNB receiver for larger packet sizes with use of a turbo code instead of TTBC (this should provide 1 dB gain according to [12]) and interference cancellation in case of UL transmissions with collision. Coded random access (CRA) proposed in [13] to reduce the collision probability and compressive sensing based multi user detection (MUD) proposed in [14] to mitigate interference in case of UL transmissions with collision would be examples of coding and MUD methods to be further considered to improve the CB UL access. This is beyond the scope of this work.

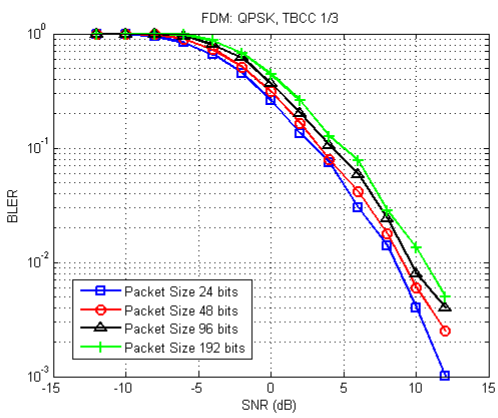

Figure 10. BLER performance for PHY abstraction model – packet size is before 1/3 TBCC coding

V. SUMMARY

This paper provided an analysis of contention-based access for scheduling request and data transmission for 5G Cellular IoT. The simulation results show that over 50% latency reduction and a 400% increase in capacity can be achieved by CB SR as compared to LTE benchmark of SR via random access procedure. Further, CB access is shown to be a viable method for small data packets compare to LTE benchmark of SR via random access procedure followed by the scheduled data transmission. The proposed method considered new numerology applied to SC-FDMA waveform to remove the need of pre-synchronization of devices on the UL. Further improvement of CB UL access could be achieved by improving the performance of eNB receiver for larger packet sizes with the use of a turbo code instead of TTBC and interference cancellation in case of UL transmissions with collision. The research showed that UL contention-based access is a promising concept to substantially improve the system performances in order to achieve the challenging latency and the capacity requirements for 5G cellular IoT.

References

[1] 5D/TEMP/548-E, "IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond", ITU-R WP5D, Feb. 2015

[2] NGMN 5G Initiative White Paper, Mar. 2015.

[3] 3GPP R1-142919, "Narrow band LTE for MTC in LTE Rel-13", MediaTek, RAN1#78, Aug. 2014

[4] 3GPP GP-140421, New Study Item on Cellular System Support for Ultra Low Complexity and Low Throughput Internet of Things, GERAN#62, May 2014

[5] 3GPP TR 37.868 Study on RAN Improvements for MTC

[6] Stefania Sesia, Issam Toufik, Matthew Baker, Wiley, "LTE – from theory to practice ", ISBN 9780470660256J.

[7] Thorsten Wild and Al, 5GNOW: Intermediate Transceiver and Frame Structure Concepts and Results, EuCNC 2014, At Bologna

[8] G. Wunder, P. Jung, C. Wang, "Compressive Random Access for Post-LTE Systems", accepted at IEEE International Conference on Communications (ICC), Sydney, Australia, June 10-14, 2014

[9] 3GPP R2-093812, "Contention based uplink transmissions", Ericsson, RAN2#66bis, July 2009

[10] 3GPP TS 36.211 "Physical Channel and modulation"

[11] Engage Consulting ltd, "High-Level Smart Meter Traffic Data Analysis", May 2010

[12] 3GPP GP-140867, NB M2M - Performance Comparison of Convolutional Coding and Turbo Coding for UL transmission, Huawei, GERAN#64, Nov 2014

[13] METIS 2020, Deliverable D2.4, Proposed solutions for random access, February 2015

[14] H. Schepker, "Improving Greedy Compressive Sensing Based MUD with Iterative Feedback, VTC Fall 2013