Impact of Preamble-Priority-Aware Downlink Control Signaling Scheduling on LTE/LTE-A Network Performance

Carlos A. Astudillo, Tiago P. C. de Andrade, and Nelson L. S. da Fonseca
Institute of Computing - State University of Campinas, Brazil
castudillo@lrc.ic.unicamp.br, tiagoandrade@lrc.ic.unicamp.br, nfonseca@ic.unicamp.br

***Abstract*—The concept of preamble-priority awareness in downlink control signaling scheduling was recently proposed to provide Quality of Service (QoS) differentiation in the Random Access (RA) procedure of the LTE/LTE-Advanced technology. This approach employs the information about preamble-priority levels used in the the initial phase of the RA procedure to schedule random access response messages. In this paper, we extend the application of this concept to the scheduling of other control messages and analyze its impact when the RACH Resource Separation (RRS) scheme and the traditional RA scheme with both the contention-free and the contention-based modes are used under heavily-loaded, highly-synchronized Machine-Type Communications (MTC) scenarios. Our results derived via extensive simulations show that the preamble-priority-aware concept provides QoS differentiation to users utilizing contention-free preambles. Furthermore, this concept helps to achieve the goal of isolation between traditional LTE users and MTC devices when the RRS scheme is used.**

Keywords*—*Handover, LTE-A networks, machine-to-machine communications, PDCCH scheduling algorithm, random-access procedure.

I. INTRODUCTION

In the fifth generation (5G) cellular networks, the support of Massive Machine-Type Communications (mMTC) will be mandatory in order to attend the expected Internet of Things (IoT) market. Such networks will use a combination of the evolution of the Long Term Evolution (LTE)[1] technology as well as new radio technologies in future scenarios. Machine-Type Communications (MTC) enables communication between geographically distributed devices on the Internet, generating traffic with reduced or even without human intervention [1].

LTE is a key technology for IoT connectivity giving its high capacity, security and ubiquity coverage. However, introducing a huge number of MTC devices transmitting sporadically small packets would put a very high pressure on the current cellular networks. Such a traffic pattern leads to shortages of radio resources, specially those associated to the Random Access (RA) procedure, which is used by User Equipments (UEs) to establish network connection, perform handover, and request uplink resources. Thus, MTC can severely jeopardize the overall network performance for both traditional Human Type Communication (HTC) users and MTC devices [2].

The LTE RA procedure can be executed by the UE either in the contention-free mode or the contention-based mode. In both modes, the UE transmits a preamble sequence (msg1) on the Physical Random Access Channel (PRACH). This preamble sequence is either explicitly allocated by the evolved NodeB (eNB) (contention free) or randomly selected from a set of available preamble sequences by the UE (contention based). Upon detecting a preamble sequence, the eNB creates a Random Access Response (RAR) message (msg2) to be transmitted on the Physical Downlink Shared Channel (PDSCH), containing an uplink grant for a transmission on the Packet Uplink Shared Channel (PUSCH). To indicate the transmission of the RAR messages, a Downlink Control Information (DCI) message on the Packet Downlink Control Channel (PDCCH) is required. If no response is received, the UE enters a backoff period and then tries to transmit a new preamble sequence. This is repeated until the UE receives an RAR message or until the maximum number of preamble sequence transmissions is achieved. Once the RAR message is received by the UE, it transmits an L2/L3 message (msg3) on the PUSCH as indicated in the uplink grant. The msg3 message contains information on the reason of the RA procedure triggering. If two or more UEs chose the same preamble sequence index in a certain Random Access Opportunity (RAO), they will receive the same RAR message. In such a case, all msg3 transmissions from the involved UEs will collide. A UE re-initiates the RA procedure after achieving the maximum number of msg3 retransmissions. Upon successful reception of a msg3 message, the eNB creates a Contention Resolution (CR) message (msg4) to be transmitted on the PDSCH. The successful reception of the CR message finishes the contention-based RA procedure. In the contention-free mode, the reception of the RAR message finishes the RA procedure.

Some Quality of Service (QoS) differentiation during the RA procedure can be provided to a particular user or a reduced group of users, mainly by reserving a set of preamble sequences for these users. By doing this, collision of preamble transmissions are reduced (contention based) or even avoided (contention free). However, as described above, this procedure also involves radio resources from PDCCH [2] as well as the data channels [1] and the available resources on these channels may not be sufficient to attend the huge demand in the presence of massive access attempts [3].

---

[1] we use LTE throughout the paper to refer to all technologies based on 3GPP LTE standards (release 8 and beyond).

[2] we use PDCCH to refer to all packet downlink control channels as defined by the 3GPP.

978-1-5090-5935-5/17/$31.00

Most of the existing PDCCH scheduling algorithms do not differentiate users utilizing high-priority preamble sequences and low-priority ones when scheduling their control messages derived from a preamble sequence detection (*e.g.*, RAR and CR messages). To cope with the shortage of radio resources and to improve QoS differentiation in the RA procedure, the concept of preamble-priority awareness was recently introduced in [4], which allows users employing high-priority preamble sequences to receive priority in the scheduling of their RAR control messages. By doing this, devices using prioritized preamble sequences can increase the chance of accessing the network during intense RA attempts [4].

In [4], the access success ratio and access delay of this new approach were assessed under low- and medium-loaded MTC scenarios. However, to fully understand the performance of the Preamble-Priority-Aware (PPA) concept, especially under heavily-loaded MTC scenarios, the relationship between RAR blocking ratio and the access success ratio is quite relevant. The reason is that high blocking of RAR messages was evinced under high loads in [3] and the PPA concept can have a positive effect on this metric for prioritized users under these conditions.

In this paper, we extend the application of the PPA concept to the scheduling of control messages others than RAR messages and analyze the impact of the introduction of this concept in the PDCCH scheduling when the traditional RA scheme and the RACH Resource Separation (RRS) scheme are used. We utilize the relationship between the access success ratio and the blocking of RAR messages to highlight its impact on the RA performance. We show that the PPA approach provides enhanced QoS support with both RA schemes, even under the challenging heavily-loaded, highly-synchronized MTC scenario defined by the 3GPP in [2], which was not previously shown.

The rest of the paper is organized as follows. Section II reviews the related work. Section III presents the PPA concept and its application to the scheduling of RAR and CR control messages. Section IV shows the performance of the PPA concept via extensive simulations and discusses the simulation results. Finally, we conclude the paper in Section V.

II. RELATED WORK

In this section, we review the literature related to RA schemes and PDCCH resource allocation and scheduling. We focus on RA schemes that prioritize a preamble or a set of preambles in the initial phase of the RA procedure. We divide the PDCCH-related work in papers that propose PDCCH policies either considering data-only or random-access-only signaling messages and those considering the whole picture of the PDCCH resource allocation process.

A. Random Access Schemes and their Preamble Priorities

Some RA schemes for LTE networks can provide prioritized access during the initial phase (preamble transmission) of the RA procedure. For instance, in the contention-free RA procedure, the Radio Access Network (RAN) explicitly informs the UEs the preamble sequence to be used in the RA procedure in order to avoid collisions. This preamble sequence is chosen from a poll of unique preamble sequences reserved for users in connected state requiring low latency access such as those performing handover or with pending downlink packets.

Another example is the RAN overload control scheme called RRS, which allows preamble sequence separation between HTC and MTC in order to alleviate the effect of the MTC on the HTC [2]. By reserving an exclusive set of preamble sequences for a small number of UEs, the collision probability in the preamble transmission phase of the RA procedure is significantly reduced in the presence of massive access attempts. Condoluci *et. al* [5] also introduced an RA scheme that reserves a set of preamble sequences for transmitting critical alarm messages.

Kim *et. al* [6] proposed the Prioritized Random Access (PRA) scheme, which enables fixed-location MTC devices to indicate their priority (either low or high) during the RA procedure by means of the transmit power level of the preamble sequence. In [7], the same authors have extended their idea to coexisting MTC/HTC scenarios, in which HTC users get high priority and MTC devices low priority. The PRA scheme includes a procedure to create first the msg2 messages for high-priority preambles and then those for low-priority preambles. In this way, the assembled msg2 messages to be scheduled contain the time advancing command of the high priority users, increasing the chances of successful transmission of msg3 messages from high priority users since they use the time advancing matching technique to reduce collisions in msg3 transmissions. However, the actual scheduling of PDCCH resources for this messages is neglected and the authors assume that unlimited PDCCH resources are available. Moreover, any PDCCH allocation framework nor PDCCH scheduler is described in their work. In addition, since this procedure is made in a TTI-basis, msg2 messages arriving in the n-th RAO are positioned in the queue after low priority msg2 messages from a previous RAO, which can impact the high priority users when PRACH configurations with various RAO per frame are used.

B. PDCCH Scheduling for Data Traffic

Villa *et. al* [8] propose one of the first existing models in the literature for the allocation of control resources in LTE and describe the interaction of the PDCCH resource allocation with the packet schedulers. Hosein [9] introduces another PDCCH resource allocation model and five scheduler algorithms. The above-mentioned papers focus on the allocation of control resources for uplink grants and downlink assignments and the schedulers allocate control resources to DCI messages based on the algorithms proposed. However, none of them includes the RA-related control messages in their model and the proposed algorithms do not consider these messages. Thus, these papers as well as several others do not explicitly consider allocation of radio resources for RAR and CR messages.

C. PDCCH Scheduling for Random Access

Cheng *et. al* [1] introduced the msg2-first PDCCH scheduling policy, in which RA-related control messages receive high priority, with msg2 having the highest priority. Osti *et. al* [10] propose an analytical evaluation framework for the the four-step RA procedure, considering limited PDCCH resources and employing the msg2-first policy. Lin *et. al* [11] propose a throughput-optimal policy in which an objective function is optimized combining scheduling of msg2 and msg4 messages to increase the PRACH throughput. They compare the performance of the throughput-optimal policy to that of the msg2-first one. All these papers, however, do not specify the scheduling for downlink assignments, uplink grants, BSR grants, TPC commands, among others.

D. PDCCH Scheduling for Data Traffic and Random Access

de Andrade *et. al* [12] introduced a PDCCH resource allocation model that includes RA-related control messages and propose three PDCCH schedulers: the RA-Priorized (RAP), Lifetime-Aware (LTA) and GBR-Priorized LTA (GBR-LTA) algorithms. The RAP algorithm prioritizes RA-related messages over other control messages whereas the last two provide QoS differentiation by taking into account both the specific waiting timer of the various control messages involved in the RA procedure and the delay requirements of the downlink and uplink users' traffic.

Astudillo *et. al* [4] proposed the PPA algorithm, which is based on the RAP algorithm and allows users employing high-priority preamble sequences to receive priority in the scheduling of their msg2 control messages.

III. ENHANCED PREAMBLE-PRIORITY-AWARE PDCCH SCHEDULING ALGORITHM

In this section, we describe the Enhanced Preamble-Priority-Aware (ePPA) PDCCH scheduling policy, an extension of the PPA algorithm that prioritizes not only RAR messages but also CR control messages based on the priority of the preamble sequences that generate them. The preamble sequence sets and their corresponding priority levels are based on the RA scheme and the RAN overload control scheme defined by the Mobile Network Operator (MNO) at the eNB.

The ePPA algorithm uses the main idea introduced by the msg2-first policy and also used in the RAP and the PPA algorithms, in which msg2 messages get high priority over msg4 messages. The ePPA algorithm sorts the msg2 and msg4 queues based on the priority of the preamble sequences that generated these messages. The ePPA policy defines at least two priority levels for scheduling msg2 and msg4 messages. The highest priority is given to messages addressed to users employing contention-free preamble sequences since this preamble set is typically used by users requiring low access latency such as those performing handover or with pending downlink transmissions. The other priority levels depend on the priority given to the contention-based preamble sequences by the RA scheme and the RAN overload control scheme being used. When the traditional RA scheme is used (without using any RAN overload control scheme or prioritized RA scheme), all contention-based preamble sequences are mapped onto the low-priority level since this scheme does not provide any differentiation among them. Medium- and low-priority levels are defined by the ePPA algorithm for those RA schemes that differentiate among contention-based preamble sequences such as the RRS scheme, the PRA scheme and Condoluci *et al*'s proposal. The high-priority set of contention-based preamble sequences is mapped onto a medium-priority level, whereas the low-priority set gets the lowest priority level. More priority levels can exist depending on the RA scheme used. Once all msg2 and msg4 messages are scheduled, the other control messages are processed as in the RAP algorithm [12].

IV. PERFORMANCE EVALUATION

In this section, we analyze the impact of the PPA concept via extensive simulations by using the LTE Simulator (LTE-Sim) [13], which is a widely-used LTE simulator developed in C++. To do this, we compare the RA performance of the ePPA algorithm to that of the RAP algorithm under PDCCH constraints in terms of the *access success ratio*, which is defined as the ratio of devices that successful perform the RA procedure and the *msg*2 *blocking ratio*, which is the number of dropped msg2 messages divided by the total number of msg2 messages that joined the msg2 queue at the eNB. To assess the effect of PDCCH constraints with different RA schemes, we implemented the traditional LTE RA scheme with both contention-free and contention-based operation modes as well as the RRS RAN overload control scheme, which was proposed to alleviate the negative effect of massive MTC on HTC. Since both scheduling algorithms use the principle of the msg2-first policy, actual data transmission is not necessary to assess their RA performance.

A. Simulation Model

Both HTC users and MTC devices are considered in the scenarios. Activation of MTC devices follows the $Beta(3, 4)$ distribution within a 10s interval to simulate a extreme scenario with MTC transmissions highly synchronized as proposed by the Third Generation Partnership Project (3GPP) in [2]. Once an MTC device is activated, it triggers the RA procedure in order to transmit its uplink data since allocation of Physical Uplink Control Channel (PUCCH) resources to several MTC devices is not feasible [14]. HTC users generating Voice-over-IP (VoIP) traffic, modeled as explained in [13], are also included. To decrease the amount of PDCCH resources used by the VoIP users and to accommodate more VoIP calls per cell [14], semi-persistent scheduling with initial random access [15] is assumed to be used by VoIP users. This means that VoIP users initiate the RA procedure to send Scheduling Request (SR) messages at the beginning of each talkspurt. The arrival of VoIP users (initial activation) follows a Poisson distribution with rate λ_{VoIP} within a 10s interval, *i.e.*, the inter-arrival time of VoIP users is modeled as a negative exponential distribution with mean μ_{VoIP}, where $\mu_{VoIP} = \frac{1}{\lambda_{VoIP}}$. Some HTC users are assumed to perform handover by using the contention-free RA procedure, following a Poisson process with rate λ_{HO} [16]. Other HTC users are assumed to have PUCCH resources allocated to send SR messages so that they do not frequently use the RA procedure to request uplink resources. As we use the enhanced LTE-Sim module described in [3], which implements the RA procedure and different RA schemes, implementation assumptions made in [3] hold here. Note that the PDCCH allocation model used is the one proposed in [12].

B. Simulation Setup

The simulation scenarios comprise a single cell with a 0.5 km radius. An eNB with 5 MHz cell bandwidth in the frequency division duplexing mode is located at the center of

TABLE I. SIMULATION PARAMETERS

Parameter	Value
System type	Single cell
System bandwidth	5 MHz
Cell radius	0.5 km
PRACH configuration index	6
RA preamble format	0
Contention-free preambles	12
Contention-based preambles	52
RRS preamble sets	30 for MTC devices 22 for HTC users
RAR messages per TTI	3
CCEs allocated for the PDCCH	16
CCEs per UE-specific DCI message	4
CCEs per common DCI message	8
Backoff	20 ms
HARQ retransmission probability	10%
preambleTransMax	10
Response Window Size	5 ms
Contention Resolution Timer	48 ms
maxHARQ-Msg3Tx	5
λ_{HO}	2.7 HOs/s
λ_{VoIP}	12 UEs/s

the cell with several UEs uniformly distributed around it. Each UE acts either as MTC device or as HTC user. Following the requirements of the LTE-Advanced (LTE-A) technology, the number of HTC users attached to the cell was fixed to 300 UEs. We assumed that 40 percent of these users (120 UEs) have VoIP traffic. Therefore, $\lambda_{VoIP} = \frac{120 \text{ UEs}}{10s} = 12$ UEs/s. Based on the total number of Radio Resource Control (RRC)-connected devices, the mean number of X2-based (inter-eNB) handovers per second is 1.8 [17]. To account for intra-eNB handovers also, we add 50 percent to the calculated value of X2-based handovers. Therefore, $\lambda_{HO} = 2.7$ HOs/s. Scenarios with $5,000$ (*light load*), $10,000$ (*medium load*), and $30,000$ (*high load*) MTC devices were executed as proposed by the 3GPP in [2]. In the traditional RA scheme, a set of 12 preamble sequences are used by handover users for contention-free RA and the remaining 52 preamble sequences are shared between MTC devices and HTC users for contention-based RA. In the RRS scheme, the contention-based preamble sequences are further separated between MTC devices (30 preambles) and HTC users (22 preambles). Table I summarizes the main configuration parameters used in the simulations.

C. Simulation Results and Discussion

The figures in this section show mean values derived by using the independent replication method with 10 replication. Confidence intervals of 95 % confidence level are also shown. The scenarios were run until all the started RA procedures were finished (either successfully or unsuccessfully).

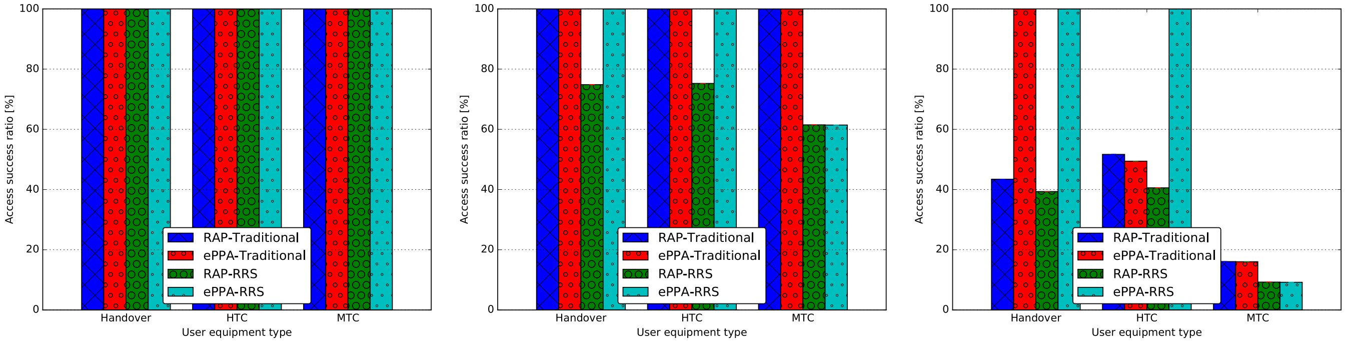
The access success ratio is shown in Fig. 1. Under *light loads*, none of the evaluated scenarios has blocking of access requests (Fig. 1(a)) since there are sufficient resources for the demand of all type of users (Fig. 2(a)). Under *medium loads*, none of the traffic types presented blocking of access requests when the traditional RA scheme is employed (Fig. 1(b)). However, under *high loads*, the ePPA algorithm with the traditional RA scheme yields no loss of access attempts for handovers users but does not provide QoS guarantees for HTC users (Fig. 1(c)). This is explained by the fact that handover users utilize contention-free RA procedure, whereas HTC users share the contention-based preambles with MTC devices in the traditional RA scheme. Thus, the ePPA algorithm gives high priority to control messages for handover users but does not prioritize control messages for HTC users over those for MTC devices.

On the other hand, a portion of the users utilizing prioritized preamble sequences is not able to get access to the network when the RAP algorithm with RRS scheme is employed under higher loads, *i.e.*, $10,000$ (Fig. 1(b)) and $30,000$ MTC devices (Fig. 1(c)). This fact shows the lack of the RAP algorithm to prioritize control messages derived from prioritized preambles. Due to the reduction of the number of preambles reserved for MTC devices in the RRS scheme, high number of collisions of preamble transmissions from MTC device occurs [3]. The msg2-related resource utilization increases with the collision probability since unsuccessful msg3 transmissions implies on new preamble transmissions and more messages arriving at the msg2 queue. As the resources for control messages is shared between MTC devices and HTC users and the RAP algorithm does not differentiate msg2 messages for any type of users, high msg2 blocking ratio is evinced for all type of users under *medium* (Fig. 2(b)) and *high loads* (Fig. 2(c)). Conversely, the preamble-priority awareness approach of the ePPA algorithm leads to no loss of access attempts for users with high and medium priority when the RRS scheme is used. The ePPA algorithm is thus able to provide QoS differentiation during the RA procedure even under massive access attempts.

Even though the RRS scheme has been considered a good option to isolate the impact of MTC on HTC [2], existing literature on RAN overload control schemes has not considered the constraints imposed by PDCCH resource allocation and scheduling on the performance. The results presented here, which take into account these constraints, evince that the RRS scheme only achieves its objective with the help of the PPA approach. Thus the ePPA PDCCH scheduler is able to guarantee the required isolation in scenarios where RACH resource separation is employed such as coexisting MTC/HTC scenarios or MTC scenarios with high and low priority MTC devices. It can also be interesting for RAN virtualization in which resource isolation is one of the key requirements.

V. CONCLUSION

In this paper, we have analyzed the impact of the Preamble-Priority-Aware (PPA) concept in the PDCCH scheduling of LTE/LTE-A networks. Specifically, we apply this concept to the scheduling of RAR and CR control signaling messages. We have highlighted the reduction of the msg2 blocking ratio for prioritized users when the PPA concept is employed. It increases the access success probability for users whose preambles sequences are prioritized by the traditional RA scheme and the RRS scheme in the initial phase of the RA procedure. In fact, we found that this concept is fundamental to achieve the goal of isolation between high-priority and low-priority users in the RRS RA scheme. Moreover, we have shown that the PPA concept provides QoS differentiation to users utilizing contention-free preambles.

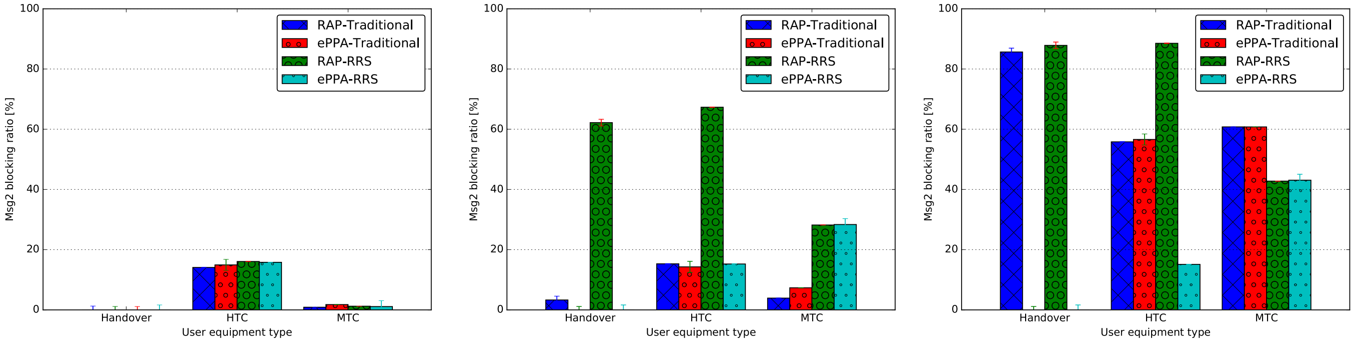


(a) 5,000 MTC devices

(b) 10,000 MTC devices

(c) 30,000 MTC devices

Figure 1. Access success ratio



(a) 5,000 MTC devices

(b) 10,000 MTC devices

(c) 30,000 MTC devices

Figure 2. Msg2 blocking ratio

ACKNOWLEDGEMENTS

This work was supported by Motorola and the Brazilian National Council for Scientific and Technological Development (CNPq).

REFERENCES

[1] M.-Y. Cheng *et al.*, "Overload control for machine-type-communications in LTE-advanced system," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 38–45, June 2012.

[2] 3GPP, "Technical Specification Group Radio Access Network; Study on RAN Improvements for Machine-type Communications," 3rd Generation Partnership Project (3GPP), TR 37.868, September 2011.

[3] T. P. C. de Andrade *et al.*, "The random access procedure in long term evolution networks for the Internet of Things," *IEEE Commun. Mag.*, vol. 55, no. 3, Mar. 2017.

[4] C. A. Astudillo, T. P. C. de Andrade, and N. L. S. da Fonseca, "Allocation of control resources with preamble priority awareness for human and machine type communications in LTE-Advanced networks," in *2017 IEEE Int. Conf. Commun. (ICC)*, May, pp. 1–6.

[5] M. Condoluci *et al.*, "Enhanced radio access and data transmission procedures facilitating industry-compliant machine-type communications over LTE-based 5G networks," *IEEE Wireless Commun.*, vol. 23, no. 1, pp. 56–63, Feb. 2016.

[6] T. Kim, K. S. Ko, and D. K. Sung, "Prioritized random access for machine-to-machine communications in OFDMA based systems," in *2015 IEEE Int. Conf. Commun. (ICC)*, June, pp. 2967–2972.

[7] ——, "Prioritized random access for accommodating M2M and H2H communications in cellular networks," in *2015 IEEE Globecom Workshops (GC Wkshps)*, December, pp. 1–6.

[8] D. L. Villa *et al.*, "Performance of downlink UTRAN LTE under control channel constraints," in *IEEE 67th Veh. Technol. Conf. (VTC Spring)*, May 2008, pp. 2512–2516.

[9] P. Hosein, "Resource allocation for the LTE physical downlink control channel," in *2009 IEEE Globecom Workshops*, November, pp. 1–5.

[10] P. Osti *et al.*, "Analysis of PDCCH performance for M2M traffic in LTE," *IEEE Trans. Veh. Technol.*, vol. 63, no. 9, pp. 4357–4371, Nov. 2014.

[11] G. Y. Lin, S. R. Chang, and H. Y. Wei, "Estimation and adaptation for bursty LTE random access," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2560–2577, Apr. 2016.

[12] T. P. C. de Andrade, C. A. Astudillo, and N. L. S. da Fonseca, "Allocation of control resources for machine-to-machine and human-to-human communications over LTE/LTE-A networks," *IEEE Internet of Things J.*, vol. 3, no. 3, pp. 366–377, June 2016.

[13] G. Piro *et al.*, "Simulating LTE cellular systems: An open-source framework," *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, pp. 498–513, Feb. 2011.

[14] T. P. C. de Andrade, C. A. Astudillo, and N. L. S. da Fonseca, "Impact of M2M traffic on human-type communication users on the LTE uplink channel," in *7th IEEE Latin-American Conf. Commun. (LATINCOM)*, November 2015, pp. 1–6.

[15] J. B. Seo and V. C. M. Leung, "Performance modeling and stability of semi-persistent scheduling with initial random access in LTE," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4446–4456, Dec. 2012.

[16] D. Hong and S. S. Rappaport, "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures," *IEEE Trans. Veh. Technol.*, vol. 35, no. 3, pp. 77–92, Aug. 1986.

[17] R. Lilius *et al.*, *Planning and Optimizing Mobile Backhaul for LTE*. John Wiley and Sons, Ltd, 2015, pp. 129–237.