Goodbye, ALOHA!

A. Laya*, C. Kalalas‡, F. Vazquez-Gallego‡, L. Alonso† and J. Alonso-Zarate‡
*KTH Royal Institute of Technology, Sweden. e-mail: laya@kth.se
†Universitat Politècnica de Catalunya (UPC), Spain. e-mail: luisg@tsc.upc.edu
‡Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Spain. e-mail: {ckalalas, francisco.vazquez, jesus.alonso}@cttc.es

Abstract

The vision of the Internet of Things (IoT) to interconnect and Internet-connect everyday objects and machines poses new challenges in the design of wireless communication networks. The design of Medium Access Control (MAC) protocols has been traditionally an intense area of research due to their high impact on the overall performance of wireless communications. The majority of research activities in this field deal with different variations of protocols somehow based on ALOHA, either with or without listen before talk, i.e., Carrier Sensing Multiple Access (CSMA). These protocols operate well under low traffic loads and low number of simultaneous devices. However, they suffer from congestion as the traffic load and the number of devices increase. For this reason, unless revisited, the MAC layer can become a bottleneck for the success of the IoT. In this paper, we provide an overview of the existing MAC solutions for the IoT, describing current limitations and envisioned challenges for the near future. Motivated by those, we identify a family of simple algorithms based on Distributed Queueing (DQ) which can operate for an infinite number of devices generating any traffic load and pattern. A description of the DQ mechanism is provided and most relevant existing studies of DQ applied in different scenarios are described in this paper. In addition, we provide a novel performance evaluation of DQ when applied for the IoT. Finally, a description of the very first demo of DQ for its use in the IoT is also included in the paper.

I. Introduction

The Internet of Things (IoT) has the potential to transform the world as we know it. The IoT entails the vision of improving industries and society by enabling the automated remote communication between objects and machines and the smart use of the exchanged data. IoT is about automating and enhancing processes to reduce expenditures and create novel services. The IoT can bring benefits to several verticals sectors, enabling concepts such as remote health care, autonomous driving, intelligent transport systems, smart-homes, smart-grids, and industry 4.0, just to mention a few.

Many challenges have to be addressed to accomplish the full potential of the IoT. This paper focuses on one of the key topics that need to be addressed; the need to enable efficient of Machine-Type Communications (MTC). For many years, wireless communication networks have been designed for Human-Type Communications (HTC) and not for MTC. However, MTC are fundamentally different from HTC. MTC are characterized by a heterogeneous variety of requirements covering both delay-tolerant to delay-critical applications, all mixed up. MTC bring new data traffic patterns; combining short and bursty traffic with periodic reporting messages. Typically, MTC is associated with a massive number of simultaneously connected devices, orders of magnitude above what current communication networks are capable of dealing with.

One of the key building blocks of a wireless communication network can be found at layer 2 of the protocol stack. The Medium Access Control (MAC) layer is responsible for deciding who, when, and how access to the shared wireless channel is granted. Among other existing options, Random Access (RA) methods have received increasing attention from the research community. RA methods share the communication channel using some kind of randomization procedures and distributed access. The great majority of existing contributions are based on variations of ALOHA, and its variation with carrier sensing, i.e., Carrier Sensing Multiple Access (CSMA). Since ALOHA was designed[1], variations of it have been used in almost all telecommunication systems, e.g., cellular systems, Wireless Local Area Networks (WLANs), Radio Frequency Identification (RFID), Bluetooth, etc. This is summarized in Table I. MAC protocols based on either ALOHA or CSMA operate very well when the number of simultaneous contending users is low and the overall traffic load is low. However, they suffer from congestion as the traffic load and number of devices increase. The challenge consists on how to efficiently handle the connectivity from massive number of devices or a number of devices which request very frequent channel accesses to transmit small data packets; even when these channel accesses may be concentrated over short periods of time, e.g. event-driven applications. One solution to this could consist in the deployment of denser access networks, i.e., using many small cells or access points to create dense deployments and reduce contention in each network cell. However, this approach may not constitute a cost-effective solution given the capacity requirements of the majority of IoT applications.

A possible solution can be found in the family of protocols based on Distributed Queueing (DQ). DQ protocols have been already studied in various wireless network use cases showing that they can:

[1]Aloha is a Hawaiian word used as an English greeting to say goodbye and hello. ALOHA is also the name of a pioneering computer network system developed at the University of Hawaii in the early 70's, effectively providing the first public demonstration of a wireless packet data network. By the naming of this manuscript, we invite the research community to welcome new approaches to address some of the challenges imposed by the Internet of Things, while recognizing the tremendous achievement that legacy technologies continue to provide to our ever evolving endeavour as researchers, scientists, and engineers.

TABLE I
OVERVIEW OF CONTENTION-BASED CHANNEL ACCESS MECHANISMS IN VARIOUS COMMUNICATION TECHNOLOGIES FOR IOT

Contention-based access mechanism	Technology
Pure ALOHA	SigFox, LoRa
Slotted ALOHA	RFID, RACH of LTE, NB-IoT (CIoT), Weightless
Non-slotted CSMA/CA	ZigBee, WiFi
Slotted CSMA/CA	ZigBee

- Attain the maximum capacity of the channel (attaining a near-optimum performance).
- Share the available resources in a fair manner, while accepting the enforcement of Quality of Service (QoS) policies.
- Ensure maximum performance independently of the number of contenting devices and traffic pattern.
- Ensure maximum performance without having *a priori* knowledge of the configuration and/or composition of the network; such flexibility is an invaluable asset for the IoT.

For all these reasons, we present DQ as a MAC protocol highly suitable for future networks that will need to provide communication capabilities for both HTC and the MTC.

This paper has a twofold contribution:

1) First, it provides a comprehensive discussion about the use of ALOHA and CSMA (and their variations) in communication systems that are becoming predominant in IoT deployments, i.e., Wireless Personal Area Networks, WLAN, public LTE, the recently introduced narrow-band IoT-tailored radio networks (NB-IoT), Sigfox, LoRa and Weightless.
2) Second, it presents DQ and its suitability to deal with a high density of devices in IoT applications. A detailed review of existing literature related to DQ is provided. A simulation experiment of the use of DQ in LTE for massive MTC is presented and a demo of a Machine-to-Machine (M2M) area network based on an implementation of DQ is described.

Key comparative studies and surveys are referenced in order to guide interested readers into detailed comparative studies of ALOHA and DQ in various technologies. In this paper, a comparison is provided between DQ and current cellular technologies, based on the suggestion that cellular technologies are emerging as strong candidates to leverage the potential of the IoT. The remainder of the paper is organized as follows: Section II describes and compares the MAC protocols used in current telecommunication networks that are relevant for the IoT. This section aims at demonstrating that contention-based access for the majority of existing technologies still relies on simple variants of ALOHA and CSMA. In Section III, the limitations of ALOHA are identified and the need for a new understanding of these protocols for the IoT is described. Section IV is devoted to describe in detail the DQ concept and relevant studies evaluating its benefits for different communication networks. In Section V, the suitability of DQ for the deployment of the IoT is presented in two parts; In Section V-A, a computer-based simulation is described to compare the performance of the RA procedure of LTE and that when using an adaptation of DQ. In Section V-B, a demonstration of an M2M area network based on DQ is presented. Finally, the paper is concluded in Section VI. Since this paper deals with many technologies and protocols, the acronyms included in this paper have been summarized in Table V in the Appendix.

II. MAC IN EXISTING IOT COMMUNICATION TECHNOLOGIES

Various communication technologies have been considered to support emerging IoT applications. Their MAC implementations mainly rely on hybrid schemes that employ both contention- and schedule-based access mechanisms, in an effort to leverage the advantages of both approaches in terms of complexity and performance. Spread spectrum techniques are also applied to provide multiple access capabilities in the frequency domain. In particular, multiple users are able to access simultaneously the same frequency band, while frequency diversity is achieved through different pseudo-random number sequences, e.g., spreading codes, or frequency-hopping patterns.

Despite the vast amount of existing studies on MAC protocols, only variations of ALOHA and CSMA are still used in the great majority of technologies being used for the IoT. In the following sections, the MAC of different technologies is reviewed. A summary is also provided in Table II.

A. Wireless Personal Area Networks

1) ZigBee: The IEEE 802.15.4 Standard, promoted by the IEEE 802.15 working group, constitutes the basis of the ZigBee Alliance specification. This standard defines the PHY and MAC layers for low data rates and low power ad hoc self-organizing networks of inexpensive fixed, portable and moving devices [1]. It operates in license-free bands and specifies two different channel access methods:

- Beacon-enabled mode for star-topology networks: a hybrid-based MAC using a slotted Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) scheme for delay-tolerant data and an optional Guaranteed Time Slot (GTS) allocation mechanism with contention-free reserved access for time-critical data.

- Non-beacon mode suitable for multi-hop deployments: a contention-based MAC using a simple non-slotted CSMA/CA mechanism based on channel sensing and random exponential backoff for contention resolution.

In a star-topology beacon-enabled network, the communication between a network coordinator and the nodes is performed during the access period which is defined by periodic beacon broadcasts. In particular, when a device needs to send data to a coordinator, it must wait for beacon synchronization and later contend for channel access. The access period is divided into a Contention Access Period (CAP) where a slotted CSMA/CA mechanism is used for channel access of delay-tolerant data and GTS requests, and an optional Contention Free Period (CFP), composed of GTSs that are assigned and administered by the network coordinator. In the CFP, the dedicated bandwidth is reserved for time-critical data. In the case a coordinator needs to communicate with a network device, it indicates the pending data in the beacon; in turn, devices periodically wake up and listen for beacon frames to identify possible data reception.

In the non-beacon mode, transmission is based on channel sensing and nodes apply a random exponential backoff mechanism for contention resolution. Each time a device wants to transmit data frames or MAC layer commands, it waits for a random time period. By the expiration of this period, if the channel is found inactive, the device sends its data; otherwise, if the channel is busy, the device waits for a random time period until it checks again the availability of the channel. Despite most of the unique features of IEEE 802.15.4 are in the beacon-enabled mode, combining the advantages of both contention-based and scheduled-based MAC, the great majority of implementations today use the non-beacon mode.

Due to the poor performance of this technology in networks with a high number of simultaneous devices, several research works have aimed to tune the IEEE 802.15.4 MAC layer operation–either by making use of PHY layer measurements or link layer information– to improve the performance in terms of reliability, delay, or throughput [2], [3]. Basic techniques include optimizations of the average backoff window size and dynamic algorithms to set the contention window size. The IEEE 802.15.4e constitutes a recent MAC amendment which adopts a time-slotted channel hopping strategy to enhance low-power operation and reliability by increasing robustness against interference and multi-path fading. This has been referred to as Time Synchronized Channel Hopping (TSCH) and is suitable for static industrial deployments. In RSCH, subsequent packets are sent using different frequencies following a pseudo-random hopping pattern, improving the successful transmission rates [4].

2) Bluetooth Low Energy: Bluetooth Low Energy (BLE) is a short-range wireless technology developed to enable a potentially large number of devices in IoT applications. BLE is gaining momentum in several control and monitoring applications [5]. Based on the IEEE 802.15.1 Bluetooth Standard, BLE defines a lightweight MAC layer that offers ultra-low power idle mode operation, simplified device discovery, and supports increased number of nodes [6]. BLE relies on a time slotted access mechanism with a time division multiplexing technique applied to coordinate the medium access. Each channel is divided into time slots to avoid packet collisions and an adaptive frequency hopping spread spectrum method is used in the ISM license-free frequency band to mitigate interference and multi-path fading.

In particular, BLE defines a master/slave network architecture, named *piconet*, where a master node manages numerous connections with multiple slave nodes and each slave node is associated with only one master. Slave nodes are by default in sleep mode and wake up periodically to listen to possible packets transmitted from the master. In turn, the master regulates the medium access using a Time Division Multiple Access (TDMA) scheme to assign the time slots when the slaves need to listen to. Upon connection establishment, the master provides information to the slave node for the selected channel frequency and the timing for the data exchange. Channel selection relies on a robust frequency hopping mechanism while knowledge about the connection duration allows for an optimization in the power consumption.

3) Radio Frequency Identification: Radio Frequency Identification (RFID) constitutes an important enabler for IoT applications such as asset tracking and remote supervision. RFID systems operate in ISM frequency bands and use radio signal broadcast to automatically identify items with attached RFID tags. Contention-based channel access for RFID mainly relies on uncoordinated Frame Slotted ALOHA (FSA) schemes. In an effort to mitigate tag collision problem, various proposals aim at the design of collision resolution techniques for the performance improvement of FSA in RFID systems [7]. The first approach refers to the dynamic adaptation of the number of slots per frame based on an estimate of the tag population derived from collisions, e.g., double the number of slots per frame if the number of collisions is high. The second anti-collision mechanism builds a query tree based on subsequently querying a sub-group of tags, e.g., first discover the tags and then query each tag independently to avoid collisions. However, both approaches are not optimal in terms of system performance and low energy consumption due to the time and energy required to estimate the number of tags from collisions or to build the query tree.

B. Wireless Local Area Networks: WiFi

The IEEE 802.11 family of standards, supported by the WiFi alliance, consists of a number of specifications that primarily define the PHY and MAC layers for WLANs [8]. WiFi is a mature and widely adopted wireless technology. In addition and now it is a promising candidate to support a diverse range of IoT applications, thanks to the recent low-power specifications in 802.11ah. Although mainly used for Internet access at residential premises, WiFi is increasingly getting deployed for other use cases as well, spanning from industrial automation, e.g. smart grids, to intelligent commercial buildings [9].

The Distributed Coordination Function (DCF) constitutes the fundamental MAC technique of the IEEE 802.11 Standard [10]. DCF is based on a CSMA/CA scheme with a slotted binary exponential backoff (BEB) mechanism for retransmissions in case of collision. Besides the basic access scheme that relies on acknowledgements, DCF provides an optional virtual carrier sensing mechanism based on the exchange of short Request-to-Send (RTS) and Clear-to-Send (CTS) control frames between source and destination nodes to reduce collisions introduced by the hidden node problem. The IEEE 802.11e MAC amendment introduces an Enhanced Distributed Channel Access (EDCA) function which defines multiple access categories and relevant configuration parameters to support MAC-level QoS provision and prioritization [11]. This technique has been also used in subsequent amendments for high throughput, i.e., 802.11ac and 802.11ad, to ensure some degree of soft-QoS guarantees.

Due to the limitations of the IEEE 802.11 in case of increasing demand for ubiquitous wireless access, the IEEE 802.11ah amendment has been recently proposed to support large-scale topologies with increased (over 8000) number of nodes associated with an Access Point (AP) via a hierarchical identifier structure [12]. In particular, three types of stations are introduced, each of them associated with different channel access mechanisms: Traffic Indication Map (TIM) stations, non-TIM stations, and unscheduled stations. For TIM stations, on top of contention-based access, the IEEE 802.11ah introduces a beacon-enabled access method with time slot reservations, named Restricted Access Window (RAW). RAW constitutes a time period among signalling beacons and consists of one or multiple time slots. The AP is responsible for assigning each time slot to a group of TIM stations and broadcasts this information within the beacon frames. In turn, the TIM stations, upon receiving the RAW information, identify whether they are allowed to contend for medium access in a time slot or not. This technique ensures a fair spectrum access among a large number of nodes, reduces the number of simultaneous access attempts and maximizes the channel utilization. On the other hand, data transmissions for non-TIM stations are scheduled during a Periodic RAW (PRAW), where access for TIM stations is prohibited. Similarly, unscheduled stations do not require any beacon listening prior to transmission and the AP allocates time slots outside both restricted windows for their sporadic channel access.

C. Public Cellular Networks and CIoT

3GPP standardization efforts aim at enabling LTE as a suitable connectivity technology for the IoT in the mid-term future, particularly for the case of massive MTC. The ubiquitous infrastructure provides benefits in terms of coverage, support for mobility, and use of licensed bands. However, in LTE technology, User Equipments (UEs) use the Random Access CHannel (RACH) to perform initial network association, request transmission resources, and re-establish a connection to the eNodeB (base station). The RACH is formed by a periodic sequence of allocated time-frequency resources, reserved in the uplink channel for the transmission of access requests. The RA procedure in LTE can be either contention-free or contention-based. In the contention-free mode, the eNodeB allocates specific access resources for requests that require high probability of success (delay-constrained access), e.g., handover [13]. On the other hand, the contention-based RA operation normally involves a four message handshake between the UE and the eNodeB and is based on (multi-channel) Frame Slotted ALOHA (FSA) medium access, i.e., mutually orthogonal preambles are used by the UEs to contend in the available RA slots. In the case of the transmission of simultaneous access requests, this may result in a severe performance degradation due to a high probability of collision in the transmission of the preambles. To this end, several methods have been proposed during the recent years to improve the contention-based RACH operation, including MAC-parameter optimizations, access class barring schemes and separation of RA resources [14].

Due to the limitations of LTE to deal with huge numbers of devices and provide the IoT with cost-efficient and energy-efficient communications, the 3GPP is approaching the suitability of releases of LTE for massive MTC with the inclusion of a new UE category-0 associated with PHY layer capabilities specifically intended for MTC support (LTE-M) [15]. LTE-M, also referred to as LTE for MTC, constitutes an evolution enabling the IoT by reducing the complexity, cost, and power requirements of the end devices. However, the access to the system remains the same, based on a FSA scheme.

In an effort to further address the heterogeneous IoT communication needs, a novel narrow-band radio access technology is being promoted within 3GPP, coined Narrow Band IoT (NB-IoT)–and formerly referred to as Cellular IoT (CIoT)–which promises to reach those vertical market applications where LTE-M cannot reach [16]. Aiming at a lower device cost and power consumption, and support of a massive number of low throughput devices, NB-IoT technology takes into account the received signal strength information for an efficient management of RACH resources. In particular, depending on the coverage conditions of each UE, a different set of RACH resources is specified while the parameters for the random access procedure can be network-configured for different coverage classes. To handle collision on the RACH, NB-IoT makes use of overlaid Code Division Multiple Access (CDMA). Orthogonal codes are used to separate users within a coverage class that attempt simultaneous system access.

D. Unlicensed Low Power Wide Area Networks

Emerging Low Power Wide Area Network (LPWAN) technologies are gaining attention as suitable solutions for IoT wireless connectivity [17]. They are becoming complementary or alternative approaches to fill the gap between local wireless and mobile wide area network technologies and address their shortcomings in IoT applications. Compared to 3GPP Standards, unlicensed spectrum is now utilized, which could make QoS requirements difficult to guarantee; however, the dedicated devices are often more power and cost efficient than 3GPP ones. In the following, we focus on the three most widely-deployed solutions, namely Sigfox, LoRa and Weightless.

TABLE II
OVERVIEW OF MAC IMPLEMENTATIONS AND CHANNEL ACCESS METHODS IN VARIOUS COMMUNICATION TECHNOLOGIES FOR IOT

	Channel access mechanisms	**Technology**								
		ZigBee	BLE	RFID	WiFi	LTE	NB-IoT[a](CIoT)	Sigfox	LoRa	Weightless
Contention-based	Pure ALOHA							✓	✓	
	Slotted ALOHA			✓		✓	✓			✓
	Non-slotted CSMA/CA	✓			✓					
	Slotted CSMA/CA	✓								
Schedule-based	Frequency Division Multiple Access				✓	✓	✓			✓
	Time Division Multiple Access		✓							✓
	Code Division Multiple Access						✓		✓	
	Time slot reservation	✓			✓	✓			✓	
Spread spectrum	Frequency Hopping Spread Spectrum		✓							✓
	Direct Sequence Spread Spectrum	✓			✓					✓
	Chirp Spread Spectrum	✓							✓	

[a] NB-IoT (CIoT) is currently under standardisation by 3GPP RAN Working Group [16].

1) Sigfox: Sigfox technology adopts an ultra-narrow band implementation, using sub-GHz frequency bands to enable long-range communication for IoT applications with very low data rates (100bps) [18]. Due to the narrow-band operation, Sigfox deployments allow large-scale network topologies with improved sensitivity values since the receiving noise impact is reduced. Concerning medium access, Sigfox does not employ any collision-avoidance mechanisms for medium access; instead, a Random FTDMA (R-FTDMA) scheme is applied, where each node asynchronously transmits at a frequency chosen randomly in the continuous available frequency band. Therefore, this is indeed an ALOHA-based procedure. Time and frequency randomness render this scheme prone to high interference. To cope with this problem, software-defined radio techniques are designed to ensure the overall adequate performance.

2) LoRa: Long Range (LoRa) technology focuses on a wideband CDMA approach and specifies the lower-layer physical layer functionality. LoRaWAN is the LoRa Alliance protocol handling the higher-layers of the communication protocol and the network system architecture [19]. LoRaWAN employs a lightweight MAC layer and defines three different classes of end-point devices to address the different requirements reflected in the wide range of applications. Bidirectional communication for class-A devices is performed using pure-ALOHA and is suitable for applications that require a downlink server response shortly after the uplink transmission. However, the achieved throughput performance is relatively low since this ALOHA-based scheme is highly susceptible to packet collisions. Class-B devices use an additional beacon-enabled time-slotted communication scheme that allows for scheduled message reception windows. A less energy-efficient scheme is considered for class-C devices which use a similar to class-A uncoordinated transmission scheme and remain always-on listening to the medium for message reception. Spectrum spreading in LoRa technology is achieved by generating a chirp signal that continuously varies in frequency and lowers the complexity of the receiver design.

3) Weightless: Weightless technology constitutes another alternative LPWAN technology designed to provide relatively low-cost MTC utilizing low-frequency spectrum and techniques that enable communications over a long range [20], [21]. Weightless systems employ a master-slave architectural model and each MAC frame consists of a downlink part followed by an uplink one. In particular, based on resource allocation information included in the downlink section, the devices (slaves) are able to transmit on the determined frequency during uplink opportunities allocated by the base station (master). Depending on the regulatory environment, two uplink multiple access modes are specified, narrowband FDMA and wideband FDMA, which constitute a combination of FDMA and TDMA schemes. In the case of initial network association or unscheduled message transmissions, contention-based channel access takes place using a variation of FSA scheme. Weightless specification employs various mechanisms to reduce the increased number of collisions. These techniques include dynamic configuration of contention-based access slots and device prioritization for access restriction to certain classes.

Weightless Special Interest Group defines three connectivity standards targeting at different use cases: Weightless-W, Weightless-N, and Weightless-P. Weightless-W, designed to operate in white space spectrum, is based on a time division duplexing operation. It uses a direct-sequence spread spectrum technique with variable spreading factors to minimize the interference. Weightless-N specification, typically deployed over the ISM bands, is designed for low power and low cost devices that perform one-way communication. It uses an ultra-narrow band technology and a frequency hopping algorithm is applied for interference/fading mitigation and enhanced security. On the other hand, Weightless-P standard allows for a bidirectional communication and applies a combined FDMA and TDMA scheme for access in 12.5kHz narrow-band channels.

III. Motivation to depart from ALOHA

The ALOHA and CSMA protocols, together with all their variants, have been comprehensively analysed in the available literature. Moreover, there are many solutions based on ALOHA that are applicable for IoT regarding the random access [22], [23], [24], [25]. Most of the existing theoretical analyses consider homogeneous networks where each device generates packets following a given random distribution [10]. They evaluate the performance of the protocols in terms of delay and throughput in steady-state conditions. Due to its mathematical tractability, stationary Poisson processes have been traditionally used to model traffic generation. However, new applications and new communication scenarios, particularly posed by the IoT, require a revision of existing models, including traffic generation models and their impact into communication protocols. Some examples of applications are:

- Structural health monitoring, where a large number of wireless sensors measure vibrations in civil infrastructures.
- Asset tracking, using active RFID to accurately track the real-time location of assets.
- Automatic meter reading, where a gateway collects readings from electricity, water or gas meters.
- Power grid protection and control (substation automation), where sporadic but time-critical data exchange is performed among monitoring units.

In networks enabling IoT applications, a diverse set of challenges and performance requirements, ranging from low latency and high reliability to the sheer scale of access attempts and energy efficiency, need to be satisfied. In many applications, the devices need to remain in sleep mode for certain periods of time in order to save energy, and wake up to transmit bursts of data with very diverse traffic patterns, e.g., triggered by events or periodically scheduled. Therefore, the network may change abruptly from idle into saturation when a devices have new data ready in a given time and wake up to transmit simultaneously; this has been referred to in the literature as delta traffic condition or batch arrival [26]. Although the amount of data generated by each device may be relatively low, the total number of devices that can attempt to get access to the wireless channel simultaneously can be potentially larger than the one manageable by traditional ALOHA and CSMA techniques.

The FSA protocol has been identified as a good alternative to handle the delta traffic due to its good performance when optimally configured [27], [28]. In fact, FSA was adopted in the ISO/IEC 18000-7 standard that is used for active RFID systems. In FSA, time is divided into frames which are further divided into slots where devices contend to transmit data. This approach is convenient when the data packets to be transmitted fit in one slot. When data packets have to be fragmented, it is possible to add a reservation mechanism to ALOHA. This is referred to as the reservation FSA protocol (RFSA) [29]. In RFSA, when a device succeeds in transmitting the first packet of a message in a given slot, that slot is reserved for that device in subsequent frames until the last packet of the message is sent. Upon completion of the entire sequence of data packets, the slot is released again for contention. Therefore, each frame can be conceptually split in two parts; one for contention-based access and one for collision-free access. A number of research works have evaluated the performance of FSA and RFSA in terms of average delay required to resolve the contention and energy consumption under delta traffic [28], [30], [31].

It has been shown in the literature that most protocols deriving from ALOHA and CSMA use data to contend and rely on waiting backoff periods for contention resolution, thus fall short to provide good performance under heavy-loaded networks with a high density of devices. Some studies show how appropriate parameter selection in ALOHA and CSMA can be optimised in terms of maximum throughput [32]. However, systems based on these protocols are prone to suffer from congestion, thus not being able to provide any service; since the selection of the backoff parameters requires an estimation of the traffic load, which may be difficult in highly dense M2M networks. Similar conditions happen in spontaneous crowd aggregations where it is hardly possible to establish mobile connectivity. The IoT is foreseen to become a constant aggregation of crowds and machines requiring connectivity.

A promising strategy to improve the maximum stable throughput of random access protocols based on ALOHA is to use a Collision Resolution Algorithm (CRA) [33]. The CRAs resolve collisions by organizing the retransmission of colliding packets in such a way that all packets are always successfully transmitted with finite delay. The basic CRA is the tree-splitting or Contention Tree Algorithm (CTA), which iteratively splits a large group of contenders into smaller sub-groups in order to reduce collisions in an efficient manner. The tree-splitting algorithms implemented in [27], [34] use the same resources (slots) to transmit data and resolve contention, thus not attaining all the potential gains that this approach can offer. Instead of using data transmissions for contention, it is possible to separate contention from data through the use of contention-based access requests using minislots. Since these minislots can be much shorter than the duration of a data packet, the performance of the network can be improved. This concept is the foundation of the DQ protocol that will be reviewed in the next Section, which combines a CTA with the logic of two distributed queues to manage the contention resolution and the collision-free data transmission, respectively. In the next section, the DQ technology is presented.

IV. A Promising Approach: Distributed Queuing

In this section, Distributed Queuing is presented as an approach to addresses many of the MAC level issues associated with IoT. First, related work on DQ is presented, followed by a detailed description of the mechanism.

A. Related work on DQ

Distributed Queuing (DQ) was first introduced by Xu and Campbell as a novel MAC protocol whose performance is independent of the number of devices sharing a common channel [35], [36]. It was originally designed for cable TV distribution (DQRAP, DQ Random Access Protocol [35]). Following the initial design, it has been adapted to different types of communication networks. Since the first DQ algorithm was proposed, several studies have demonstrated the stability of its performance and the near-optimum behaviour in terms of channel utilization, access delay, and energy consumption for many system layouts. Relevant studies have provided extensions of the basic protocol mechanisms, including:

- Wired centralized networks: extended DQRAP [37] and prioritized DQRAP [38].
- Satellite communications: adapted for long propagation delays on interleaved DQRAP [39].
- Code-Division Multiple Access: in the context of 3G cellular networks, DQRAP/CDMA [40], improves the capacity of random access channels in terms of throughput stability and delay characteristics.
- Wireless Local Area Networks: cross-layer enhancements, referred to as DQ Collision Avoidance [41], where the key benefit lies in a better handling of heterogeneous traffic constituted by voice and data streams.
- Cooperative communications: to coordinate the relay retransmissions in a Cooperative Automatic Retransmission Request (C-ARQ) scheme for wireless network [42]).
- Wireless ad-hoc Networks: for half-duplex radio stations in single-hop networks to improve throughput and average transmission delay; DQ MAC protocol for Ad Hoc Networks [43].
- Low-Power Wireless Networks: for data collection scenarios with a large number of nodes that generate bursty traffic using low-power commercial radio transceivers [44].
- Body Area Networks: adaptation for body sensor networks, referred to as DQ Body Area Network [45], which considers restrictive latency requirements in the healthcare domain and limited energy availability.

All these works consider that devices generate packets following a random Poisson distribution and study the steady-state performance of the protocol. Under these conditions, results illustrate the key features of DQ, which can be summarized as follows:

- DQ behaves as a random access scheme under low traffic and switches to a reservation-based access scheme when the traffic load increases.
- DQ provides near-optimum performance in terms of throughput and delay.
- DQ achieves maximum performance using three access request slots regardless of the traffic load.

DQ allows almost full channel utilization independently of the number of the transmitting devices and the traffic pattern. What is more important, this can be achieved without knowledge of the composition, topology and members of the networks. The details are presented in the following sections.

B. Mechanism Overview

The fundamental idea of DQ is to gather access requests into short contention windows, keeping data transmissions free of collisions. The transmission resources are divided into two parts; a small part is used for access requests and a predominant part is used for collision-free data. The frame structure of DQ is then composed of the parts illustrated in Figure 1: (i) m slots for collision resolution, and (ii) one slot for collision-free data. For the operation of the DQ protocol, it is necessary to count with the presence of a network coordinator that can provide periodic feedback on the result of the contention resolution and data transmission processes. The coordinator will process every frame and transmit a corresponding Feedback Packet (FBP) with the state of the contention slots. The behaviour of DQ starts as follows: at the first frame, devices randomly choose one of the m available contention slots to transmit an Access Request Sequence (ARS). With the FBP, devices receive information on whether the ARS was successfully decoded or collided. Depending on the feedback, each device enters into one of two distributed queues:

1) Colliding devices enter the Contention Resolution Queue (CRQ). A tree-splitting algorithm is then used to resolve the contention.
2) Succeeding devices enter the Data Transmission Queue (DTQ). In this case, a first-in first-out queue allows devices to transmit data in subsequent frames using the data slot.

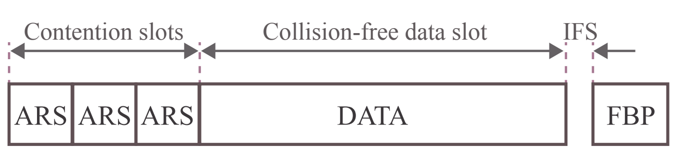

Fig. 1. DQ Frame structure, consisting of m slots for contention resolution and one slot for collision-free data transmission.

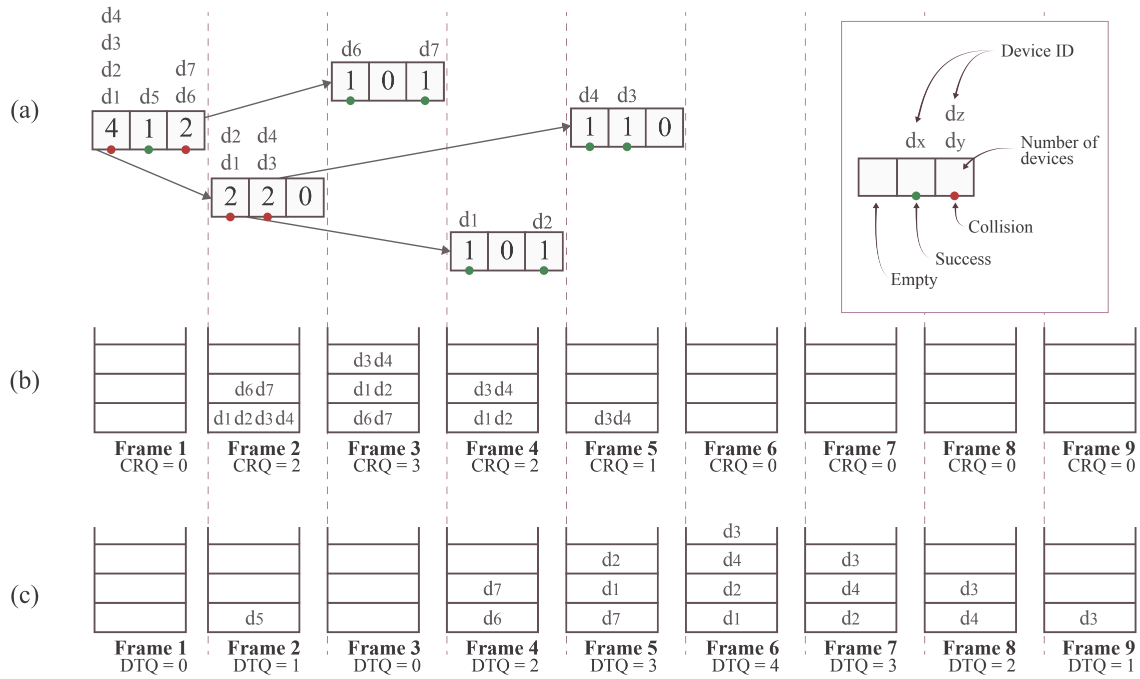

Fig. 2. Example of the DQ protocol, with 7 devices: (a) Tree-splitting algorithm (b) CRQ behaviour per frame in the contention resolutions (c) DTQ behaviour per frame in the data transmission. Three contention slots are available on each frame.

Each queue is represented at each device by two integer numbers which indicate: (i) the length of the queue, and (ii) the position of each device in the queue. The length of each queue is updated by the coordinator after each frame. In the next sections, the operation of DQ is explained. The description is divided into two separate stages; namely, the contention resolution queue and the data transmission queue.

C. Contention Resolution Queue (CRQ)

The first stage corresponds to the contention resolution, where a tree-splitting algorithm is used to resolve the contention in groups; Fig. 2.a depicts a representation of the tree-splitting algorithm execution, considering 7 devices and 3 contention slots. On the first frame, devices select a contention slot to request access with an ARS. In the case where more than one device selects the same contention slot, a subsequent contention slot will be assigned to the group of colliding devices. The length of CRQ then represents the number of sub-groups of devices waiting to retransmit an ARS.

Devices must compute the length of the CRQ and their position in it. To achieve so, the FBP provides the contention status and the CRQ length. The feedback information must consider differentiation of three states for each contention slot: empty, collision and success. Based on this feedback information, each device computes its representation of the CRQ by means of two integer numbers:

1) Calculation of the CRQ length (RQ counter): the value of the counter is increased by one unit for each collision state accounted in the previous frame. At each frame, the counter is decreased by one, to account for the frame execution. The RQ counter and the state of the m contention slots are updated by the coordinator and signalled in the FBP.
2) Calculation of the device position in the CRQ (pRQ counter): if the device is waiting in the CRQ, it must first decrease its representation of the pRQ counter by one unit at each frame. In case the device has attempted an access on the previous frame and collided, then it sets its pRQ counter to point at RQ, i.e., to the end of the CRQ.

The devices that occupy the first position in CRQ will transmit a new ARS in the next frame. Since the length of the CRQ is decremented by one after each frame, the devices only need to receive the FBP in those frames where they transmit the ARS. Therefore, the devices can switch to sleep mode during those frames where they do not transmit access requests. Figure 2.b illustrates the example of the CRQ with 7 devices (d1 to d7) and 3 contention slots. At frame 1, all the devices contend: d1, d2, d3 and d4 collide in slot 1; d5 succeeds in slot 2; and d6 and d7 collide in slot 3. Thus, d1, d2, d3 and d4 enter in the first position of CRQ; d6 and d7 enter in the second position of CRQ; and d5 enters in the first position DTQ. At frame 2, only d1, d2, d3 and d4 contend (they are in the first position of the CRQ), and d5 transmits data. d1 and d2 collide to each other on slot 1, while d3 and d4 collide on slot 2. Both groups enter at the end of the CRQ on positions 2 and 3, respectively. At frame 3, d6 and d7 contend; both succeed and leave the CRQ. At frame 4, d1 and d2 contend again and succeed. Finally, d3 and d4 succeed at frame 5. At frames 6 to 9 the CRQ is empty, no device contends.

TABLE III
MAPPING BETWEEN DQ CONCEPTS AND LTE TERMINOLOGY

DQ concept	Adaptation for LTE
Access Request Sequence (ARS)	Preamble sequence
Contention Slot	Random Access Slot (RA Slot)
Feedback Packet (FBP)	Random Access Response (RAR)

D. Data Transmission Queue (DTQ)

After a contention is resolved and the device has received a success feedback, the device is virtually organised into a Data Transmission Queue (DTQ). The CRQ and the DTQ procedures work in parallel, but a device must first successfully exit the CRQ in order to enter the DTQ. The behaviour below describes the DTQ when the data transmission is performed on a fixed-size resource, i.e., there is no dynamic resource allocation and all transmissions are granted for the same predefined resource, shared on a time basis. Devices use two counters in order to keep track of the DTQ:

1) Calculation of the DTQ length (TQ counter): the value of the counter is increased by one unit for each success state accounted in the previous frame. After a data transmission occurrence, the counter is decreased by one. The TQ counter is updated by the coordinator and signalled in the FBP.
2) Calculation of the device position in the DTQ (pTQ counter): When a device enters the DTQ, it points the pTQ at the end of the queue, which corresponds to the TQ value. If the device is waiting in the DTQ, it must first decrease its representation of the pRQ counter by one unit every frame (at the occurrence of each transmission).

The device that occupies the first position in DTQ will transmit a data packet in the next frame. Since the length of the DTQ is decremented by one after each frame, the devices only need to receive the FBP in those frames where they transmit data. Therefore, the devices can switch to sleep mode during those frames where they do not transmit.

Figure 2.c shows the example of the DTQ with 7 devices (d1 to d7) and 3 contention slots. At frame 1, all the devices contend and no device is transmitting. At frame 2, d5 transmits data. At frame 3, no device is able to transmit data due to unresolved contentions. At frame 4, d6 transmits data and d7 remains in the DTQ. At frame 5, d7 transmits data; d1 and d2 enter the DTQ. At frame 6, d1 transmits data, d2 remains in the DTQ; d4 and d3 enter the DTQ after resolving the contention. At frames 7, 8 and 9, d2, d4 and d3 transmit data, respectively.

In the next section, specific research regarding the use of DQ in the IoT context is presented.

V. DQ IN THE IOT CONTEXT

As it has been previously discussed in Section III, IoT brings new challenges in terms of traffic patterns, even imposing abrupt changes from idle into saturation when a large number of devices transmit data simultaneously. The consideration of massive MTC in 3GPP systems has motivated the proposal of multiple amendments and alternative solutions to efficiently resolve congestion based on large number of devices, as compared in [13]. However, the majority of the proposals fall short to provide a fair balance between access delay, access probability rate, and energy consumption. In the following subsections, two cases are described: a proposal to implement the CRQ principles in LTE networks and a DQ demo on active RFID communications at 433 MHz. Both cases address the same challenges of efficiently handle a massive number of devices.

A. DQ for the Random Access procedure in LTE

The authors have studied the possibility of implementing DQ principles in the RA procedure of the LTE standard [46]. The detailed LTE RA procedure in standard LTE networks is explained in [46]. As discussed in Section II-C, it is based on FSA scheme, where devices use orthogonal preamble sequences over RA slots to contend for network access. Devices randomly select one of the available preambles (with a maximum of 64 possibilities) and transmit it over the RA slot. The base station then process the received preambles and provides a feedback in a message referred to as the Random Access Response (RAR). The RAR informs devices if a collision was detected for their preamble, in which case devices are signalled to perform a backoff time before the next contention attempt. The RAR also conveys information for successfully decoded preambles, which include the resource grant for devices to transmit a connection request.

Since the DQ and LTE terminologies might create confusion, Table III provides a mapping between the relevant DQ concepts and the interpretation given for the LTE implementation. The DQ principles can be adapted to the operation of the LTE standard, leveraging the availability of the orthogonal preambles used for the initial access[2]; this means that more than one preamble can be detected over the same RA slot, and collisions occur when more than one device selected the same preamble and transmits it over the same RA slot.

[2]DQ implementations can vary on the resource they use for contention. Most of the adaptations use contention slots (or control minislots) where transmitting devices send a signature, i.e., a pseudo-random sequence. The LTE adaptation makes use of the orthogonal sequences over the same slot; in such case, the contending resource corresponds to the sequences thereof.

The DQ implementation for LTE networks behaves as follows: upon initial access, devices select an RA slot and wait for the corresponding feedback message before attempting to request access. This way, devices are not allowed to use a RA slot where previous collisions are being resolved (blocked access). The feedback message is provided with some modifications to the RAR. In [46], a CRQ sub-header is proposed as the solution to provide the three feedback states required by the DQ principles (success, collision and empty states).

In order to verify the DQ proposal for the LTE RA procedures, system simulations have been conducted. To be able to efficiently simulate the large number of devices considered for IoT scenarios, independent LTE RA modules in FDD mode have been developed in ns-3 simulator[3]. The modules where validated in [13], [47] by replicating the simulation conditions and parameters provided by the 3GPP in [48] and comparing the performance results. In particular, the CRQ mechanisms have been implemented in these modules to compare the performance of the contention resolution with the standard LTE RA procedure.

The simulation scenario assumes a cellular LTE network where devices are cell-synchronized and have already received all configuration parameters related to the RA procedure. Transmissions related to system information are not considered for the simulation modules. As described in Section III, a delta traffic condition or batch arrival is considered [26] for a varying number of simultaneous access attempts, up to 1500. We consider different number of available preambles to show the scalability of each procedure. Details on the simulation parameters are provided in Table IV. Four performance metrics are used to compare the standard LTE RA procedure and the DQ proposal:

1) *Blocking Probability*: the probability of a device reaching the maximum number of attempts and being unable to complete an access process.
2) *Average Access Delay*: the average time elapsed between the RA procedure initiation and the reception of the contention resolution message by the eNodeB. Only successful accesses are considered for the average calculation.
3) *Average Energy Consumption*: the average energy spent between the RA procedure initiation and the reception of the contention resolution message by the eNodeB. Only successful accesses are considered for the average calculation.
4) *Average Number of Preamble Retransmissions*: the average number of access attempts that a device executes before receiving an access. If a device reaches the maximum number of retransmissions attempts and it is not able to resolve the contention, it is considered to be blocked by the network.

Fig. 3 shows the performance comparison between the standard LTE RA procedure and the CRQ implementation for contention resolution. Results demonstrate the superior performance of the DQ discipline with realistic amendments to the standard operation. The standard LTE RA procedure is capable to support large number of simultaneous arrivals by increasing different backoff indications in order to spread in time subsequent attempts. Increasing the backoff indication or setting more restrictive barring factor may ease the congestion experienced on the network side; however, our experiment reveals the negative effect on the device side. Increasing the backoff indications not only affects the average access delay, but it also results in negative implications for the energy consumption on the device side. The CRQ implementation performance is also affected

[3]This paper contains supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors. The material corresponds to the RA modules described in this work and used to create the results presented in this section. The material includes a readme file with usage instructions and links to the official ns-3 simulator installation and requirement guidelines.

TABLE IV
SIMULATION PARAMETERS

Parameter	Value				Unit
No. Available preambles	56	36	18	6	int.
Barring Factor[a]	80	60	40	40	%
Barring Time[a]	2				s
PRACH Configuration Index[b]	3				int.
Backoff Indicator[b]	480				ms
Preamble duration	1				ms
Max. Preamble retransmissions[b]	20				int.
RAR Window Size[a]	5				ms
Contention Resolution Timer[a]	48				ms
Power consumption values[c]					
Transmission	500				mW
Active Period (Reception mode)	150				mW
Accurate clock (Idle mode)	10				mW
Number of iterations	500				int.
Simulation time per iteration	20				s

[a] Standard values available in 3GPP TS 36.331 [49].
[b] Standard values available in 3GPP TS 36.321 [50].
[c] Values taken from the description given in [51], assuming that the power consumption on transmission mode is equal to the radiated power.

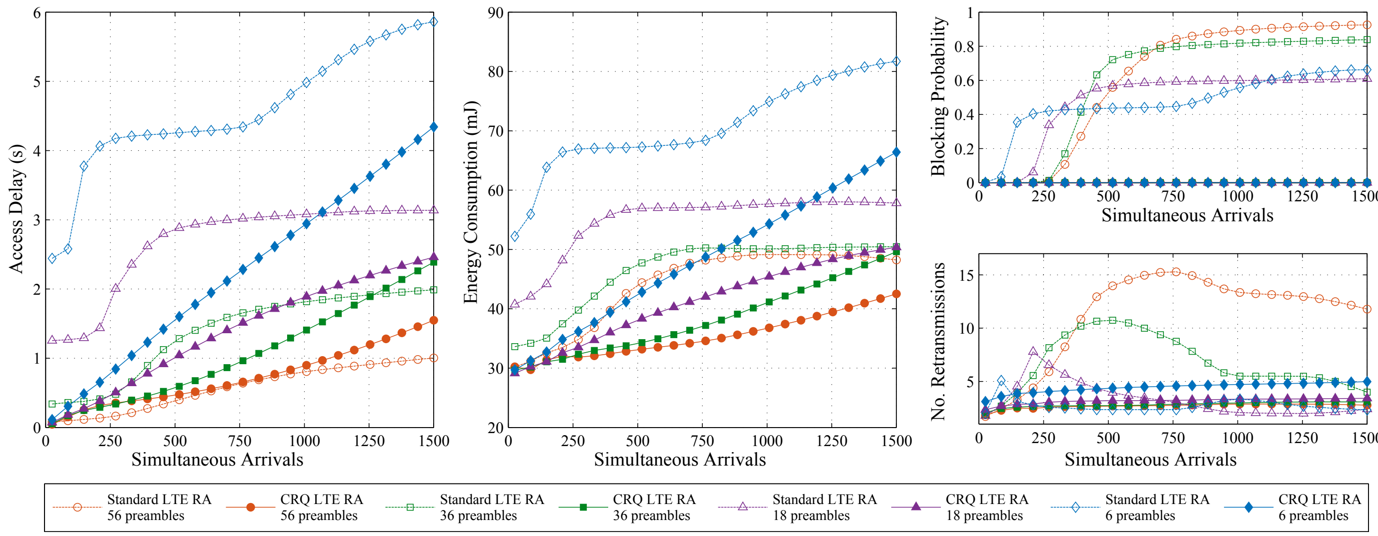

Fig. 3. Comparison between the LTE standard RA procedure and the DQ-based adaptation for contention resolution with up to 1500 simultaneous arrivals. Results show that the average access delay, average energy consumption, average number of preamble retransmissions and the blocking probability when using different number of orthogonal preambles.

by the increase on the simultaneous arrivals, but it provides a sustained performance of the blocking probability, which is not affected by the increase in simultaneous arrivals; illustrating the efficient performance independently of the number of contending devices, even in the extreme case of 1500 simultaneous arrivals while only using 6 orthogonal preambles for contention. Moreover, the average number of retransmissions is lower than 5 for all the conditions presented in Fig. 3.

To date, there is no study assessing a feasible adaptation of the DTQ in LTE systems. However, based on the DQ principles, the idea should be to allocate predefined transmission resources in data uplink that devices can access following the DTQ order. Such alternative would provide the additional benefit of reducing the signalling transmissions related to the connection set-ups, which has been widely discussed as a challenge for devices that transmit short data streams under limited energy availability.

B. Implementation of DQ in M2M area networks

The first proof-of-concept of the DQ technology in a wireless system was development in 2014. The work in [52], [53] presents a demonstration of the operation of DQ in a real M2M area network targeting data collection scenarios using active RFID systems operating at the 433 MHz band. The protocol implemented is named Low-Power DQ (LPDQ). LPDQ is based on a packet-based preamble sampling to achieve tag synchronization and DQ as the channel access mechanism. In [52], LPDQ is compared to the MAC protocol defined in the ISO 18000-7 standard for RFID, which is based on an analog preamble sampling and FSA scheme.

The experimental demo presented is composed of up to 30 active tags (or devices) and 1 reader connected to a computer that acts as coordinator. The reader and the tags are implemented using the OpenMote-433, which is based on the CC430 System-on-Chip from Texas Instruments. The CC430 includes an MSP430 16bit RISC microcontroller and a CC1101 radio transceiver, which operates at sub-GHz bands with data rate up to 600 kbps and supports ASK, OOK, FSK and MSK modulations. Fig. 4 shows a picture of the devices used in the demonstration.

Each test of FSA and LPDQ consists of two phases: *i*) synchronization and *ii*) data collection. During the synchronization phase, the tags are in preamble sampling mode, switching periodically between sleep and receive modes in order to detect wake-up packets from the reader. The reader transmits a sequence of wake-up packets to synchronize the tags. Once the tags are synchronized, the data collection phase starts. During the data collection phase, each tag executes the rules of the configured MAC layer and transmits a predefined number of data packets to the reader.

The results from the test measurements conclude that LPDQ outperforms FSA in terms of delay and energy consumption. In LPDQ there are no collisions during data packet transmission, which reduces the energy consumption of the tags by more than a 50% because no energy is wasted in the retransmission of data packets. In addition, the performance of LPDQ is independent of the number of tags, which means that it is not needed to adjust the frame length based on the number of collisions as in FSA. And finally, LPDQ reduces the delay in data collection because the collision resolution and the data transmission are interleaved in time and thus it is not necessary to wait until the query tree is build to start receiving data from the tags that are already in the DTQ. LPDQ represents therefore a major breakthrough in terms of delay, throughput and energy consumption.

Fig. 4. Picture showing the OpenMotes used for the LPDQ demo. On the bottom-left corner is the reader for the coordinator, the remaining devices are active tags.

This demo is an integral part of the IoT device tier of the IoTWORLD end-to-end experimental platform. To explain the platform composition, the authors have included a supplementary MOV file which describes the complete architecture of the platform (506 MB in size). This will be available at http://ieeexplore.ieee.org.

VI. Conclusions

IoT will create a new technology landscape, as mobile communications have already done it. In order for this new paradigm to actually arise, wireless communications become a cornerstone while relevant engineering challenges appear in the design of such kind of networks. Established MAC protocols are a core part of them. The great majority of proposals and solutions proposed during the previous years for contention-based channel access constitute protocol variations of ALOHA with or without listening before talk. It is well known that this fact leads to congestion and energy waste when the traffic load and the number of devices increases, remarking the bottleneck challenges at MAC layer level for IoT.

In the first part of this paper, we review the MAC implementations in currently technologies considered for IoT, aiming to corroborate the former reasoning. Based on the identified limitations of ALOHA-based systems in the context of IoT, we emphasize the benefits of a solution proposed some years ago, namely Distributed Queueing (DQ) protocol. Intensive research related with the benefits of DQ in various communication networks has already been carried out. In particular, DQ operates for any number of devices generating any traffic load and pattern, it always attains the maximum capacity of the channel, it does not need any *a priori* knowledge of the configuration and/or composition of the network, it allows to apply QoS rules and therefore it is considered a promising catalyst for IoT.

In the second part of the paper, we further elaborate on the potential of DQ as a key enabling protocol to address the main challenges of IoT. In particular, a technical feasibility study of applying DQ principles in the RA procedure of LTE standard is performed. A system-level simulation framework is then built for performance evaluation and comparison between the standardized RA method and the DQ-based adaptation for contention resolution. The simulation results reveal the superior performance of DQ discipline assessed through four performance metrics. In an effort to apply DQ mechanism in different communication technologies and evaluate its performance, a demonstration of an M2M area network based on DQ for RFID systems is also presented. The experimental results are discussed to illustrate a clear performance improvement in terms of delay, throughput and energy efficiency, compared with the default FSA scheme.

Considering all these features, any reader may wonder why is not DQ an integral component of communication systems. It is indeed a very interesting question. First of all, the simplicity of the current technology works properly in a wide variety of situations; making it unnecessary to substitute unless a critical situation forces the change. Even if legacy technologies may have shortcomings and inefficiencies, it is possible to overcome some issues in the short term by over-providing resources. It is quite reasonable to expect that IoT may be the key to force the need of a real MAC evolution. Secondly, the implementation of DQ requires certain degree of connection between physical layer and MAC layer developers, between electronic and signal theory engineers. In addition to the MAC change, some simple modifications in the physical layer are needed to actually achieve the maximum benefits of DQ. This is a challenge that must be seen by electronic engineers as a worthy effort, and it is a concrete line of future research and development effort.

Appendix

The acronyms included in this paper along with their definitions are summarized in Table V.

TABLE V
LIST OF ACRONYMS ALONG WITH DEFINITIONS

Acronym	Definition
AP	Access Point
ARS	Access Request Sequence
BEB	Binary Exponential Backoff
BLE	Bluetooth Low Energy
CAP	Contention Access Period
CDMA	Code Division Multiple Access
CFP	Contention Free Period
CIoT	Cellular Internet of Things
CRA	Collision Resolution Algorithm
CRQ	Contention Resolution Queue
CSMA	Carrier Sensing Multiple Access
CSMA/CA	Carrier Sensing Multiple Access with Collision Avoidance
CTA	Contention Tree Algorithm
CTS	Clear-To-Send
DCF	Distributed Coordination Function
DQ	Distributed Queueing
DTQ	Data Transmission Queue
EDCA	Enhanced Distributed Channel Access
FBP	Feedback Packet
FSA	Frame Slotted ALOHA
GTS	Guaranteed Time Slot
HTC	Human-Type Communications
IoT	Internet of Things
ISM	Industrial, Scientific and Medical
LPWAN	Low Power Wide Area Network
LTE	Long Term Evolution
MAC	Medium Access Control
MTC	Machine-Type Communications
NB-IoT	Narrow-Band IoT
QoS	Quality of Service
R-FTDMA	Random Frequency and Time Division Multiple Access
RA	Random Access
RACH	Random Access CHannel
RAW	Restricted Access Window
RFID	Radio Frequency IDentification
RTS	Request-To-Send
TDMA	Time Division Multiple Access
UE	User Equipment
WLANs	Wireless Local Area Networks
WSNs	Wireless Sensor Networks

ACKNOWLEDGEMENT

This work has been partially funded by Wireless@KTH, projects ADVANTAGE (FP7-607774), CellFive (TEC2014-60130-P) and DEFINE5G (TEC2014-60258-C2-2-R) and the Catalan Government under grants (2014-SGR-1551) and (2014-SGR-1160).

REFERENCES

[1] IEEE Std. 802.15.4, "IEEE 802.15.4 Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)," 2006.
[2] P. Park, P. Di Marco, C. Fischione, and K. Johansson, "Modeling and Optimization of the IEEE 802.15.4 Protocol for Reliable and Timely Communications," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 24, no. 3, pp. 550–564, Mar. 2013.
[3] Y. Zhu, K. Chi, X. Tian, and V. Leung, "Network Coding Based Reliable IPv6 Packet Delivery over IEEE 802.15.4 Wireless Personal Area Networks," *Vehicular Technology, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.
[4] M. Palattella, N. Accettura, L. Grieco, G. Boggia, M. Dohler, and T. Engel, "On Optimal Scheduling in Duty-Cycled Industrial IoT Applications Using IEEE802.15.4e TSCH," *Sensors Journal, IEEE*, vol. 13, no. 10, pp. 3655–3666, Oct. 2013.
[5] C. Gomez, J. Oller, and J. Paradells, "Overview and Evaluation of Bluetooth Low Energy: An Emerging Low-Power Wireless Technology," *Sensors*, vol. 12, no. 9, pp. 11 734–11 753, 2012. [Online]. Available: http://www.mdpi.com/1424-8220/12/9/11734
[6] Bluetooth specification version 4.0. [Online]. Available: https://www.bluetooth.org/en-us/specification/adopted-specifications.
[7] D. K. Klair, K. W. Chin, and R. Raad, "A Survey and Tutorial of RFID Anti-Collision Protocols," *IEEE Communications Surveys Tutorials*, vol. 12, no. 3, pp. 400–421, Third 2010.

[8] IEEE Standard 802.11, "IEEE 802.11 Wireless LAN Medium Access Control and Physical Layer Specifications)," 1999.
[9] G. Cena, L. Seno, A. Valenzano, and C. Zunino, "On the Performance of IEEE 802.11e Wireless Infrastructures for Soft-Real-Time Industrial Applications," *Industrial Informatics, IEEE Transactions on*, vol. 6, no. 3, pp. 425–437, Aug. 2010.
[10] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE J.Sel. A. Commun.*, vol. 18, no. 3, pp. 535–547, Sep. 2006. [Online]. Available: http://dx.doi.org/10.1109/49.840210.
[11] S. Mangold, S. Choi, G. Hiertz, O. Klein, and B. Walke, "Analysis of IEEE 802.11e for QoS support in wireless LANs," *Wireless Communications, IEEE*, vol. 10, no. 6, pp. 40–50, Dec. 2003.
[12] T. Adame, A. Bel, B. Bellalta, J. Barcelo, and M. Oliver, "IEEE 802.11ah: the WiFi approach for M2M communications," *Wireless Communications, IEEE*, vol. 21, no. 6, pp. 144–152, Dec. 2014.
[13] A. Laya, L. Alonso, and J. Alonso-Zarate, "Is the Random Access Channel of LTE and LTE-A Suitable for M2M Communications? A Survey of Alternatives," *Communications Surveys Tutorials, IEEE*, vol. 16, no. 1, pp. 4–16, First 2014.
[14] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: issues and approaches," *Communications Magazine, IEEE*, vol. 51, no. 6, pp. 86–93, Jun. 2013.
[15] 3GPP TS 22.368 v12.1.0, "Service requirements for Machine-Type Communications (MTC) Stage 1," Release 12, Dec. 2012.
[16] 3GPP TR 45.820 V13.0.0, "Cellular system support for ultra-low complexity and low throughput Internet of Things (CIoT)," 2015.
[17] S. Andreev, O. Galinina, A. Pyattaev, M. Gerasimenko, T. Tirronen, J. Torsner, J. Sachs, M. Dohler, and Y. Koucheryavy, "Understanding the IoT connectivity landscape: a contemporary M2M radio technology roadmap," *Communications Magazine, IEEE*, vol. 53, no. 9, pp. 32–40, Sep. 2015.
[18] Sigfox. [Online]. Available: http://www.sigfox.com/.
[19] LoRa Alliance. [Online]. Available: https://www.lora-alliance.org/.
[20] Weightless. [Online]. Available: http://www.weightless.org/.
[21] Weightless System specification Version 0.8. [Online]. Available: http://usgroup.eu/wp-content/uploads/2013/02/Weightless-Specification-v0.8.pdf
[22] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted aloha," *IEEE Transactions on Communications*, vol. 59, no. 2, pp. 477–487, February 2011.
[23] F. Clazzer and C. Kissling, "Enhanced contention resolution aloha - ecra," in *Systems, Communication and Coding (SCC), Proceedings of 2013 9th International ITG Conference on*, Jan 2013, pp. 1–6.
[24] C. Stefanovic and P. Popovski, "Aloha random access that operates as a rateless code," *IEEE Transactions on Communications*, vol. 61, no. 11, pp. 4653–4662, November 2013.
[25] E. Paolini, G. Liva, and M. Chiani, "Coded slotted aloha: A graph-based method for uncoordinated multiple access," *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6815–6832, Dec 2015.
[26] L. Kleinrock, *Queueing Systems*. Wiley Interscience, 1975, vol. I: Theory.
[27] Y. Xinqing and L. Xuemei, "Evaluating the energy consumption of RFID tag collision resolution protocols," in *International Conference on Mobile Ad-hoc and Sensor Networks*, Feb. 2010, pp. 215–221.
[28] V. Namboodiri, M. DeSilva, K. Deegala, and S. Ramamoorthy, "An extensive study of slotted Aloha-based RFID anti-collision protocols," *Elsevier Computer Communications*, vol. 35, no. 16, pp. 1955–1966, 2012.
[29] W. Crowther, R. Rettberg, D. Walden, S. Ornstein, and F. Heart, "A system for broadcast communication: Reservation-ALOHA," in *International Conference on Systems Sciences*, 1973, pp. 371–374.
[30] F. Vazquez-Gallego, L. Alonso, and J. Alonso-Zarate, "Energy and Delay Analysis of Contention Resolution Mechanisms for Machine-to-Machine Networks based on Low-Power WiFi," in *IEEE International Conference on Communications (ICC)*, Jun. 2013.
[31] ——, "Modeling and Analysis of Reservation Frame Slotted-ALOHA in Wireless Machine-to-Machine Area Networks for Data Collection," *Sensors*, vol. 15, no. 2, p. 3911, 2015. [Online]. Available: http://www.mdpi.com/1424-8220/15/2/3911.
[32] L. Dai and X. Sun, "A unified analysis of ieee 802.11 dcf networks: Stability, throughput, and delay," *IEEE Transactions on Mobile Computing*, vol. 12, no. 8, pp. 1558–1572, Aug 2013.
[33] J. I. Capetanakis, "Tree Algorithms for Packet Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 505–515, Sep. 1979.
[34] V. Namboodiri and L. Gao, "Energy-Aware Tag Anticollision Protocols for RFID Systems," *IEEE Transactions on Mobile Computing*, vol. 9, no. 1, pp. 44–59, Jan. 2010.
[35] W. Xu and G. Campbell, "A Near Perfect Stable Random Access Protocol for a Broadcast Channel," in *IEEE International Conference on Communications (ICC)*, vol. 1, Jun. 1992, pp. 370–374.
[36] ——, "A distributed queueing random access protocol for a broadcast channel," in *Conference Proceedings on Communications Architectures, Protocols and Applications*, ser. SIGCOMM '93. New York, NY, USA: ACM, 1993, pp. 270–278. [Online]. Available: http://doi.acm.org/10.1145/166237.166263
[37] W. Chien-Ting and G. Campbell, "Extended DQRAP (XDQRAP). A cable TV protocol functioning as a distributed switch," in *International Workshop on Community Networking Integrated Multimedia Services to the Home*, Jul. 1994, pp. 191–198.
[38] L. Harn-Jier and G. Campbell, "PDQRAP-prioritized distributed queueing random access protocol," in *IEEE Conference on Local Computer Networks*, 1994, pp. 82–91.
[39] W. Chien-Ting and G. Campbell, *Interleaved DQRAP with Global TQ*. Dept. of Comp. Sci., Illinois Inst. of Technology, 1995.
[40] L. Alonso, R. Agusti, and O. Sallent, "A near-optimum MAC protocol based on the distributed queueing random access protocol (DQRAP) for a CDMA mobile communication system," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 9, pp. 1701–1718, Sep. 2000.
[41] E. Kartsakli, J.Alonso-Zarate, C. Verikoukis, and L. Alonso, "Cross-Layer Enhancement for WLAN Systems with Heterogeneous Traffic based on DQCA," in *IEEE Communications Magazine*, vol. 46, no. 6, Jun. 2008, pp. 60–66.
[42] J. Alonso-Zarate, L. Alonso, C. Skianis, and C. Verikoukis, "Analysis of a Distributed Queuing Medium Access Control Protocol for Cooperative ARQ," in *IEEE GLOBECOM*, Dec. 2010, pp. 1–5.
[43] J. Alonso-Zarate, E. Kartsakli, L. Alonso, and C. Verikoukis, "Performance Analysis of a Cluster-Based MAC Protocol for Wireless Ad Hoc Networks," in *EURASIP Journal on Wireless Communications and Networking*, 2010.
[44] P. Tuset-Peiro, F. Vazquez-Gallego, J. Alonso-Zarate, L. Alonso, and X. Vilajosana, "LPDQ: A self-scheduled TDMA MAC protocol for one-hop dynamic low-power wireless networks," *Pervasive and Mobile Computing*, vol. 20, pp. 84 – 99, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574119214001576
[45] B. Otal, C. Verikoukis, and L. Alonso, "Highly Reliable Energy-saving MAC for Wireless Body Sensor Networks in Healthcare Systems," in *IEEE Journal on Selected Areas in Communications, Special Issue on Wireless and Pervasive Communications in Healthcare*, vol. 27, no. 4, May 2009, pp. 553–565.
[46] A. Laya, L. Alonso, and J. Alonso-Zarate, "Contention Resolution Queues for Massive Machine Type Communications in LTE," in *IEEE PIMRC, Workshop on Machine-to-Machine Communications*, Jul. 2015.
[47] A. Laya, L. Alonso, P. Chatzimisios, and J. Alonso-Zarate, "Massive access in the random access channel of lte for m2m communications: An energy perspective," in *Communication Workshop (ICCW), 2015 IEEE International Conference on*, Jun. 2015, pp. 1452–1457.
[48] 3GPP TR 37.868 V11.0.0, "Study on RAN Improvements for Machine-type Communications," Sep. 2011.
[49] 3GPP TS 36.331 V10.5.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC)," Mar. 2012.
[50] 3GPP TS 36.321 V9.3.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC)," Jun. 2010.
[51] T. Tirronen, A. Larmo, J. Sachs, B. Lindoff, and N. Wiberg, "Machine-to-machine communication with long-term evolution with reduced device energy consumption," *Transactions on Emerging Telecommunications Technologies*, vol. 24, no. 4, pp. 413–426, 2013.

[52] P. Tuset-Peiro, F. Vazquez-Gallego, J. Alonso-Zarate, X. Vilajosana, and L. Alonso, “Demonstrating Low-Power Distributed Queuing for Active RFID communications at 433 MHz,” in *IEEE INFOCOM*, May 2014.

[53] P. Tuset-Peiro, F. Vazquez-Gallego, J. Alonso-Zarate, L. Alonso, and X. Vilajosana, “Experimental energy consumption of frame slotted aloha and distributed queuing for data collection scenarios,” *Sensors*, vol. 14, no. 8, p. 13416, 2014. [Online]. Available: http://www.mdpi.com/1424-8220/14/8/13416