Distributed Queueing-Based Random Access Protocol for LoRa Networks

Wei Wu, Yan Li, Yanhe Zhang, Bin Wang, and Wennai Wang

***Abstract*—Among many others, LoRa/LoRaWAN is one of the most widely employed technologies to support Internet of Things applications for low-power and long-range wireless devices. LoRaWAN is a medium access control (MAC) layer protocol based on the traditional Aloha algorithm which reveals a great degraded throughput when a large number of devices attempt to communicate over the shared channel at the same time. Aiming at this scalability issue, we propose to introduce the distributed queueing (DQ) algorithm into the MAC protocol over the LoRa physical layer, named DQ-LoRa. The details of DQ-LoRa are provided about the frame structure and access procedure, followed by the performance evaluations on system throughput, average delay, and energy consumption. Our analytical results show that the maximum throughput of DQ-LoRa is independent of the number of devices, and the delay is insensitive to the number of random access minislots for device contending. The numerical computation shows the gain of throughput up to 2.6-fold, and the savings of energy and latency up to 48% and 54%, respectively, in comparing DQ-LoRa with a pure Aloha system.**

***Index Terms*—Distributed queueing (DQ), LoRa, medium access control (MAC), performance analysis.**

I. Introduction

LOW-POWER wide-area networking (LPWAN) enables power efficient wireless communications over very long distances, which offers a significant potential to support a large number of Internet of Things (IoT) applications. LoRa, a physical (PHY) layer modulation technique based on the chirp spread spectrum (CCS), is one of the most successful and adopted LPWAN technologies [1], [2]. Since battery powered nodes in LoRa networks share the unlicensed ISM band and require a long lifetime, an effective medium access control (MAC) mechanism is critical for LoRa. A widely used MAC layer over the LoRa PHY layer is based on pure Aloha fashion, which is specified by LoRaWAN. However, the Aloha's scalability issue is notorious, i.e., it suffers from congestion under high traffic load.

An alternative solution to Aloha is distributed queueing (DQ), which ensures high throughput regardless of the traffic load and traffic pattern [3]. DQ was first introduced by Xu and Campbell as a novel MAC protocol for cable TV distribution [4]. The stability of its performance and near-optimum behavior in terms of channel utilization, access delay, and energy consumption have been demonstrated in [5] and [6].

With the motivation to eliminate the limitations of LoRa networks resulting from LoRaWAN Aloha MAC, in this article, we introduce the DQ algorithm into the MAC layer over LoRa, named DQ-LoRa. We believe that DQ-LoRa substituted for the Aloha-based MAC layer of commercial LoRaWAN devices is not so difficult, since most LoRaWAN MAC functionalities are implemented in software [7]. In addition to that our proposal is designed to be compatible to Class A by utilizing beacon signals defined for Class B in traditional LoRaWAN [8]. Therefore, our proposal requires LoRa end-devices (EDs) that can switch between Classes A and B modes according to the beacon signal.

The following are our main contributions.

1) DQ-LoRa is proposed to address the scalability issue of LoRaWAN, as a candidate for IoT applications with unpredictable number of connected devices and bursty traffic. Its frame structure as well as the access procedure are provided in detail.
2) The performance of DQ-LoRa is analyzed theoretically, in terms of the system throughput, average delay, and average energy consumption. It is also compared with pure Aloha and CSMA.
3) The results demonstrate that with a large number of active devices, DQ-LoRa outperforms pure Aloha in all the aspects of system performance. The throughput gain can be up to 2.6 times, as well as the delay and energy savings up to 54% and 48%, respectively.

The rest of this article is organized as follows. Related works are discussed in Section II. In Section III, an overview of LoRa and LoRaWAN is provided. Section IV is devoted to describing our design of DQ-LoRa in detail. Section V presents the performance analysis of DQ-LoRa. The conclusions are made in Section VI.

Manuscript received May 28, 2019; revised September 4, 2019; accepted September 24, 2019. Date of publication October 3, 2019; date of current version January 10, 2020. This work was supported in part by the Postgraduate Research and Practice Innovation Programm of Jiangsu Province under Grant KYCX18_0897 and Grant SJCX18_0278. *(Corresponding author: Wennai Wang.)*

W. Wu, B. Wang, and W. Wang are with the Key Laboratory of Broadband Wireless Communication and Sensor Network Technology, Ministry of Education, Nanjing University of Posts and Telecommunications, Nanjing 210003, China, and also with the College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: 2017010205@njupt.edu.cn; wangbin7062@sohu.com; wangwn@njupt.edu.cn).

Y. Li and Y. Zhang are with the College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: 1249873373@qq.com; 1018010324@njupt.edu.cn).

Digital Object Identifier 10.1109/JIOT.2019.2945327

II. Related Works

A. Medium Access Control Protocols Over LoRa

Most of the existing researches focus on LoRaWAN Aloha MAC and its throughput performance is analyzed in [1], and [9]–[11]. In general, at a very low transmission rate the throughput of LoRa networks is impacted by packets collisions, whereas at higher transmission rates it is impacted by duty cycling [1]. Additionally, it has been shown that although LoRaWAN uses Aloha, due to LoRa's robust modulation technique only 32% more packet losses in the case of 1000 nodes per gateway (GW), whereas for the same case the losses are up to 90% in other pure Aloha-based networks [10]. A comprehensive model of LoRaWAN LPWANs in the ns-3 network simulator is presented in [11], which forms the basis for a scalability analysis of LoRaWAN. The simulation results show that downstream traffic has a negative impact on the delivery ratio of confirmed upstream traffic and increasing GW density can ameliorate but not to eliminate this effect.

In addition to that a few studies analyze the throughput of LoRa networks utilizing other MAC protocols. Farooq and Pesch [2] modified the Aloha-based mechanism by using a simple systematic approach. They found that this modification positively impacts the performance of the different communication settings. Besides, a range of MAC protocols have been analyzed in order to search for a suitable MAC protocol for LoRa networks [12], namely, pure Aloha, delay before transmit, random frequency hopping, and carrier sensing multiple access (CSMA). The results demonstrate that CSMA has significant better performance and scalability features in terms of the number of nodes and data traffic generation models.

B. DQ Extensions in Various Networks

DQ has been adapted to many types of the communication systems, such as wired centralized networks [13], [14], satellite communications [15], CDMA [16], WLAN [17]–[19], body area networks [20], wireless ad-hoc networks [21], cooperative networks [22], massive machine type communications in LTE [23], and low-power wireless networks [7], [24]. The results of all these works demonstrate that the superiority of DQ over Aloha-based protocols and its stable and scalable performance for densely loaded networks.

Low-power DQ (LPDQ) [24] is based on the low-power listening for network synchronization, DQ for channel access, and channel hopping against multipath propagation and external interference. It is not intended for low-rate IoT networks, but for high-frequency, high-bandwidth (BW) wireless links. In addition, the synchronization phase is a heavy overhead, which adversely affects channel utility and latency.

DQ-N was proposed for crowdsourcing LPWAN [7] and its extension for LoRa (a demo) is described in [25]. Its main difference from LPDQ is that devices can request multiple data slots in each frame by the transmission request (TR). The 4-bit ID field of TR was not explained in [7], and we infer from [25] that it is used to indicate device ID. However, the 4-bit length is far from enough since IoT is expected to accommodate billions of connected devices in the future. On the other hand, TR is a heavy overhead, for example, its size is up to 26 symbols while the size of data packet is only 50 symbols (in LoRa packet radios with payload size of 20 bytes). Additionally, in this protocol, a downlink transmission must be requested by an uplink receive request, which increases the probability of collisions in the access procedure. The synchronization mechanism of DQ-N is also absent, but strict timing mechanism is important since several events occur within a frame.

III. LoRa/LoRaWAN Overview

A. LoRa Physical Layer

LoRa is a PHY layer modulation technique, which makes the communication robust against noise and interference with the usage of CCS and forward error correction (FEC). It provides several customizable parameters, including spreading factor (SF), BW, code rate (CR), and transmission power (TP). For each LoRa transmission, the available bit rate, resilience against interference, and the ease of decoding can be changed by tuning these parameters. Three typical BWs (125-, 250-, and 500 KHz) in the unlicensed frequency bands (433-, 868-, or 915 MHz) are available, as well as six different SFs (from 7 to 12) and four different CRs (4/5, 4/6, 4/7, and 4/8). The higher SF implies lower bit rate but longer communication distance, while the higher CR indicates higher protection against interference and vice versa.

LoRa PHY packet structure includes a preamble, an optional header, a data payload, and an optional CRC field. The preamble, whose length can be configured from 10 to 65 536 symbols, is used to synchronize the receiver with the transmitter. The optional header contains the payload length, the CR used for payload, and the header CRC. The presence of the header depends on the header mode (explicit or implicit mode), and it is disabled in the implicit mode when the payload length, CR and CRC presence are fixed or known in advance. More details of LoRa can be found in [8].

B. LoRaWAN

LoRaWAN is an MAC layer protocol and network architecture to be used with the LoRa PHY layer. As described in LoRaWAN specification, devices access the channel randomly in a pure Aloha fashion and must obey a strict duty cycle. Utilizing the Aloha MAC protocol can be justified if to account for the hidden node problem and omitting listen-before-talk phase for energy savings [26].

An LoRaWAN network has a star-of-star topology, in which communications occur between three kinds of components: 1) ED; 2) GW; and 3) network server (NS). GW transparently relays the messages between EDs and NS. Communication is bidirectional, yet uplink communication is strongly preferred. Note that as GW receives every message from EDs within its communication range, collisions occur when it receives more than one message in the same channel at the same time.

Three types of EDs are defined by LoRaWAN [8].

1) *Class A:* Devices in this class open two short receive windows following each uplink transmission. Downlink transmissions are only allowed after uplink messages.

Class A has the lowest power consumption and can be implemented by every device.

2) *Class B:* The purpose of this class is to have a device available for reception at scheduled times (ping slots), in addition to the reception windows in Class A. It requires GW to broadcast a beacon (BCN) periodically to synchronize all the devices in the network. Class B has the medium power consumption.
3) *Class C:* Devices in this class have nearly continuously open receive windows, which can only be closed when transmitting. Hence, Class C has the maximum power consumption.

Note that all the EDs start and join the network in Class A, after that any ED can switch to Class B by searching for a BCN signal. If a network BCN was found, it switches to Class B successfully, otherwise, it fails. Once in Class B mode, the ED should periodically receive BCN and inform its location to NS. If no BCN has been received for a given period, the ED cannot synchronize with the network and it switches back to Class A.

IV. DQ-Based MAC Protocol for LoRa

A. DQ for Contention Resolution

DQ is an MAC protocol based on a tree-splitting algorithm, whose performance is independent of the number of devices sharing a common channel [4]. Different from Aloha, DQ mechanism requires active devices (i.e., the devices which have data to be transmitted) to contend in the contention slots before collision-free data transmission (DT). It uses a set of rules are to be organize each device into one of the two logical queues: 1) collision resolution queue (CRQ) and 2) DT queue (DTQ). If collisions are detected, the devices are split into groups and organized into the CRQ, otherwise, they are organized into the DTQ. The devices in the CRQ wait for contention resolution in the subsequent contention slots while those in the DTQ wait for collision-free DTs. The splitting of the colliding devices reduces the probability of collisions in the later access procedures since the simultaneous attempts are decreased.

The queues are distributed in the sense that each device uses internal counters to represent the queue length and its position within the queue [23]. In order to update the internal counters, the coordinator must provide feedback information, which includes the status of each contention slot as well as the lengths of CRQ and DTQ. There are three states of each contention slot: 1) empty (no device contends); 2) success (only one device contends); and 3) collision (two or more devices contend). Upon receiving the feedback information, the device checks the state of the slot it chooses. If it is a success, the device enter the DTQ, otherwise, it enters the CRQ. At the same time it updates the internal counters as follows [3].

1) *RQ Counter (CRQ Length):* The counter is increased by one for each contention slot with collision state in the previous frame. At each frame, the counter is decreased by one for the resolution attempt of the devices at the head of the CRQ.
2) *pRQ Counter (Device Position in the CRQ):* If the device is waiting in the CRQ, it must decrease the counter by one at each frame. In addition, if the device has attempted an access in the previous frame and collided, it sets the counter to the end of CRQ.
3) *TQ Counter (DTQ Length):* The counter is increased by one for each success state accounted in the previous frame. After a DT, the counter is decreased by one.
4) *pTQ (Device Position in the DTQ):* If a device enters the DTQ, it sets the counter to the end of DTQ. If the device is waiting in the DTQ, it must decrease the counter by one every frame.

We introduce the DQ algorithm into the MAC protocol over the LoRa PHY layer, named DQ-LoRa. Similar to the operation of DQ, this protocol requires EDs to contend the shared channel by sending access requests before transmitting the data packets. GW will act as a coordinator to broadcast the feedback information, then EDs can distributedly execute the rules of DQ-LoRa and autonomously decide when to transmit.

At first, we make some assumptions.

1) Every ED has no more than a single packet to transmit in one BCN period and they arrive at the beginning of this period.
2) Each BCN period is fixed and all the collisions occurring in one BCN period must be resolved before the next period.
3) The length of each data slot (packet) is fixed.
4) There are no transmission errors in the data slot.
5) The number of contention slots is no less than 2.

B. Frame Structure

The general frame structure of DQ-LoRa is composed of three parts in the time domain: 1) a contention window (CW) consisting of m minislots (contention slots) for contention resolution; 2) a data slot for collision-free DT; and 3) a feedback slot for downlink feedback packet (FBP) broadcasting. Apart from that a guard time—inter slot space (ISS), is reserved for the compensation of switching time between reception and transmission. It is worth to mention that the data slot is optional and it can be disabled in any frame when the DTQ is empty. Hereafter, we call the frame which contains a data slot "Type-I frame" (Fig. 1, upper side), and the frame without a data slot "Type-II frame" (Fig. 1, lower side).

In order to accommodate the frame structure to the LoRa standard, the length of each part can be set as follows.

1) *CW Length:* In order to significantly lower the overhead and inspired by the small preamble used in LTE [27], we design the access request as a short preamble of 2 symbols, referred to as random access preamble (RAP). If SF $= 12$, there exits 2^{24} RAPs that can be allocated to different devices. Therefore, the length of each minislot must be at least two symbols and the length of CW is $2m$ symbols. Note that the number of minislots m affects system performance. The optimal setting of m in CW depends on the optimization objective (maximum throughput or minimum energy consumption) rather than the number of EDs, which will be demonstrated in

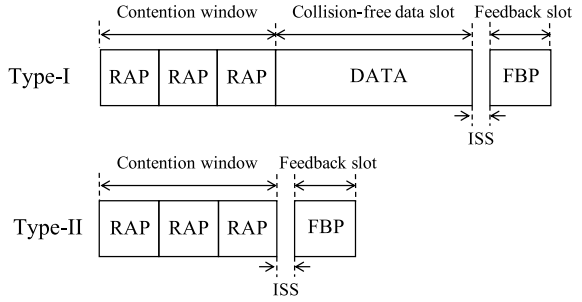


Fig. 1. Frame structures defined in DQ-LoRa. The upper side is Type-I frame and the lower side is Type-II frame.

Section V. Therefore, a tradeoff between the throughput and energy consumption must be considered when setting m.

2) *Data Slot Length:* The length of the data packet is the sum of preamble and PHY payload, and the number of payload symbols can be calculated as [28]

$$\max\left(\left\lceil \frac{8\mathrm{PL}-4\mathrm{SF}+28+16\mathrm{CRC}-20\mathrm{IH}}{4(\mathrm{SF}-2\mathrm{DE})}\right\rceil(\mathrm{CR}+4),0\right)$$

where PL and SF are the size of payload (in bytes) and the SF, respectively, IH indicates the type of header mode (1 for implicit mode and 0 for explicit mode), DE indicates the usage of low data rate optimization (1 when used and 0 when disabled), CRC indicates the presence of the payload CRC (1 when on and 0 when off), CR is the programmed CR from 1 to 4. Considering a data packet with explicit header mode, CR of 4/8, and payload size of 20 bytes, the length of the data packet is 50 symbols (10-symbol preamble and 40-symbol payload).

3) *Feedback Slot Length:* Each FBP must have $2m$ bits to indicate the results of m minislots since each slot has three kinds of states. Apart from that the additional 4 bytes are required to indicate the length of CRQ and DTQ [6]. We consider FBP with implicit header mode to reduce the overhead and transmission time, then the length of FBP is

$$L_{\mathrm{FBP}}=18+8\left\lceil\frac{\lceil 0.25m+4\rceil-5}{6}\right\rceil. \tag{1}$$

It is known from (1) that L_{FBP} is a step function of m and the step height is eight symbols. More intuitively, the length of FBP is 18 symbols when $m<5$, then it rises up to 26 symbols when $5\le m<29$ and further to 34 symbols when $m=29$. The great step height will account for the steep drops of system throughput or the steep rises of energy consumption at some m, which will be exhibited in Section V.

C. Access Procedure

Fig. 2 shows an example to illustrate the access procedure of DQ-LoRa, where three minislots are configured in the CW and three active devices arrive in the current BCN period. Colliding EDs enter the CRQ using a tree-splitting algorithm to resolve the contention, whereas, succeeding EDs enter the DTQ using a first-in-first-out (FIFO) principle to arrange their DT.

TABLE I
COUNTER VALUES

Step[a]	Counter	ED1	ED2	ED3
Frame 1	RQ	1	1	1
	pRQ	0	1	1
	TQ	1	1	1
	pTQ	1	0	0
Frame 2	RQ	0	0	0
	pRQ	0	0	0
	TQ	2	2	2
	pTQ	0	2	1
Frame 3	RQ	0	0	0
	pRQ	0	0	0
	TQ	1	1	1
	pTQ	0	1	0
Frame 4	RQ	0	0	0
	pRQ	0	0	0
	TQ	0	0	0
	pTQ	0	0	0

[a] The steps are shown in Fig. 2

It is worth to mention that the BCN period is variable since a new BCN period can start immediately after all active devices in one BCN period completing their DTs. However, the BCN period is not unlimited, which is restricted by a minimum and a maximum thresholds. If there is no contending device in one BCN period, another period will start after the elapse of the minimum threshold. If there is a large number of contending devices and collisions have not been resolved until the elapse of the maximum threshold, another BCN period will start immediately and the collided devices will recontend in that period. In order to simplify the performance analysis, we make the assumption that each BCN period is fixed (Section IV-A). The performance analysis under variable BCN period is left as a future work.

As we mentioned in the previous section, the length of CW (which relies on the number of minislots) is independent of the number of EDs, but a tradeoff between the throughput and energy consumption must be considered when setting it. Besides, the configuration of minislots can either be adaptable or determined. If it is adaptable, it can be broadcasted by the GW at the beginning of each BCN period, otherwise, it can be optimized offline and preconfigured in the EDs.

The access procedure (Fig. 2) is as follows, and the counter values of each device in each step are shown in Table I.

1) At the beginning of the BCN period, the GW broadcasts a BCN for synchronization, the devices which receive the BCN switch to or maintain in Class B mode.
2) At frame 1, three active EDs (in Class B mode) contend in the CW. ED1 succeed in the second minislot and enter in the first position of DTQ; ED2 and ED3 collide in the third minislot and enter in the first position of CRQ. Note that there is no data slot in this frame because the DTQ is always empty before the first FBP broadcast.
3) At frame 2, ED2 and ED3 contend in the CW as they are in the first position of CRQ. They succeed in the

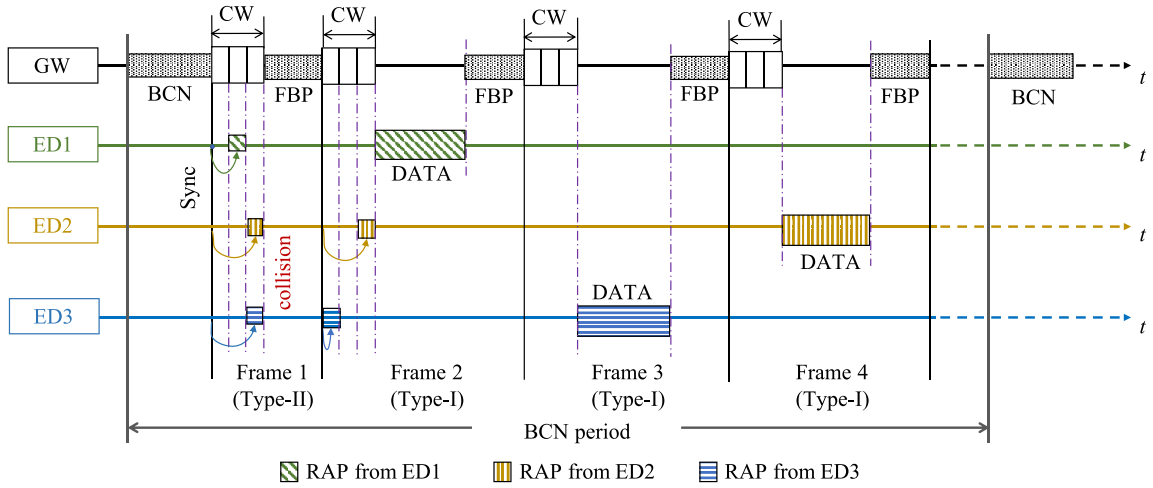


Fig. 2. Example to illustrate the access procedure of DQ-LoRa. Three minislots are configured in the CW and three active EDs arrive in this BCN period.

third and first minislots, respectively, thus they enter at the end of the DTQ according to the order of minislots. Then, ED1 transmits its data in the data slot since it is at the head of the DTQ. After that the GW broadcasts an FBP.

4) At frame 3, no EDs contend in the CW since the CRQ is empty. Subsequently, ED3 at the head of the DTQ transmits its data, and an FBP from the GW follows.
5) At frame 4, the CW is still empty, followed by the DT of ED2, then the FBP from the GW.

It is worth to mention that the influence of propagation delay on the collision detection in minislots is negligible The reason is that LoRa enables the long-range transmissions between 5 and 15 km [7], hence the maximum delay of the edge device is about $(15 \times 10^3)/(3 \times 10^8) = 50$ μs, which is much less than the minislot duration 96 ms (when the bit rate is 250 b/s).

V. Performance Evaluation

We evaluate the performance of DQ-LoRa by the following metrics.

1) *System Throughput:* The time occupation of the data slots in a BCN period.
2) *Average Delay:* The average time required for an ED to complete its DT from its first contention.
3) *Average Energy Consumption:* The average energy consumed by an ED in a BCN period.

A. System Throughput

As we defined before, the system throughput can be represented as

$$S = \frac{nT_{\text{DATA}}}{T_{\text{overhead}} + nT_{\text{DATA}}} \tag{2}$$

where T_{DATA} is the transmission time of a data packet, n is the number of active devices (i.e., the number of Type-I frames), and

$$T_{\text{overhead}} = T_{\text{BCN}} + (n + n')(mT_{\text{RAP}} + T_{\text{FBP}}) \tag{3}$$

TABLE II
Evaluation of n' From Numerical Experiments

m \ n	100	1000	10000
2	50 (50%)	453 (45.30%)	4451 (44.51%)
3	13 (13%)	110 (11%)	1017 (10.17%)
4	7 (7%)	67 (6.70%)	497 (4.97%)
10	2 (2%)	12 (1.20%)	112 (1.12%)
20	2 (2%)	2 (0.20%)	22 (0.22%)
50	1 (1%)	2 (0.20%)	2 (0.02%)

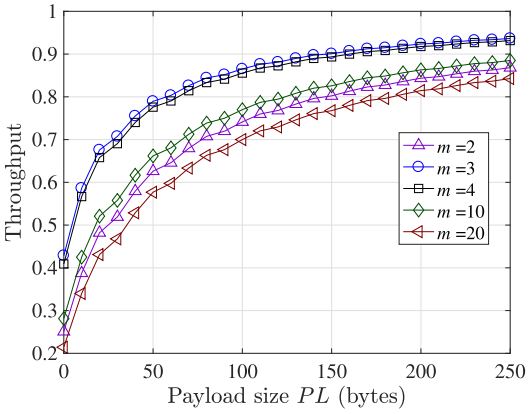
[a] The values in parentheses are n'/n

where T_{BCN}, T_{RAP}, and T_{FBP} are the transmission time of BCN, RAP, and FBP, respectively, and n' is the number of Type-II frames in one BCN period.

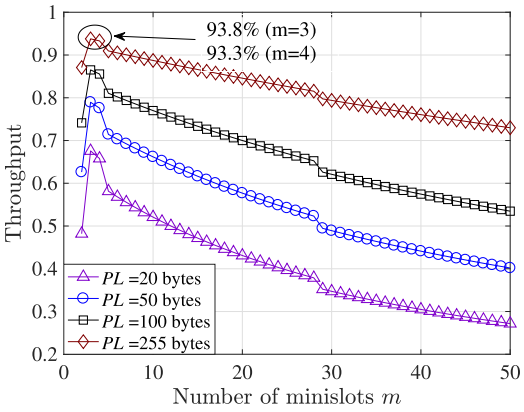
1) Simplified Expression for S: Since the access procedure is a stochastic process, the type of current frame depends on the number of successful minislots in all the previous frames. Therefore, it is cumbersome to obtain an accurate closed-form expression for the number of Type-II frames n'.

As the authors demonstrated in [29], the speed of contention resolution is faster than the speed of DT when the number of minislots is no less than 3, which indicates that $n' < n$ when $m \geq 3$. According to the access procedure, we surmise that n' is much less than n when n is large. In addition to that the traditional LoRaWAN can operate well under the low traffic loads and only suffers from congestion as the traffic load increases [3], [10], and [11]. Van den Abeele *et al.* [11] showed that when the number of EDs is greater than 100 the packet delivery ratio declines significantly. For that we evaluate n' for $n \geq 100$ in the numerical experiments, and show it in Table II.

It can be observed from Table II that: when $m = 2$, n' is about 50% of n and it is much larger than n' in other cases; 2) when $m = 3$, n' is about 10% of n under the condition of $n = 10\,000$; 3) when $m \geq 4$, n'/n is less than 10%; 4) the portion of n' is reduced when n increases. Therefore, we can assume that: 1) $n + n' \approx n$ under the condition of $m \geq 4$, or the condition of $m = 3$ and $n \geq 10\,000$ and 2) $n + n' \approx 1.5n$



(a)



(b)

Fig. 3. Throughput performance of DQ-LoRa for any n. Throughput versus (a) payload size and (b) number of minislots.

when $m = 2$. Hence, for $n \geq 100$, (2) can be simplified to

$$S = \begin{cases} \frac{T_{\text{DATA}}}{3T_{\text{RAP}}+1.5T_{\text{FBP}}+T_{\text{DATA}}}, & m = 2 \\ \frac{T_{\text{DATA}}}{mT_{\text{RAP}}+T_{\text{FBP}}+T_{\text{DATA}}}, & m \geq 3. \end{cases} \tag{4}$$

The error introduced by the approximation of $n' + n$ is that the throughput calculated by (4) is greater than the actual value. Nevertheless, the error is reduced when n increases. Besides, for $m = 2$ and $m = 3$, the throughput calculated by (4) can be used as a reference, but it is less accurate. Later in this section, we focus on analyzing the numerical results for $m \geq 4$.

2) Numerical Results: The impacts of the payload size and the number of minislots on the system throughput are shown in Fig. 3. It can be observed from Fig. 3(a) that the throughput performance is improved when the payload size increases. The trend of improvement becomes flat when the payload size is larger than 150 bytes. In Fig. 3(b), it can be seen that the maximum throughput occurs at $m = 3$ and the throughput at $m = 4$ is slightly lower. Additionally, the throughput experiences degradation when $m \geq 4$. One explanation for the degradation is that although increasing the number of minislots speeds up contention resolution, it also burdens the overhead, thus lowers the system throughput. Note that the steep drops at $m = 5$ and $m = 29$ are caused by L_{FBP}, which has been explained in the previous section.

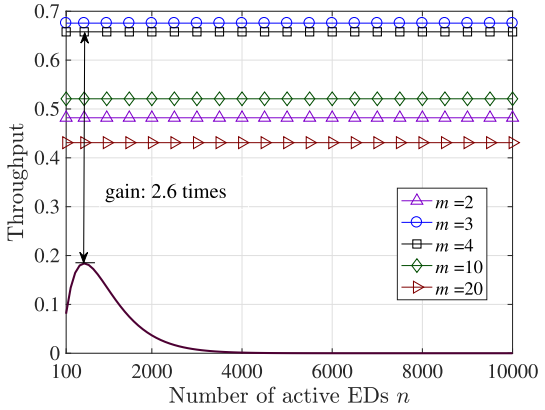


Fig. 4. Throughput comparison between DQ-LoRa and pure Aloha, with the payload size of 20 bytes.

It is well know that the throughput of pure Aloha can be expressed as $S_{PA} = G\exp(-2G)$, where $G = \lambda T_{\text{frame}}$ is the traffic load. Here, $\lambda = n/T$ is the input rate, T and T_{frame} are, respectively, the BCN period and frame duration.

Fig. 4 compares the system throughput of DQ-LoRa with pure Aloha. From Fig. 4, the throughput of DQ-LoRa is independent of the number of EDs due to the simplified expression (4). The result is consistent with the conclusion made in [5] and [6]—DQ ensures high throughput regardless of the traffic load. Besides, the throughput performance of DQ-LoRa is always greater than that of pure Aloha. When $m = 4$, the minimum throughput gain is up to 2.6 times.

B. Average Delay

The delay model of DQ-LoRa is similar to that of DQ [5], which is composed of two queue subsystems: 1) collision resolution (CR) subsystem and 2) DT subsystem. Therefore, the average delay ($\overline{T}_d$) for an active ED to complete its DT can be broken down into two terms: 1) the delay of CR subsystem ($\overline{T}_{\text{CR}}$) and 2) the delay of DT subsystem ($\overline{T}_{\text{DT}}$), that is,

$$\overline{T}_d = \overline{T}_{\text{CR}} + \overline{T}_{\text{DT}}. \tag{5}$$

1) Delay of CR Subsystem: We assume that the arrival process of active EDs is memoryless and the input rate λ is Poisson distributed. Since $\lambda = n/T$, the higher input rate implies larger number of active devices in a BCN period.

The probability of an arrival device finding a free minislot to access in a frame, $P(\lambda)$, has been derived in [16]. However, the expression of $P(\lambda)$ is inaccurate due to the wrong calculation of the probability that n_f devices contend in a given frame. We correct it as follows:

$$\begin{aligned} P(\lambda) &= \sum_{n_f=0}^{\infty} P\big(\text{success}|k = n_f\big)P\big(k = n_f\big) \\ &= e^{-\lambda} + \sum_{n_f=1}^{\infty} \frac{\lambda^{n_f}}{n_f!} e^{-\lambda} m \left(\frac{1}{m}\right)\left(1 - \frac{1}{m}\right)^{n_f-1} \\ &= \frac{\text{me}^{-\lambda/m} - e^{-\lambda}}{m-1} \end{aligned} \tag{6}$$

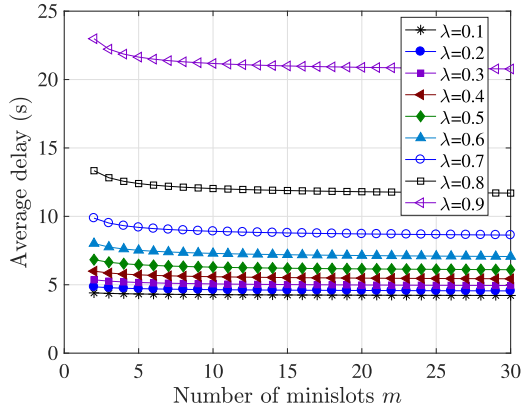


Fig. 5. Impact of the number of minislots on the system throughput, under different input rates.

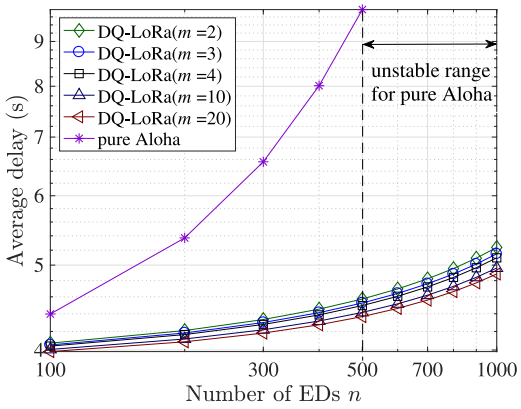


Fig. 6. Delay comparison of DQ-LoRa and pure Aloha.

where $P(\text{success}|k = n_f)$ is the probability of an active device randomly choosing a free minislot when there are n_f active devices in a given frame, and $P(k = n_f)$ is the probability that n_f active devices contend in that frame.

According to [16], the service time of CR subsystem is a Poisson distributed random variable with mean

$$\bar{t} = \left[\ln\left(\frac{1}{1 - P(\lambda)}\right)\right]^{-1}. \tag{7}$$

We can therefore model the system as an M/M/1. Thus, the total delay in CR subsystem is [30]

$$\overline{T}_{\text{CR}} = \frac{1}{\mu_{\text{CR}} - \lambda} = \left[\ln\left(\frac{1}{1 - P(\lambda)}\right) - \lambda\right]^{-1} \tag{8}$$

where $\mu_{\text{CR}} = \ln[1/(1 - p(\lambda))]$ is the service rate of CR subsystem.

2) Delay of DT Subsystem: Since both arrival and service time processes are Poisson distributed, CR subsystem output traffic pattern will also be Poisson distributed with the same rate λ as the input traffic [5]. This output traffic is directly the input traffic of DT subsystem, hence the input rate of DT subsystem is also λ.

As the service time for every ED in the DTQ is one frame, DT subsystem can be modeled as an M/D/1, with a service rate $\mu_{\text{DT}} = 1$. We can then obtain $\overline{T}_{\text{DT}}$ immediately

$$\overline{T}_{\text{DT}} = \frac{1}{\lambda}\left[\rho_{\text{DT}} + \frac{\rho_{\text{DT}}^2}{2(1 - \rho_{\text{DT}})}\right] = \frac{2 - \lambda}{2(1 - \lambda)} \tag{9}$$

where $\rho_{\text{DT}} = \lambda/\mu_{\text{DT}}$. Substituting (8) and (9) to (5), the average delay can be obtained and it is in the unit of T_{frame}.

3) Numerical Results: The curves of average delay against the number of minislots under different input rates are exhibited in Fig. 5. As we expected, the higher input rate results in longer delay. However, under the same input rate, the number of minislots has little influence on the delay and the delay is almost constant when the number of minislots is larger than 10. One explanation is that the duration of minislots, compared to the data slot, only covers a small portion of the frame duration.

The average delay of pure Aloha, according to [31], can be expressed as $\overline{T}_{PA} = \exp(2G)T_{\text{frame}}$. As we known, the total traffic load G in a pure Aloha system consists of two parts: 1) the new arrival packets and 2) the retransmitted packets. For the Aloha access in LoRaWAN, however, it is limited to the duty cycle regulations (e.g., 1% in EU 868), which means if an ED transmits a packet and collides on the sub-band 868 MHz, it must wait for an off-period ($99T_{\text{frame}}$) to retransmit. Therefore, the collided packets in one duty period do not arrive immediately in this period, and the total traffic can be considered to comprise only the new arrival packets, that is, $G = nT_{\text{frame}}/T$. Besides, it is worth to mention that the expression for delay calculation is only applicable when the system is stable (i.e., $G \leq 0.5$), because the delay performance degrade dramatically when $G > 0.5$ [31].

The delay comparison between DQ-LoRa and pure Aloha is shown in Fig. 6. Note that for the reasons we mentioned above, the delay curve of pure Aloha is only plotted for $n \leq 500$ (i.e., $G \leq 0.5$). Fig. 6 exhibits that the delay of pure Aloha increases exponentially with n, while that of DQ-LoRa increases slowly and keeps at a low level. When $n = 500$, the average delay of DQ-LoRa is about 54% lower than that of pure Aloha. The results indicate that DQ-LoRa provides significant improvement of delay performance comparing to pure Aloha.

C. Energy Consumption

From the perspective of an active device, the duration of a BCN period can be represented as [6]

$$\overline{T} = T_{\text{BCN}} + \overline{T}_{\text{active}} + \overline{T}_{\text{sleep}} \tag{10}$$

where T_{BCN} is the transmission time of a BCN, $\overline{T}_{\text{active}}$ is the average time elapsed since a BCN round starts (after the BCN transmission) until an active ED successfully transmit its data, and $\overline{T}_{\text{sleep}}$ is the average time from the end of the successful transmission to the end of the BCN round. The middle part can be expressed as

$$\begin{aligned}\overline{T}_{\text{active}} = {} & \left(\overline{\text{CR}}_{\text{RAP}} + \overline{\text{CR}}_{\text{sleep}}\right)T_{\text{frame}} \\ & + \left(\overline{\text{DT}}_{\text{DATA}} + \overline{\text{DT}}_{\text{sleep}}\right)T_{\text{frame}}\end{aligned} \tag{11}$$

where $\overline{\text{CR}}_{\text{RAP}}$ is the average number of frames, where the ED contends in the CW, $\overline{\text{CR}}_{\text{sleep}}$ is the average number of frames, where the ED is sleeping in the CRQ waiting for the next

contending, $\overline{\mathrm{DT}}_{\mathrm{sleep}}$ is the average number of frames, where the ED is sleeping in the DTQ waiting for its DT, $\overline{\mathrm{DT}}_{\mathrm{DATA}}$ is the average number of frames, where the ED is in the first position of DTQ to transmit data, and the frame duration $T_{\mathrm{frame}} = mT_{\mathrm{RAP}} + T_{\mathrm{DATA}} + T_{\mathrm{FBP}}$.

1) Expression of Average Energy Consumption: The average energy consumed by an ED in a BCN period can be represented as

$$\overline{E} = E_{\mathrm{BCN}} + \overline{E}_{\mathrm{active}} + \rho_{\mathrm{sleep}}\overline{T}_{\mathrm{sleep}} \tag{12}$$

where E_{BCN} is the energy consumption during T_{BCN}, $\overline{E}_{\mathrm{active}}$ is the average energy consumption during $\overline{T}_{\mathrm{active}}$, and ρ_{sleep} is the power consumption in the sleeping mode. The middle part can be further expressed as

$$\begin{aligned}\overline{E}_{\mathrm{active}} = \overline{\mathrm{CR}}_{\mathrm{RAP}}E_{\mathrm{RAP}} + \left(\overline{\mathrm{CR}}_{\mathrm{sleep}} + \overline{\mathrm{DT}}_{\mathrm{sleep}}\right) \\ \times \rho_{\mathrm{sleep}}T_{\mathrm{frame}} + \overline{\mathrm{DT}}_{\mathrm{DATA}}E_{\mathrm{DATA}}\end{aligned} \tag{13}$$

where E_{RAP} and E_{DATA} are the energy consumed by the ED in a frame where it contends and transmits data, respectively.

At the beginning of each BCN period, every active ED listens to the channel to receive the BCN from GW, hence

$$E_{\mathrm{BCN}} = \rho_{rx}T_{\mathrm{BCN}} \tag{14}$$

where ρ_{rx} is the power consumption in the receiving mode.

When the ED contends in a frame, it executes the following actions: 1) transmits an RAP in one of m minislots, and keeps in the standby mode in the other $m-1$ minislots; 2) keeps in the standby mode in the data slot; and 3) receives FBP in the feedback slot. Therefore, E_{RAP} can be represented as

$$E_{\mathrm{RAP}} = \left(\rho_{tx} + (m-1)\rho_{\mathrm{stby}}\right)T_{\mathrm{RAP}} + \rho_{\mathrm{stby}}T_{\mathrm{DATA}} + \rho_{rx}T_{\mathrm{FBP}} \tag{15}$$

where ρ_{tx} and ρ_{stby} are the power consumption in the transmitting and standby modes, respectively.

When the ED transmits data in a frame, it executes the following operations: 1) remains in the standby mode in the m minislots; 2) transmits data in the data slot; and 3) receives FBP in the feedback slot. Then, E_{DATA} can be expressed as

$$E_{\mathrm{DATA}} = m\rho_{\mathrm{stby}}T_{\mathrm{RAP}} + \rho_{tx}T_{\mathrm{DATA}} + \rho_{rx}T_{\mathrm{FBP}}. \tag{16}$$

Since we assume that there are no transmission errors, $\overline{\mathrm{DT}}_{\mathrm{DATA}} = 1$. Additionally, $\overline{\mathrm{CR}}_{\mathrm{RAP}}$ has already been derived in [32], that is,

$$\overline{d_n} = \log_m(n-1) + \left(0.5 + \frac{\gamma}{\ln m}\right) + \frac{1}{2n\ln m} \tag{17}$$

where $\gamma \approx 0.5772$ is Eulers constant. Then, substituting (13)–(17) into (12)

$$\begin{aligned}\overline{E} = E_{\mathrm{BCN}} + \overline{d_n}E_{\mathrm{ARS}} + E_{\mathrm{DATA}} + \left(\overline{\mathrm{CR}}_{\mathrm{sleep}} + \overline{\mathrm{DT}}_{\mathrm{sleep}}\right) \\ \times \rho_{\mathrm{sleep}}T_{\mathrm{frame}} + \rho_{\mathrm{sleep}}\overline{T}_{\mathrm{sleep}}.\end{aligned} \tag{18}$$

TABLE III
SYSTEM PARAMETERS

Packet Size	symbol	Power Consumption	mW
BCN	25	ρ_{tx}	94.2
RAP	2	ρ_{rx}	34.65
FBP	$18 + x^{\mathrm{a}}$	ρ_{stby}	4.62
Data	154	ρ_{sleep}	0.33

[a] $x = 8\lceil(\lceil 0.25m + 4\rceil - 5)/6\rceil$

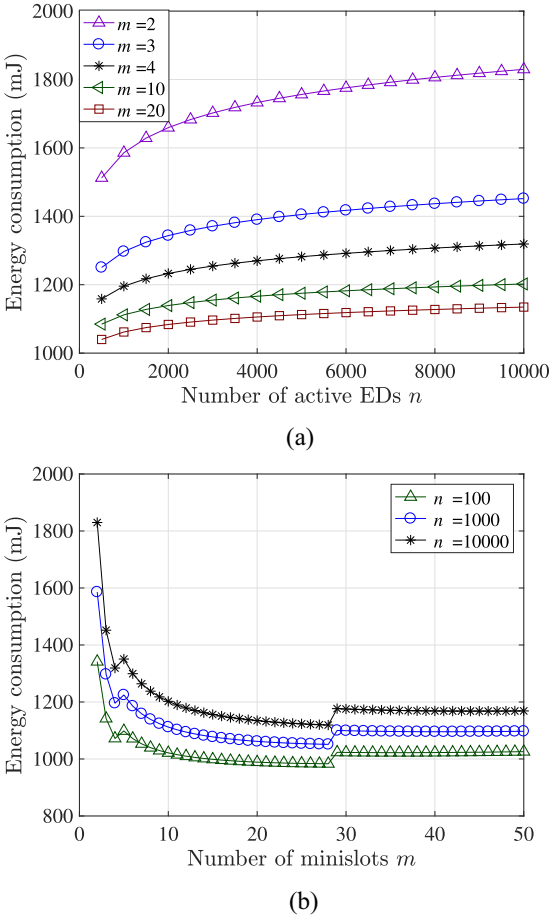


Fig. 7. Impacts of the numbers of active EDs and minislots on average energy consumption, with payload size of 100 bytes. Energy consumption versus (a) number of active EDs and (b) number of minislots.

2) Numerical Results: System parameters are summarized in Table III, which are selected according to the LoRa SX1272 data sheet [28]. The number of active EDs and the number of minislots impact the average energy consumption, which is shown in Fig. 7. The energy consumption is growing with the increased number of active devices, in Fig. 7(a), but the growth trend tends to be flat when the number of active devices is more than 6000. On the other hand, the energy consumption does not vary monotonically with the number of minislots, shown in Fig. 7(b). There exists an optimal setting, $m = 28$, to achieve the minimum energy consumption and it is independent of the number of EDs. Additionally, the energy consumption increases exponentially up to a large value when m decreases to less than 3. Note that there exist steep rises at $m = 5$ and $m = 29$, and the reason has been explained before.

TABLE IV
SYSTEM PERFORMANCE COMPARISON

Target[a]: 10^6 packets		
Protocol	Energy consumption (mJ)	TTR[b]
pure Aloha	2273	6.94
CSMA	432	1.3
DQ-LoRa	1185	1

[a] Number of successful packets received by the gateway.

[b] Transmit to Target Ratio (TTR): the ratio of the total number of packets transmitted to achieve the target.

It is worth to mention that according to Section V-A, the optimum setting of minislots is $m = 4$, which is not consistent with the optimum setting here. Therefore, a tradeoff between the throughput and energy consumption must be considered while optimizing the number of minislots.

System performance comparisons among pure Aloha, CSMA, and DQ-LoRa are listed in Table IV. Note that the energy consumptions and transmit to target ratios (TTRs) of pure Aloha and CSMA were measured in [12] and the performance of DQ-LoRa is calculated under the condition of $m = 4$. TTR refers to transmit to target ratio, which is defined in [12] as the ratio of the total number of packets transmitted to achieve the target. Hence, the closer the value of TTR to 1, the higher the reliability and efficiency of one protocol and vice versa.

We can observe from Table IV that DQ-LoRa consumes about 48% lower energy than pure Aloha, whereas it consumes more energy than CSMA. However, the TTR of DQ-LoRa equals to 1 while the TTRs of pure Aloha and CSMA are higher than 1. Therefore, DQ-LoRa is more reliable than pure Aloha and CSMA. For CSMA, its TTR is 0.3 times higher than DQ-LoRa, which means 30% of packets are retransmitted. This indicates that the throughput and delay performance of CSMA would be inferior to DQ-LoRa.

VI. CONCLUSION

LoRa is one of the most successful and adopted LPWAN technologies, however, the Aloha-based MAC protocol used by LoRaWAN limits the scalability of LoRa networks. In this article, we develop a DQ-based MAC protocol over LoRa, referred to as DQ-LoRa, in order to improve the scalability and reliability of LoRa networks. The details of DQ-LoRa are provided, including the frame structure and access procedure. To evaluate its performance, we theoretically analyze the system throughput, average delay, and average energy consumption. Results show the following.

1) The throughput of DQ-LoRa outperforms pure Aloha and it is independent of the traffic load. With four minislots, the minimum throughput gain can be up to 2.6 times.
2) The delay performance of DQ-LoRa is insensitive to the number of minislots. It is also superior to pure Aloha, with the delay savings up to 54% in the case of 500 active devices.
3) DQ-LoRa is more energy efficient than pure Aloha and it provides energy savings up to 48% when the number of transmitted packets is large.
4) The optimal setting of the number of minislots contributes to an optimum frame length. For the maximum throughput the optimal setting is 4 (the frame length is 76 symbols), while for the minimum energy consumption it is 28 (the frame length is 132 symbols). These two optimal values are not coincident, thus a tradeoff between the system throughput and energy efficiency must be considered when setting the number of minislots.

In order to reduce the overhead, as we mentioned before, the design of RAP is a short preamble rather than a complete chirp signal. It, therefore, can be generated directly in the PHY layer and the design of optimum minislot duration is one of the future directions. Additionally, we only focus on the performance analysis under fixed BCN period and leave the performance analysis under variable BCN period as a future work. Furthermore, the joint optimization problem of throughput and energy consumption is also a future direction.

REFERENCES

[1] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, and T. Watteyne, "Understanding the limits of LoRaWAN," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 34–40, Sep. 2017.

[2] M. O. Farooq and D. Pesch, "Analyzing LoRa: A use case perspective," in *Proc. IEEE 4th World Forum Internet Things (WF-IoT)*, Singapore, 2018, pp. 355–360.

[3] A. Laya, C. Kalalas, F. Vazquez-Gallego, L. Alonso, and J. Alonso-Zarate, "Goodbye, ALOHA!" *IEEE Access*, vol. 4, pp. 2029–2044, 2016.

[4] W. Xu and G. Campbell, "A near perfect stable random access protocol for a broadcast channel," in *Proc. Disc. New World Commun. (SUPERCOMM/ICC)*, vol. 1. Chicago, IL, USA, 1992, pp. 370–374.

[5] X. Zhang and G. Campbell, "Performance analysis of distributed queueing random access protocol—DQRAP," DQRAP Res. Group, Comput. Sci. Dept., Illinois Inst. Technol., Chicago, IL, USA, Rep. 93-1, 1994.

[6] F. Vázquez-Gallego, J. Alonso-Zarate, P. Tuset-Peiró, and L. Alonso, "Energy analysis of a contention tree-based access protocol for machine-to-machine networks with idle-to-saturation traffic transitions," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, 2014, pp. 1094–1099.

[7] K. Zhang and A. Marchiori, "Crowdsourcing low-power wide-area IoT networks," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Kailua-Kona, HI, USA, 2017, pp. 41–49.

[8] *LoRaWANTM Specification, V1.0.2*, LoRa Alliance, Fremont, CA, USA, 2017.

[9] M. C. Bor, U. Roedig, T. Voigt, and J. M. Alonso, "Do LoRa low-power wide-area networks scale?" in *Proc. 19th ACM Int. Conf. Model. Anal. Simulat. Wireless Mobile Syst. (MSWiM)*, 2016, p. 5967.

[10] J. Haxhibeqiri, F. Van den Abeele, I. Moerman, and J. Hoebeke, "LoRa scalability: A simulation model based on interference measurements," *Sensors*, vol. 17, no. 6, p. 1193, 2017.

[11] F. Van den Abeele, J. Haxhibeqiri, I. Moerman, and J. Hoebeke, "Scalability analysis of large-scale LoRaWAN networks in ns-3," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2186–2198, Dec. 2017.

[12] M. O. Farooq and D. Pesch, "A Search into a suitable channel access control protocol for LoRa-based networks," in *Proc. IEEE 43rd Conf. Local Comput. Netw. (LCN)*, Chicago, IL, USA, 2018, pp. 283–286.

[13] C.-T. Wu and G. Campbell, "Extended DQRAP (XDQRAP). A cable TV protocol functioning as a distributed switch," in *Proc. Int. Workshop Community Netw. Integr. Multimedia Services Home*, Jul. 1994, pp. 191–198.

[14] H.-J. Lin and G. Campbell, "PDQRAP-prioritized distributed queueing random access protocol," in *Proc. IEEE Conf. Local Comput. Netw.*, Minneapolis, MN, USA, Oct. 1994, pp. 82–91.

[15] C.-T. Wu and G. Campbell, "Interleaved DQRAP with global TQ," Dept. Comput. Sci., Illinois Inst. Technol., Chicago, IL, USA, Rep. 94-4, 1995.
[16] 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 J. Sel. Areas Commun.*, vol. 18, no. 9, pp. 1701–1718, Sep. 2000.
[17] L. Alonso, R. Ferrus, and R. Agusti, "WLAN throughput improvement via distributed queuing MAC," *IEEE Commun. Lett.*, vol. 9, no. 4, pp. 310–312, Apr. 2005.
[18] J. Alonso-Zarate, E. Kartsakli, A. Cateura, C. Verikoukis, and L. Alonso, "A near-optimum cross-layered distributed queuing protocol for wireless LAN," *IEEE Wireless Commun.*, vol. 15, no. 1, pp. 48–55, Feb. 2008.
[19] E. Kartsakli, A. Cateura, L. Alonso, J. Alonso-Zarate, and C. Verikoukis, "Cross-layer enhancement for WLAN systems with heterogeneous traffic based on DQCA," *IEEE Commun. Mag.*, vol. 46, no. 6, pp. 60–66, Jun. 2008.
[20] B. Otal, L. Alonso, and C. Verikoukis, "Highly reliable energy-saving MAC for wireless body sensor networks in healthcare systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 4, pp. 553–565, May 2009.
[21] J. Alonso-Zárate, E. Kartsakli, L. Alonso, and C. Verikoukis, "Performance analysis of a cluster-based MAC protocol for wireless ad hoc networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2010, no. 1, 2010, Art. no. 625619.
[22] J. Alonso-Zarate, L. Alonso, C. Skianis, and C. Verikoukis, "Analysis of a distributed queuing medium access control protocol for cooperative ARQ," in *Proc. IEEE GLOBECOM*, Miami, FL, USA, Dec. 2010, pp. 1–5.
[23] A. Laya, L. Alonso, and J. Alonso-Zarate, "Contention resolution queues for massive machine type communications in LTE," in *Proc. IEEE 26th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Hong Kong, 2015, pp. 2314–2318.
[24] 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 Mobile Comput.*, vol. 20, pp. 84–99, Jul. 2015.
[25] K. Zhang and A. Marchiori, "Demo Abstract: PlanIt and DQ-N for low-power wide-area networks," in *Proc. IEEE/ACM 2nd Int. Conf. Internet-of-Things Design Implement. (IoTDI)*, Pittsburgh, PA, USA, 2017, pp. 291–292.
[26] J. Petäjäjärvi, K. Mikhaylov, M. Pettissalo, J. Janhunen, and J. Iinatti, "Performance of a low-power wide-area network based on LoRa technology: Doppler robustness, scalability, and coverage," *Int. J. Distrib. Sensor Netw.*, vol. 13, no. 3, pp. 1–16, Mar. 2017.
[27] L. Tello-Oquendo, I. Leyva-Mayorga, V. Pla, J. Martinez-Bauset, and V. Casares-Giner, "Analysis of LTE-A random access procedure: A foundation to propose mechanisms for managing the M2M massive access in wireless cellular networks," in *Proc. Workshop Innov. Inf. Commun. Technol.*, 2015, pp. 95–101.
[28] *Semtech SX1272*. Accessed: Mar. 16, 2018. [Online]. Available: https://www.semtech.com/apps/product.php?pn=SX1272
[29] W. Xu and and G. Campbell, "A distributed queueing random access protocol for a broadcast channel," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 23, no. 4, pp. 270–278, 1993.
[30] L. Kleinrock, *Queueing Systems*. New York, NY, USA: Wiley, 1976.
[31] M. Reyes-Ayala, E. A. Andrade-Gonzlez, J. A. Tirado-Mendez, and H. J. Aguilar, "Average packet delay in random multiple access for satellite systems," in *Proc. Int. Conf. Telecommun.*, 2008, pp. 120–124.
[32] A. J. E. M. Janssen and M. J. de Jong, "Analysis of contention tree algorithms," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2163–2172, Sep. 2000.