RT-LoRa: A Medium Access Strategy to Support Real-Time Flows Over LoRa-Based Networks for Industrial IoT Applications

Luca Leonardi, Filippo Battaglia, and Lucia Lo Bello, *Senior Member, IEEE*

***Abstract*—Low power wide area networks (LPWANs) are suitable for many applications that require low energy consumption, support for a high number of nodes, and large coverage range. Long range (LoRa) is one of the most successful LPWAN technologies, as it enables robust long-distance low power communications and it is proven to be effective in Internet of Things (IoT) applications, such as environmental monitoring and smart metering. LoRa is also promising for Industrial IoT scenarios, but its adoption is impaired by the relevant standardized medium access control protocol, LoRaWAN, that does not offer support to real-time data flows. For this reason, this article proposes RT-LoRa, a medium access strategy for LoRa that provides support for real-time flows, thus enabling the implementation of LoRa-based LPWAN for Industrial IoT applications. This article describes RT-LoRa, presents a simulative assessment in a realistic Industrial IoT scenario and provides some guidelines for the configuration of an RT-LoRa network.**

***Index Terms*—Industrial Internet of Things (IoT), long range (LoRa), low power wide area networks (LPWANs), medium access control (MAC) protocols.**

I. Introduction

LOW POWER wide area networks (LPWANs), such as the ones based on the long range (LoRa) technology represent a novel communication paradigm that will replace or complement traditional cellular and short-range wireless technologies in several applications.

In the Internet of Things (IoT) field, LPWA networks are expected to offer energy-efficient connectivity to a high number of low power devices, distributed over very large geographical areas, that do not require to transmit a large amount of traffic [1]. LoRa offers notable properties, such as long range, low data rates, and low energy consumption. Several IoT applications requiring those properties are found in smart cities, smart metering, fleet/goods tracking, security, and health monitoring [2], [3]. For this reason, industry and academia see LoRa as one of the "rising stars" of LPWAN-based IoT technologies [4], [5]. Moreover, LoRa is suitable for distributed measurement systems [6], i.e., typical IoT applications based on millions of sensors that collect data from the real world. LoRa is an attractive solution also for Industrial IoT, thanks to its high robustness [7]. However, the LoRaWAN [8], [9] medium access control (MAC) protocol for LoRa-based networks[1] is intended for sporadic nontime-constrained communications between low-cost long-lived battery-powered devices. As a result, LoRaWAN adopts an ALOHA-based medium access protocol, that cannot provide bounded latency to the real-time flows typical of industrial applications. Conversely, a centralized approach, in which a central node manages the medium access according to a predefined order, would be more appropriate for the real-time data flows generated from Industrial IoT applications.

For this reason, previous work in [10] proposed Industrial LoRa, a centralized MAC protocol for star topologies working over LoRa that is able to support both real-time and nonreal-time communications for Industrial IoT applications. The work in [10] shows interesting simulation results obtained by Industrial LoRa in a realistic industrial scenario.

Motivation: However, the paper [10] does not explore several aspects, i.e., the way to configure a generic network and to analytically derive the upper bounds on the message latency. Moreover, the pure ALOHA-based medium access strategy for aperiodic messages does not provide sufficient reliability. The RT-LoRa protocol proposed in this article builds upon the Industrial LoRa protocol proposed in [10], but it overcomes the limitations of Industrial LoRa through a number of extensions and improvements that also increase the communication reliability.

Contributions: This article contribution consists of the detailed design of the RT-LoRa MAC scheme itself and the guidelines on setting the network parameters. The guidelines enable the network designer to correctly configure the superframe structure and assess whether a feasible schedule for a given set of flows (i.e., a schedule compliant with the application constraints) can be found. This article also provides a simulative performance assessment in terms of packet loss ratio (PLR) and end-to-end (e2e) delay.

Paper Overview: This article is organized as follows. Section II deals with background and related works. Section III

Manuscript received May 1, 2019; revised July 20, 2019 and September 10, 2019; accepted September 13, 2019. Date of publication September 20, 2019; date of current version December 11, 2019. This work was supported by the University of Catania, through the program "FONDI PER LA RICERCA DI ATENEO - PIANO PER LA RICERCA 2016/2018." *(Corresponding author: Lucia Lo Bello.)*

The authors are with the Department of Electrical, Electronic and Computer Engineering, University of Catania, I-95125 Catania, Italy (e-mail: lobello@unict.it).

Digital Object Identifier 10.1109/JIOT.2019.2942776

[1] In this article, the term LoRa refers to the physical layer, that is based on a proprietary spread spectrum modulation scheme patented by Semtech, while the term LoRaWAN refers to the MAC layer, an open standard promoted by the LoRa Alliance.

outlines the LoRa technology and constraints. Section IV presents the RT-LoRa design, while Section V provides guidelines for setting the superframe in an RT-LoRa network complying with the protocol restrictions. Section VI addresses a simulative assessment of RT-LoRa and discusses the results obtained. Finally, Section VII gives the conclusions and hints for future works.

II. Background and Related Work

A. Background

Among the wireless communication technologies for low power IoT communication, two main categories can be identified, i.e., Low Power Local Area Networks (LPLANs) and Low Power Wide Area Networks (LPWANs) technologies, as discussed in [11].

The LPLANs include several technologies, such as IEEE 802.15.4 and Bluetooth low energy (BLE). They are suitable for short-range personal area networks, body area networks or for larger areas when a mesh topology is used [12]–[15]. In particular, the IEEE 802.15.4 and BLE standards are widely used for wireless sensor networks. The IEEE 802.15.4 defines a physical (PHY) and an MAC layer used in several wireless protocol stacks (such as ZigBee and 6LoWPAN [16]–[18]), while BLE defines a complete protocol stack. BLE is mainly used in indoor human-oriented applications (home entertainment, health monitoring, and personal security) [19], [20], while IEEE 802.15.4 is adopted in industrial and environmental monitoring, security, and process automation [21]–[23].

The LPWANs can be seen as low power competitors of cellular networks, so they are suitable for LoRa applications where each "cell" covers thousands of end-devices. LPWANs include several technologies, such as LoRaWAN, SigFox, and NB-IoT [24]. They provide very low data rates, but support large areas (up to several kilometers), e.g., a smart city. The power consumption of an LoRa transceiver is similar to the one of an IEEE 802.15.4 transceiver [25]. Typical applications for LoRa technologies are smart metering, smart grid and environmental monitoring.

LoRaWAN is considered as one of the most successful LPWAN technologies [26], [27]. Comparing with other technologies, such as NB-IoT and Sigfox, LoRaWAN operates in the license-free spectrum, whereas NB-IoT uses licensed frequency bands. Moreover, LoRaWAN supports a higher bit rate than SigFox. LoRaWAN networks work over a star-of-stars topology and exploit a mechanism that enables multiple end devices (EDs) to communicate with a central network server through gateway nodes. The EDs send messages to gateways through a single-hop LoRa communication using an ALOHA-based MAC mechanism. In turn, the gateways relay data to the central network server. Such a relay operation raises the developer from the need to use a peer-to-peer protocol for IoT that performs routing over subnetworks in the application layer [28]. The LoRa physical layer allows multiple EDs to communicate simultaneously with the same gateway, using different spreading factors (SFs) and channels. The communication is bidirectional, but uplink transmissions from the EDs to the network server are strongly favored. As it was previously mentioned, the LoRaWAN ALOHA-based MAC protocol is not able to support real-time communications. For this reason, this article proposes an alternative medium access protocol on top of LoRa. Further details about LoRaWAN can be found in [8].

B. Related Work

Recent works have dealt with LoRaWAN [8] capabilities [27], [29], performance [11], and parameter setting [1], [30] for indoor industrial monitoring applications. For example, RS-LoRa [31] exploits a *lightweight scheduling* upon LoRaWAN, in which the gateway periodically sends some messages to the EDs, specifying communication parameters, such as the allowed transmission power on each channel. Next, each ED independently determines the channel and *time offset* to use for the deferred transmission. Although the probability of collision between the transmitting EDs is reduced exploiting different channels and other features, RS-LoRa is not able to support timeslots and time-bounded communications, as the time-offset is randomly determined.

In [32], a slotted ALOHA approach for LoRaWAN is proposed, as an alternative to the classic ALOHA, to improve the packet loss and throughput performance. While the approach in [32] proposes a different way to rule transmissions in LoRaWAN, RT-LoRa is not LoRaWAN, but a different MAC protocol that, unlike LoRaWAN, is able to guarantee bandwidth (BW) to support real-time and nonreal-time flows over LoRa. Conversely, the approach in [32] cannot provide bounded latency to real-time flows, as it does not allow to reserve bandwidth.

In [33], the use of a network synchronization and scheduling entity (NSSE) integrated in the LoRaWAN network server is proposed. Each ED synchronizes to the NSSE, sending a request that contains the traffic periodicity. In turn, the NSSE replies with a data structure that contains the timeslot assigned to the node for communicating with the gateway. However, in [33] the ED communication with the gateway use the transaction model defined by LoRaWAN for Class A devices [8]. Consequently, the gateway is unable to start downlink communications before a message from the ED is received. In fact, Class A devices open two receive windows (for downlink) after an uplink transmission.

On-demand LoRa [25] is an MAC-layer, alternative to LoRaWAN, that uses two different time division multiple access (TDMA) strategies, called unicast and broadcast TDMA, respectively. Each ED is equipped with a low power transceiver, compliant to the Wake-up radio standard [34], that is normally maintained in deep listening state. With Unicast TDMA, the gateway sends a wake-up beacon to one specific ED that starts uplink communications. With Broadcast TDMA, the gateway sends a wake-up beacon to multiple EDs that start deferred uplink communications using scheduled timeslots. Although on-demand LoRa is energy-efficient, it requires a nonstandard LoRa transceiver.

In [35], a slightly modified version of the LoRaWAN upper layer that implements a TSCH-like approach was proposed to make the protocol suitable for industrial wireless

networks. There are several differences between RT-LoRa and the approach in [35]. First, the TSCH-like approach works on top of LoRaWAN, while RT-LoRa works on top of LoRa as an alternative option to LoRaWAN. Second, the TSCH-like approach supports real-time flows only, while RT-LoRa both real-time and nonreal-time traffic. Third, the approach in [35] only supports uplink unconfirmed communications, whereas RT-LoRa supports uplink/downlink confirmed/unconfirmed communications, thus being more flexible. Moreover, in [35] each flow transmits only in one sub-band, as each flow is assigned one specific slot with a fixed channel. Conversely, RT-LoRa adopts a frequency rotation mechanism, according to which the slot assigned to each flow has a varying channel, so the flow transmissions occur on different sub-bands. This improves bandwidth exploitation while complying with the duty cycle (DC) constraints that are imposed on each sub-band by the regulations.

C. LoRa for Industrial Applications

Several recent works have addressed LoRa properties for industrial applications. For instance, in [36] the LoRa performance and noise robustness for a specific industrial application was assessed and the best configuration and tradeoff between data rate and packet loss were determined. The work in [7] presents an accurate timing synchronization system for TDMA scheduling implemented on LoRa and shows some experimental results. The main idea consists of using the LoRa physical layer, but without the LoRaWAN MAC layer, in order to obtain a new protocol stack that maximizes the radio performance and complies with all the requirements of industrial IoT. Using this approach LoRa may become a candidate technology for low-bandwidth Industrial IoT applications [37].

In particular, to the best of our knowledge, Industrial LoRa [10] is the first work that proposed a mechanism to provide support for both real-time and nonreal-time communications over LoRa. In Industrial LoRa, the network access is organized in cyclically repeated superframes, consisting of five sections, i.e., the beacon section, the contention access period (CAP), the contention-free period (CFP), the downlink period, and the CFP Ack section. In the beacon section, a beacon synchronizes all the network nodes and indicates the start of the superframe. During the CAP, that is intended for nonreal-time communications, the end nodes compete for channel access using pure ALOHA. The CFP, that is intended for periodic real-time data flows, consists in a set of timeslots and exploits a multichannel and multi-SF TDMA protocol. The downlink period is for the communications from the network sink to the end nodes. In the CFP Ack section, the sink broadcasts one acknowledgement message to confirm the messages received during the CFP. Compared to Industrial LoRa, the RT-LoRa protocol here proposed adds the following features.

1) *Free Mobility:* All mobile nodes (MNs) can move within the coverage range of the sink without any restriction.
2) *Smart Mobile Nodes:* An innovative mechanism based on multiple transmissions of the beacon frames by the sink allows the MNs to know the SF values that are recommended for transmission at a given time.
3) *Smart CAP:* A novel MAC strategy, based on slotted ALOHA with SF selection strategy for aperiodic nonreal-time transmissions, enables more reliable transmissions in the CAP.
4) *Quality of Service (QoS) Classes for the Flows of MNs:* Three classes of QoS for the flows of MNs allow to choose the desired tradeoff between reliability and energy consumption on the basis of the data flow requirements.
5) *Frequency Rotation for CFP Slots:* Each node that schedules a periodic real-time flow and has a timeslot assigned in the CFP performs a frequency rotation at every superframe. This mechanism improves the communication robustness and bandwidth efficiency, while complying with the DC restrictions.

III. LoRa Overview

A. LoRa Physical Layer

The LoRa physical layer [38] enables LoRa, low power communications in the unlicensed sub-GHz ISM band and exploits the Chirp spread spectrum (CSS) technique in order to offer robustness to the transmissions. A typical LoRa radio is characterized by some customizable parameters [29], such as SF, coding rate (CR), and BW. The values allowed for these parameters depend on the region where the LoRa devices are deployed [9].

The SF is the base-2 logarithm of the number of chirps per symbol [11], therefore an LoRa symbol, composed of 2^{SF} chirps, can encode SF bits of information. The SF configuration allows to tune the bit rate, covered distance, and energy consumption. A higher SF increases the signal to noise ratio (SNR) and therefore the sensitivity and coverage range, but it also increases the time on air (ToA), i.e., the transmission duration of a packet. In particular, each SF increase approximately halves the transmission rate and, hence, it approximately doubles the ToA and energy consumption. Consequently, a tradeoff between bit rate and communication range is needed. According to several works in the literature, such as [5], [38], and [39], radio communications with different SFs are orthogonal to each other, so a receiver can successfully receive distinct signals that are sent over a given channel at the same time using different SFs. Recently, the work in [40] demonstrated the quasi-orthogonality of the SFs through both simulations and an implementation based on the universal software radio peripheral (USRP). Consequently, transmissions with different SFs are not completely immune to the adjacent SFs. However, the messages simultaneously transmitted on the same channel with different SFs can be correctly decoded when the signal-to-interference ratio (SIR) of the received packet is above the isolation threshold (see Tables I and II in [40]).

The CR is the forward error correction (FEC) rate used by the LoRa receiver to improve the robustness against interference. It can be set to $4/(4+Z)$, with $Z \in \{1, 2, 3, 4\}$. A higher value of Z offers more robustness, but increases the ToA, as a better error coding introduces a higher transmission overhead (i.e., extra bits in the payload of the LoRa frame).

TABLE I
NOTATIONS USED FOR THE TOA CALCULATION IN (1)

Symbol	Definition
Z	Value of the parameter Z being $CR = \frac{4}{4+Z}$.
NP	Number of preamble symbols.
SW	Length of synchronization word.
PL	Number of PHY payload bytes.
CRC	CRC Presence (1=yes ; 0=no).
IH	PHY header Presence (1=no ; 0=yes).
DE	Use of data rate optimization (1=enabled ; 0=disabled).

TABLE II
PARAMETERS FOR THE TOA CALCULATION

Parameter	Z	NP	SW	PL	CRC	IH	DE
Value	1	8	8	50	1	0	0

The bandwidth represents the range of frequencies in the transmission band. A higher BW gives a higher bit rate (thus, a shorter ToA), but a lower sensitivity, due to additional noise. LoRa networks typically operate at 500, 250, or 125 kHz.

LoRa modulation can transmit arbitrary frames. Semtech's transmitters and receivers use a physical frame format that includes a preamble, an optional header, a payload (limited to 255 bytes), and an optional payload CRC.

The work in [41] provides (1) to calculate the TOA for the transmission of messages using the LoRa modulation. In (1) SF, BW, and CR are the Spreading Factor, Bandwidth, and Coding Rate, respectively. The other notation used in (1) is summarized in Table I

$$\text{ToA} = \frac{2^{\text{SF}}}{\text{BW}}(\text{NP} + 4.25 + \text{SW} + \max(\mathcal{H}, 0)) \quad (1)$$
$$\mathcal{H} = \left\lceil \frac{8\text{PL} - 4\text{SF} + 28 + 16\text{CRC} - 20\text{IH}}{4(\text{SF} - 2\text{DE})} \right\rceil (Z + 4).$$

For example, the transmission of a message with a physical payload of 50 bytes, with SF = 7 and BW = 125 kHz, takes about $97 \cdot 10^{-3}$ s using the parameter values shown in Table II.

B. Regional Parameters EU863-870-MHz ISM Band

ETSI regulations impose to limit the LoRa transmitter activities using either a DC limitation or the so-called listen before talk adaptive frequency agility (LBT AFA) transmission management. The proposed RT-LoRa uses duty-cycle limited transmissions to comply with the ETSI regulations. The maximum DC is defined as the maximum percentage of time during which a transmitter node can occupy a sub-band per hour. Table III shows the limitations on DC and maximum transmission power (expressed in dBm) in the EU863-870MHz ISM band. The number of available LoRa channels (n_{CH}) that are reported in Table III refers to channels with BW of 125 kHz (every channel needs at least 200 kHz considering some guard band [42]).

The four sub-bands do not correspond to four channels, e.g., the sub-band h1.4 is divided into three channels, each one with a bandwidth of 125 KHz.

TABLE III
EU863-870-MHZ ISM BAND LIMITATIONS RELEVANT TO THIS ARTICLE

Sub-band	*Frequency Band (MHz)*	*Available Channels*	*Maximum TX Power (dBm)*	*Duty Cycle*
h1.4	868.00 - 868.60	3	14	1%
h1.5	868.70 - 869.20	2	14	0.1%
h1.6	869.40 - 869.65	1	27	10%
h1.7	869.70 - 870.00	1	14	1%

TABLE IV
EU863-870 PHYSICAL BIT RATES RELEVANT TO THIS ARTICLE

LoRa configuration	Indicative physical bit rate (bit/s)
SF9 / 125kHz	1760
SF8 / 125kHz	3125
SF7 / 125kHz	5470

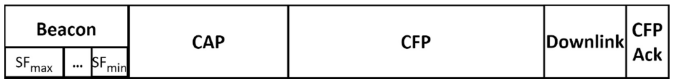

Fig. 1. Superframe structure.

Table IV shows the physical bit rates that are relevant to the configurations used in the assessment.

IV. RT-LoRa DESIGN

RT-LoRa uses a star topology in which the end nodes and the sink communicate through bidirectional links. The sink synchronizes the end nodes with periodic beacons, collects data from the entire network and sends data/commands to the end nodes when needed. Two different types of end nodes are supported, i.e., stationary nodes (SNs), that are fixed in the sensing area, and mobile nodes (MN), that move in the sensing area. The sink is an SN located approximately in the center of the sensing area. Each node can send data belonging to one or more periodic real-time flows.

The centralized medium access mechanism proposed in this article supports both periodic real-time and aperiodic nonreal-time traffic. In RT-LoRa, the network time is organized into superframes that cyclically repeat. Each superframe is composed of five main sections: *Beacon*, CAP, CFP, *Downlink*, and *CFP Ack*. Fig. 1 shows the superframe structure (the size of the sections is not in scale for graphic reasons).

The CAP and CFP contain several *timeslot sets*, where each timeslot is a conventionally defined time interval in the schedule. The timeslots are scheduled over different channels and SFs. Messages that are sent on the same channel with different SFs do not collide, as in our design we make the assumption commonly found in the relevant literature (e.g., in [35]) that the transmissions performed using different SFs are orthogonal.

To easily describe RT-LoRa, we assume that the maximum size of the messages is set by the application. The timeslot duration is set as a function of the maximum message size and of the physical bit rate, so as to ensure that a maximum-sized message can be transmitted in each timeslot. Consequently, the timeslots that schedule transmissions with different SF values have a different duration, as the SF affects the bit rate and ToA. The notation τ_{SF_x} indicates a timeslot for transmissions

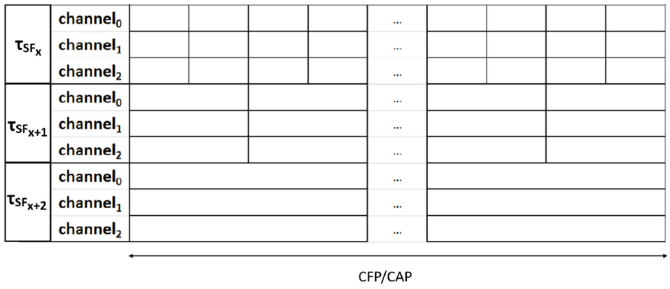

Fig. 2. CFP/CAP structure.

with the SF value set to x, while the notation τ_{s_x} indicates the timeslots belonging to the *timeslot set x*. The duration of the timeslots is calculated according to (1), that imposes a lower bound on the timeslot duration. For the sake of simplicity, we start calculating the duration of the timeslots belonging to the first set (τ_{s_0}), assuming that the transmissions are performed with the minimum SF value that is allowed by the application ($\mathrm{SF}_{\tau_{s_0}} = \mathrm{SF}_{\min}$). Then, we set the duration of a second set of timeslots (τ_{s_1}) for transmissions with the next SF value ($\mathrm{SF}_{\tau_{s_1}} = 1 + \mathrm{SF}_{\min}$). The operation is repeated for each of the n allowed SF values, thus defining n timeslot sets, where the following condition holds: ($\mathrm{SF}_{\tau_{s_{i+1}}} = 1 + \mathrm{SF}_{\tau_{s_i}}$). As each SF increase approximately doubles the ToA, the duration of $\tau_{s_{i+1}}$ is twice the duration of τ_{s_i}, and so on. This way, the timeslots durations are multiple of each other.

The RT-LoRa design assumes that the number n of SF values allowed by the application is less than or equal to the number of used sub-bands. We suggest to set the duration of the CAP and CFP equal to a multiple of the maximum-sized timeslot duration. For example, let us consider an application that allows the SF values $s \in \{7; 8; 9\}$ and transmits messages using the parameters shown in Table II on channels with BW of 125 kHz. In this case, the duration of τ_{s_0} should be set to 0.101 s, i.e., the 0.097 s value obtained by (1) plus 0.004 s, as suggested in [42]. Then, the duration of τ_{s_1} and τ_{s_2} should be equal to 0.202s and 0.404s, respectively.

Fig. 2 shows the structure of CFP/CAP for an application scenario that allows 3 SF values and uses 3 channels. In Fig. 2, τ_{SF_x} indicates the timeslots belonging to the first set (i.e., τ_{s_0}), therefore, $\tau_{\mathrm{SF}_{x+1}}$ corresponds to τ_{s_1}, while $\tau_{\mathrm{SF}_{x+2}}$ corresponds to τ_{s_2}.

The following sections provide details about each superframe section.

A. Beacon Section

The first section of a superframe is reserved to beacon transmissions. Differently from Industrial LoRa [10], in which the Beacon section is only for synchronization purposes and only one beacon is broadcast by the sink (using the highest SF among the values allowed by the application), RT-LoRa proposes a novel approach based on multiple beacon transmissions. Given a set of n SF values allowed by the application requirements, these are used by the sink to broadcast n beacons at the beginning of each superframe. Each node updates a dynamic list of recommended SF values according to the SFs of the beacons received during the current superframe. The list is called the $l_{\mathrm{SF}}^{(a)}$ list. The update period of the list is equal to the superframe duration. If no beacons are received during the Beacon section of a superframe, the list will be automatically updated with the highest available SF value allowed by the application. The proposed multiple beacon strategy provides several advantages to MNs. In fact, the nodes are synchronized to the network time and they are able to determine the SF values currently recommended for transmission to the sink. However, this strategy results in a larger energy consumption, as it requires the transmission of multiple beacon messages. The beacon format is implementation-specific. The lower bound for the Beacon section duration corresponds to the sum of the ToA of the beacons.

B. Contention Access Period

In the CAP, aperiodic unconfirmed messages are sent from the end nodes to the sink. The end nodes compete for the channel access using a slotted ALOHA-based mechanism. RT-LoRa randomly generates the 3-tuple of parameters (channel, SF, and timeslot) that will be used for transmission. The SF is randomly selected among the currently recommended values, i.e., those in the $l_{\mathrm{SF}}^{(a)}$ list described in the previous section. The set of available channels is prepared using a DC check function, that verifies the transmission feasibility on the basis of the expected ToA of the message according to the DC restrictions.[2] If the list of available channels is not empty, the message is transmitted using the selected channel and SF, so as to reduce the collision probability. Otherwise, the message is queued. The CAP is made up of a set of timeslots on different channels and SFs, as shown in Fig. 2. We suggest to set the duration of the CAP as a multiple of the maximum-sized timeslot duration. A longer CAP duration gives a lower collision probability, thus achieving better performance (i.e., lower PLR) for aperiodic transmissions. During the CAP, each end node remains in sleep mode, except when it has an aperiodic message to transmit. In that case, the end node changes its state from sleep to transmit mode right before the message transmission. After the transmission, the end node returns to the sleep mode.

C. Contention-Free Period

The CFP is devised for the real-time unidirectional communications from the end nodes to the sink (uplink). It consists of n sets (one for each SF value allowed by the application) of timeslots (see Fig. 2). The timeslots are used to schedule the confirmed transmission of messages belonging to periodic flows, using a multi-CH (channel) and multi-SF (spreading factor) TDMA strategy. Each ith timeslot in the s_k set is characterized by three items: the *identifier* $\mathcal{I}_{(i,k)}$, that determines the start time of the scheduled transmission within the CFP, the *channelset* $\mathcal{C}_{(i,k)}$, that contains the set of frequencies used, and the SF *value* $\mathcal{S}_{(i,k)}$. All the timeslots in the same set s_k are characterized by the same SF value, i.e., $\mathcal{S}_{(i,k)} = \mathrm{SF}_{\tau_{s_k}}$. During

[2]The DC check function takes into account both the bandwidth reserved for the communications in CFP per hour and the one consumed in the CAP in the last hour.

the timeslot $\mathcal{I}_{(i,k)}$, the transmission is performed using the SF $\mathcal{S}_{(i,k)}$ and the channel $\mathcal{C}_{(i,k)}(0)$ (i.e., the first frequency in the channelset). Next, the frequencies in the channelset rotate. It is advisable that the frequencies in the channelset belong to different sub-bands. This mechanism, that runs every superframe, improves communication robustness and bandwidth utilization, while remaining compliant with the DC restrictions of the selected sub-bands. For the sake of simplicity, in RT-LoRa the timeslots in the CFP are assigned offline.

In the CFP, each real-time flow generated by an SN is assigned to one timeslot. The latter is chosen so that the related SF value ensures a reliable communication between the stationary source node and the sink. Conversely, the real-time data flows generated by an MN may need multiple timeslots to transmit one message. RT-LoRa supports three QoS-classes for these flows, each associated to a different strategy for the transmission over one or multiple timeslots. The available QoS classes are as follows.

1) *Normal (N):* These flows are scheduled in x nonoverlapping timeslots (with different SF values) in the CFP, where x is the number of SFs allowed by the application. The channelsets of the timeslots go through the frequency rotation mechanism. Every message is transmitted in the timeslot that, among the ones assigned to the *N*-flow, is characterized by the lowest recommended SF value. An SF-value is "recommended" when it belongs to the $l_{\mathrm{SF}}^{(a)}$ list. This flow class optimizes energy efficiency, as it takes the shortest ToA among the available options.
2) *Reliable (R):* These flows are scheduled in one timeslot in the CFP, with the highest SF value allowed by the application. While representing the best option in terms of bandwidth utilization, as it needs one transmission per message, this class results in higher energy consumption (due to the longer ToA).
3) *Most Reliable (R+):* These flows are scheduled in x nonoverlapping timeslots (with different SF values) in the CFP, where x is the number of SFs allowed by the application. Every message is transmitted multiple times (*replicas*) during the scheduled timeslots whose SF is in the $l_{\mathrm{SF}}^{(a)}$ list. The flows in the $R+$ class achieve the highest reliability, thanks to the redundant transmissions, but the highest energy consumption.

The QoS-class is assigned to each data flow offline, thus realizing a tradeoff between reliability in transmission and energy consumption. Note that the application layer may not be synchronized with the MAC layer, so every message will be sent within the superframe that follows the superframe in which the message has been generated.

Each end node is in transmit mode during the timeslots used for the transmission of periodic real-time flows, while it remains in sleep mode during the other timeslots in the CFP.

D. Downlink Section

The Downlink section is used by the sink for transmitting to the end nodes *Unicast* or *Broadcast* messages. The first ones are sent by the sink to a specific end node. A bit in the payload indicates if the message is confirmed. If so, the end node acknowledges immediately after receiving the message. The sink transmits to the end node using the same SF of the last successful transmission from the end node, thus increasing the success probability comparing with a random choice. Conversely, Broadcast messages are sent by the sink to all the end nodes in the network, using the maximum SF value among the ones allowed by the application. Broadcast messages are unconfirmed. The end nodes must be on for listening to the downlink channel (i.e., the established channel for downlink communications) during this section. The Downlink section is optional and its duration is fixed by the network designer so as to find a tradeoff between the sink downlink throughput and the energy consumption of the end nodes.

E. CFP Ack Section

The CFP Ack section is used for *global-Ack* (GACK) frames that the sink broadcasts to confirm the messages received from the end nodes in the CFP of the same superframe. The GACK message contains a data structure of z bits per each end node in the network. The format of such a data structure is implementation-specific. If one GACK message of $r*z$ bits cannot be sent broadcast, as it is larger than the maximum allowed physical payload, it can be split into multiple frames, consecutively transmitted in the CFP Ack section. The end nodes that send messages in the CFP Ack section of the last superframe, if they are interested in the ack, must be on and listening during this section. The duration of this section is implementation dependent.

V. RT-LoRa Network Configuration

This section proposes guidelines for the superframe configuration in an RT-LoRa network for an application with a specific cycle time (CT), here defined as the time interval in which all nodes can transmit once for each locally generated data flow, highlighting the protocol limits due to the DC restrictions. The equations provided in this section were used to set simulation parameters in Section VI-B and to verify the simulation results in Sections VI-C and VI-D.

In the following (Section V-A), a schedulability assessment of RT-LoRa is provided in order to verify the schedule feasibility for the network flows, with a specific set of configuration parameters. If no feasible schedule is found, a change of the application parameters is required, such as a new set of SFs to be used for the application and/or a different QoS class for the flows (*N*, *R*, *R+*).

For the schedulability assessment, we need to calculate the minimum superframe duration (Section V-A) for a given set of periodic flows. To accomplish this, we calculate the minimum CFP duration needed to schedule all the periodic flows in the network (Section V-B) and the timing constraint, in terms of the minimum superframe duration, that derives from the DC restriction (Section V-C).

The equations provided in the following are relevant to the end nodes, so we do not consider beacons, downlink and CFP Ack communications. However, a DC check function, implemented in the sink, takes into account all the sink

TABLE V
SUMMARY OF NOTATIONS

Symbol	Definition
n_{SB}	Number of used sub-bands.
$\{\mathcal{S}\}$	Set of SF values allowed by the application.
s_{min}	Minimum SF value allowed by the application.
s_{max}	Maximum SF value allowed by the application.
$ToA_{(SF=x)}$	Time on Air of a maximum-sized messages transmitted with SF value equal to x.
$\{\mathcal{N}_S\}$	Set of stationary nodes (SNs).
$\{\mathcal{N}_M\}$	Set of mobile nodes (MNs).
$\mathcal{T}_{sl(SF=x)}$	Duration of a timeslot with transmissions performed with SF value x.
$\mathcal{T}^{(k)}_{sl(X,i,SF=x)}$	Duration of the timeslot assigned to the i-th flow of the k-th node characterized by SF value equal to x. X can be N, R or R+ for MNs and Q for SNs.
$\{\mathcal{S}\}^{(k)}_{(X,i)}$	Set of SFs in the timeslots assigned to the i-th flow generated by the k-th node. X can be N, R or R+ for MNs and Q for SNs.
$\sigma^{(k)}_{(X,i)}$	Time interval within which the first and the last timeslot for the N (or R+) i-th flow generated by the k-th node are allocated.
$D^{(k)}_{(X,i)}$	Deadline of the i-th flow generated by the k-th node. X can be N, R or R+ for MNs and Q for SNs.
DC_{min}	Minimum duty cycle value among those allowed by all the used sub-bands.
n_N	Overall number of N-flows generated by the MNs.
n_R	Overall number of R-flows generated by the MNs.
n_{R^+}	Overall number of R+-flows generated by the MNs.
$n_{(SF=x)}$	Overall number of flows that need a timeslot with SF value equal to x generated by the SNs.
$n^{(j)}_X$	Number of X-flows required by the j-th MN (X can be N, R or R+).
$n^{(h)}_{(SF=x)}$	Number of flows generated by the h-th SN with SF value equal to x.
$\Delta_S(h)$	Overall time needed by the h-th SN to send one maximum-sized message for each of the $n^{(h)}_{(SF=x)}$ locally generated flows.
$\Delta_M(k)$	Overall time needed by the k-th MN to send one maximum-sized message for each of the $n^{(k)}_X$ locally generated flows.
η_{tx}	Minimum value within the set containing the number of times that each node can transmit the real-time flows in the CFP per hour, according to the duty cycle restrictions.
$\mathcal{T}_{supfrm}$	Superframe duration.
$\mathcal{T}_{CFP}$	CFP duration.
$\mathcal{T}^{ALL}_{CFP}(s)$	Time needed to schedule the transmission of the real-time flows generated by all the SNs and MNs in the CFP for the SF s.
$\mathcal{T}_{id}$	Implementation-dependent time duration.

transmissions. Consequently, in the simulations all the needed checks are in place.

Table V summarizes the notations used in this section.

A. Schedulability Analysis

This section presents a schedulability assessment of RT-LoRa. In particular, a methodology to calculate the minimum superframe duration for a given set of periodic flows is provided. Then, the superframe duration is compared with the timing constraint deriving from the DC restriction, that defines the longest time interval (in percentage) during which a transmitter node can occupy a specific sub-band per hour. The value of the maximum e2e delay for a message is calculated as the sum of the superframe duration and the application-defined value $\sigma^{(k)}_{(X,i)}$ corresponding to the time interval, within a superframe, between the start of the first scheduled timeslot for a specific flow and the end of the last timeslot scheduled for the same flow (if only one timeslot is scheduled for the flow, $\sigma^{(k)}_{(X,i)}$ is equal to the timeslot duration). If the maximum e2e delay value is lower than or equal to the application timing constraint, then the schedule is feasible, as the flows will meet their deadline while complying with the DC constraint. Otherwise, a change of the application parameters is required, for instance, a new set of SFs to be used for the application and/or a different QoS class for the flows (N, R, $R+$), etc., so as to find a feasible schedule with RT-LoRa.

It is mandatory that $\mathcal{T}_{\text{supfrm}} \geq \mathcal{T}^{\text{MIN}}_{\text{supfrm}}$, i.e., the superframe duration $\mathcal{T}_{\text{supfrm}}$ must be larger than or equal to the minimum admissible value for a superframe duration $\mathcal{T}^{\text{MIN}}_{\text{supfrm}}$, calculated as

$$\mathcal{T}^{\text{MIN}}_{\text{supfrm}} = \max\left[\mathcal{T}_{\text{CFP}},\ \mathcal{T}^{(\text{DC})}_{\text{supfrm}}\right] + \mathcal{T}_{\text{id}}. \tag{2}$$

In (2), we describe the following.

1) $\mathcal{T}_{\text{CFP}}$ is the CFP duration needed to schedule the transmission to the sink of one maximum-sized message for each real-time flow in the network (Section V-B).
2) $\mathcal{T}^{(\text{DC})}_{\text{supfrm}}$ is the superframe duration ($\mathcal{T}_{\text{supfrm}}$) according to the DC limitation that considers the maximum number of times that each node can transmit the real-time flows per hour (Section V-C).
3) $\mathcal{T}_{\text{id}}$ is an implementation-dependent time duration that includes the duration of the Beacon, CAP, Downlink, and CFP Ack sections.

The $\mathcal{T}_{\text{CFP}}$ is a "network limitation," as it considers all the flows in the network, while $\mathcal{T}^{(\text{DC})}_{\text{supfrm}}$ is a "local limitation," as it is relevant to the node schedule. Consequently, the $\mathcal{T}^{\text{MIN}}_{\text{supfrm}}$ is calculated using the maximum between these values.

In detail, the flow set can be scheduled when $\mathcal{T}_{\text{supfrm}}$ is lower than the required cycle-time.

$$\mathcal{T}_{\text{supfrm}} \leq CT. \tag{3}$$

We assume that the message generation period is equal to the CT in all the network end nodes. As discussed in Section IV-C, the number of nonoverlapping timeslots for the normal flow (N) or the most reliable ($R+$) ith flow generated by the k-th node, i.e., w, is equal to the number of values in the $\{\mathcal{S}\}$ set. These w timeslots are allocated in the CFP within an application-defined time interval $\sigma^{(k)}_{(X,i)}$ between the sum of the durations of the timeslots scheduled for the flow and the CFP duration

$$\sum_{s\in\{\mathcal{S}\}^{(k)}_{(X,i)}} \mathcal{T}^{(k)}_{sl(X,i,\text{SF}=s)} \leq \sigma^{(k)}_{(X,i)} \leq \mathcal{T}_{\text{CFP}} \tag{4}$$

where $X \in \{N, R^+\}$. In fact, $\sigma^{(k)}_{(X,i)}$ will be equal to the sum of the durations of the timeslots scheduled for the flow if these timeslots are scheduled consecutively. On the other hand, $\sigma^{(k)}_{(X,i)}$ will be equal to $\mathcal{T}_{\text{CFP}}$ if the first and the last timeslots of the CFP are scheduled for the flow. The message belonging to the ith flow is delivered within the deadline $D^{(k)}_{(X,i)}$ if

$$D^{(k)}_{(X,i)} \geq \mathcal{T}_{\text{supfrm}} + \sigma^{(k)}_{(X,i)} \Rightarrow \mathcal{T}_{\text{supfrm}} \leq D^{(k)}_{(X,i)} - \sigma^{(k)}_{(X,i)}. \quad (5)$$

Equation (5) can be used also for an R-flow or a flow generated by an SN, with $\sigma^{(k)}_{(X,i)}$ equal to the duration of the timeslot scheduled for that flow.

The worst case in terms of e2e delay is represented by specific conditions on an N-flow due to the features of this type of flow. In fact, every message of an N-flow is transmitted in the scheduled timeslot that is characterized by the lowest recommended SF value. Then, it is possible that in a superframe the timeslot recommended for the transmission is the first among the ones assigned to the N-flow, while in the next superframe the timeslot recommended for the transmission is the last one. Such a case can be described as follows. We assume that the first scheduled timeslot for that N-flow is selected for the transmission in a specific superframe. We also assume, to take the worst case, that the application layer of the nodes is not synchronized with the MAC layer and, for this reason, the flow message is generated right after the assigned timeslot. Consequently, the message cannot be sent in the current superframe. As we are looking for the worst case, we assume that the message is generated right after the end of the timeslot it is assigned. Then, the message transmission is postponed to the next superframe. In the worst case, the timeslot assigned for the transmission of the message in the next superframe will be the last one.

B. CFP Duration

In this section, we deal with the minimum CFP duration that is needed to schedule the transmission to the sink of one maximum-sized message for each real-time flow in the network.

The overall time $\mathcal{T}^{\text{ALL}}_{\text{CFP}}(s)$ needed to schedule the transmission of the real-time flows generated by all the stationary and MNs in the CFP, calculated for a given SF value s in $\{\mathcal{S}\}$, is

$$\mathcal{T}^{\text{ALL}}_{\text{CFP}}(s) = \left\lceil \frac{n^{\diamond}_R(s) + n_N + n_{R^+} + n_{(\text{SF}=s)}}{n_{\text{SB}}} \right\rceil * \mathcal{T}_{sl(\text{SF}=s)}$$
$$\text{where } n^{\diamond}_R(s) = \{n_R \text{ if } (s = s_{\max}); 0 \text{ otherwise}\}. \quad (6)$$

Note that $\mathcal{T}^{\text{ALL}}_{\text{CFP}}(s)$ takes into account the number of used sub-bands (i.e., n_{SB}). In fact, the number of CFP timeslots required by all the nodes for each SF allowed by the application is divided by n_{SB}. The CFP duration $\mathcal{T}_{\text{CFP}}$ is the maximum value among the times $\mathcal{T}^{\text{ALL}}_{\text{CFP}}(s)$, calculated for each SF s in $\{\mathcal{S}\}$, i.e.,

$$\mathcal{T}_{\text{CFP}} = \max\left[\mathcal{T}^{\text{ALL}}_{\text{CFP}}(s)\right]_{\forall s \in \{\mathcal{S}\}}. \quad (7)$$

C. Superframe Duration According to the Duty Cycle Restrictions

In this section, we address the minimum superframe duration according to the DC limitation.

Equations (8) and (9) provide $\Delta_S(h)$ and $\Delta_M(k)$, i.e., the overall time needed by the hth SN and the k-th MN, respectively, to send one maximum-sized message for each of its real-time flows, according to the transmission strategies described in Section IV-C

$$\Delta_S(h) = \left\{ \sum_{i=s_{\min}}^{s_{\max}} \left[\left(n^{(h)}_{(\text{SF}=i)} \right) \cdot \text{ToA}_{(\text{SF}=i)} \right] \right\} \quad (8)$$

$$\Delta_M(k) = \left\{ \sum_{i=s_{\min}}^{s_{\max}} \left[\left(n^{(k)}_N + n^{(k)}_{R+} \right) \cdot \text{ToA}_{(\text{SF}=i)} \right] \right\} + n^{(k)}_R \text{ToA}_{(\text{SF}=s_{\max})}. \quad (9)$$

In (8) and (9), the aperiodic communications in the CAP and the ack messages of confirmed downlink communications are not considered, as they are unpredictable. However, a DC check function, implemented in each end node, takes into account all the transmissions.

In (10), we calculate η_{tx} as the minimum value within the set containing the number of times that each node can transmit the real-time flows in the CFP per hour in the sub-band with $\text{DC} = \text{DC}_{\min}$, i.e., we also consider the worst case in terms of DC. If the DC limit is met in the sub-band in which $\text{DC} = \text{DC}_{\min}$, it is also met in all the other sub-bands in which $\text{DC} > \text{DC}_{\min}$. In (10), $\Delta_S(h)$ and $\Delta_M(k)$ are calculated $\forall h \in \{N_S\}$ and $\forall k \in \{N_M\}$, respectively. This way, (10) takes into account all the possible nodes in the network and chooses the minimum value, i.e., the worst case in terms of the number of times that each node can transmit the real-time flows in the CPF per hour. Note that the worst case is the one in which $\Delta_S(h)$ or $\Delta_M(k)$ takes the highest value among the ones calculated for all the nodes in the network. Furthermore, (10) takes into account the frequency rotation for CFP timeslots introduced by RT-LoRa through the parameter n_{SB}, that makes the number of allowed transmissions increase n_{SB} times

$$\eta_{tx} = \min \left\{ \begin{matrix} \left[\left\lfloor \frac{3600 \cdot \text{DC}_{\min} \cdot n_{\text{SB}}}{\Delta_S(h)} \right\rfloor \right]_{\forall h \in \{\mathcal{N}_S\}} \\ \left[\left\lfloor \frac{3600 \cdot \text{DC}_{\min} \cdot n_{\text{SB}}}{\Delta_M(k)} \right\rfloor \right]_{\forall k \in \{\mathcal{N}_M\}} \end{matrix} \right\}. \quad (10)$$

Equation (11) provides the superframe duration $\mathcal{T}^{(\text{DC})}_{\text{supfrm}}$, calculated according to the worst case in terms of the number of times a node can transmit real-time flows per hour (i.e., every 3600 s)

$$\mathcal{T}^{(\text{DC})}_{\text{supfrm}} = \frac{3600}{\eta_{tx}}. \quad (11)$$

VI. Simulative Assessment

In this section, a performance assessment of RT-LoRa obtained using the OMNeT++ environment is presented. The module that simulates the RT-LoRa MAC layer was developed as an extension of the Industrial LoRa simulator [10], while the LoRa physical layer and the wireless channel were simulated

using FLoRa [43]. Consequently, as discussed in [43], simultaneous transmissions on different SFs are considered orthogonal using the FLoRa physical layer. A DC check function was implemented in each node (both sink and end nodes) in order to be compliant with the DC limitations. Two performance metrics were used: 1) the PLR and 2) the e2e delay.

The PLR, measured at the application layer, is expressed as a percentage according to

$$\mathrm{PLR} = \left(\frac{n_{\mathrm{lostMsg}}}{n_{tx\mathrm{Msg}}}\right) * 100 = \left(1 - \frac{n_{rx\mathrm{Msg}}}{n_{tx\mathrm{Msg}}}\right) * 100 \quad (12)$$

where $n_{tx\mathrm{Msg}}$, n_{lostMsg}, and $n_{rx\mathrm{Msg}}$ are the number of transmitted, lost, or correctly received messages, respectively, measured over all the end nodes in the network. In (12) a message is counted once even if it is transmitted or received in multiple replicas.

The e2e delay is the time that a message takes since its generation at the application layer of the source node and its reception at the application layer of the sink, calculated according to

$$\mathrm{e2eDelay} = \mathrm{ArrivalTime} - \mathrm{GenTime}. \quad (13)$$

In the simulation the processing delays were not considered, as they are implementation-dependent.

A. Simulated Scenario

We performed the simulative assessment in a realistic industrial use case similar to the ones presented in [10] and [37]. The considered scenario involves a large number of end nodes that communicate with one sink. An end node can be either stationary (SN) or mobile (MN).

The simulated network topology consists of a network with 101 nodes, i.e., a sink (located in the center of the sensing area to ensure the maximum coverage range), 25 SNs and 75 MNs. All the nodes are placed within an area, here called sensing area, with a range of 250 m around the sink.

The simulated scenario consists of sensor nodes that generate a periodic message every 30 s. The deadline D by which the message must be delivered is 30 s, too. This is a realistic use case of a distributed measurement system for Industrial IoT applications (e.g., automation ones) in which CTs up to 60 s are required [6], [35]. RT-LoRa, due to the LoRa DC limitation, is devised for applications with no very tight time requirements. For example, RT-LoRa may fit well process control applications, as they feature long CTs. Conversely, as in discrete manufacturing the CTs can be quite short, it could happen that, with a given set of flows and configuration parameters, no feasible RT-LoRa schedule is found. In this case, the network designer needs to try different configuration parameters, such as a new set of SFs to be used for the application and/or a different QoS class for the flows (N, R, $R+$), and so on.

The considered physical payload is 50 bytes long, a typical value for industrial applications [35], [37].

Every end node has one periodic real-time data flow scheduled in the CFP. Furthermore, aperiodic nonreal-time messages are generated by the end nodes every t seconds, where t is a random variable that varies according to an exponential distribution with mean 70 s. The nonreal-time messages are sent in the shared bandwidth of the CAP. The application layer of the end nodes is not synchronized with the MAC layer.

TABLE VI
END NODES CONFIGURATION SETTINGS

Number of nodes	*Type of nodes*	*Distance from the sink*	*Allocated CFP timeslots*
10	SN	0-125 m	τ_{SF_7}
10	SN	125-180 m	τ_{SF_8}
5	SN	180-250 m	τ_{SF_9}
75	MN	0-250 m	According to the QoS classes of the flows

B. Simulation Settings

The transmissions were performed on three sub-bands (h1.4, h1.6, and h1.7, see Table III) in order to achieve a duty-cycle limitation of 12%. As the h1.5 sub-band is not used, in (10) $\mathrm{DC}_{\min} = 1\%$. CFP frequency rotation involves only one channel for each of the three sub-bands. The h1.4 sub-band contains three available channels, but only one of these (randomly chosen) was used. In fact, as the h1.4 sub-band is subject to a DC limit of 1%, transmitting one message for each of its three channels (i.e., using all the five channels available in all the three sub-bands at the same time) would not be advisable. The use of five channels would be suitable only with a high CT and a very large number of end nodes, as it allows to schedule more simultaneous transmissions in the CFP.

The RT-LoRa protocol was set to use values of SF $s \in \{7; 8; 9\}$ in order to work with a bit rate complying with the application requirements (i.e., from 1760 to 5470 bit/s, with a BW of 125 kHz).

The configuration settings of the end nodes in the simulated scenario are summarized in Table VI.

The 75 MNs moved within the sensing area with speed between 0.5 and 1 m/s according to the ChiangMobility model, available in the INET framework.

In the simulations it was assumed that 25 MNs schedule a *normal flow* (N), 25 MNs schedule a *reliable flow* (R) and 25 MNs schedule a *most reliable* ($R+$) flow. Since the number of SFs allowed by the application is three (i.e., 7–9), every message of $R+$ flows is transmitted three times. Equation (1) was used to calculate the lower bound for time slot duration with the parameter values in Table II. About 0.004 s were added to that value, as suggested in [42]. In particular, the duration $\mathcal{T}_{sl(\mathrm{SF}=7)}$ of the τ_{SF_7} timeslots, the duration $\mathcal{T}_{sl(\mathrm{SF}=8)}$ of the τ_{SF_8} timeslots and the duration $\mathcal{T}_{sl(\mathrm{SF}=9)}$ of the τ_{SF_9} timeslots were set to 0.101 s, 0.202 s, and 0.404 s, respectively. The three nonoverlapping slots allocated in the CFP for an N-flow or a $R+$-flow were scheduled within a time interval of $\sigma^{(k)} = 3 * \mathcal{T}_{sl(\mathrm{SF}=9)}$ for each k-th node [see (4)].

The lower bound for the superframe duration was calculated using (2). The calculated value for $\mathcal{T}^{\mathrm{MIN}}_{\mathrm{supfrm}}$ is 20.112 s, lower than the required deadline D (30 s). Moreover, such a value is also lower than $D - \sigma^{(k)}$ for each k-th node, therefore the system was able to schedule the given periodic real-time flows

TABLE VII
CONFIGURATIONS SETTING

Parameter	*Configuration (A)*	*Configuration (B)*
Beacon section	0.707s	0.707s
CAP	6.060s	14.140s
CFP	10.908s	10.908s
Downlink section	0.808s	0.808s
CFP Ack	2.0s	2.0s
$\mathcal{T}_{supfrm}$	$\mathcal{T}_{supfrm}^{cfgA}$= 20.483s	$\mathcal{T}_{supfrm}^{cfgB}$= 28.563s

TABLE VIII
SIMULATION PARAMETERS

Parameter	*Value*
Coding Rate	4/5
Bandwidth	125 kHz
Transmission Power	14 dBm
$\{S\}$	$\{7; 8; 9\}$
Sub-bands	$\{h1.4; h1.6; h1.7\}$
Physical payload	50 bytes
Generation period of the periodic messages	30s
$\mathcal{T}_{sl(SF=7)}$	0.101s
$\mathcal{T}_{sl(SF=8)}$	0.202s
$\mathcal{T}_{sl(SF=9)}$	0.404s
$\mathcal{T}_{sim}$	36000s
Propagation model	LoRa log-normal shadowing model

in the CFP [see (5)]. The $\mathcal{T}_{\text{CFP}}$ duration was set to 10.908 s, according to (7).

We considered two case studies corresponding to different configurations in the same scenario in order to show how different superframe configurations affect the performance of aperiodic messages in terms of PLR and the e2e delays of periodic messages. In particular, in the configuration (A) the $\mathcal{T}_{\text{supfrm}}$ is set at a value close to $\mathcal{T}_{\text{supfrm}}^{\text{MIN}}$ ($\mathcal{T}_{\text{supfrm}} \geq \mathcal{T}_{\text{supfrm}}^{\text{MIN}}$), i.e., the minimum allowed value of the superframe duration. Conversely, in the configuration (B) the $\mathcal{T}_{\text{supfrm}}$ is set at a value close to $D - \sigma^{(k)}$ ($\mathcal{T}_{\text{supfrm}} \leq D - \sigma^{(k)}$), i.e., the maximum allowed value of the superframe duration. The larger $\mathcal{T}_{\text{supfrm}}$ in the configuration (B) than the configuration (A) allows to extend the CAP. This way, we expect to obtain a lower PLR for aperiodic messages and a higher (but always less than the deadline) maximum e2e delay for periodic messages. The configurations setting are shown in Table VII.

Note that the duration of the beacon section was set to $\mathcal{T}_{sl(\text{SF}=9)} + \mathcal{T}_{sl(\text{SF}=8)} + \mathcal{T}_{sl(\text{SF}=7)}$, but the sink only transmits for a time interval equal to $\text{ToA}_{(\text{SF}=9)} + \text{ToA}_{(\text{SF}=8)} + \text{ToA}_{(\text{SF}=7)}$, according to (1), where the PL parameter is the physical payload of a beacon message (implementation dependent).

The simulation time was set to $\mathcal{T}_{\text{sim}} = 36000$ s (i.e., 10 h) in order to collect a significant amount of data. The LoRa log-normal shadowing model [43], a propagation model provided by the FLoRa framework, was adopted for the simulations.

The relevant parameters are summarized in Table VIII.

C. Simulation Results—Aperiodic Transmissions

Fig. 3 shows the average PLR measured for the configurations (A) and (B) for the aperiodic messages transmitted in the CAP. The flows of SNs are grouped by the maximum distance between the end node and the sink. The flows of MNs are grouped by their QoS class.

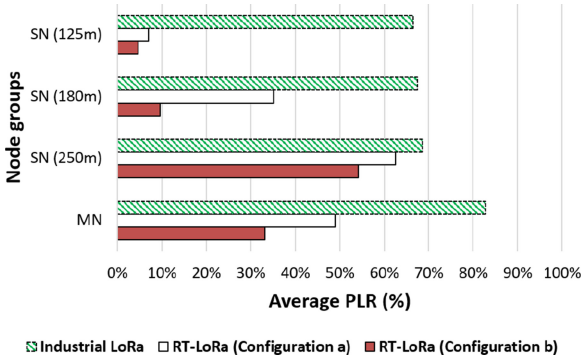

Fig. 3. Average PLR for aperiodic messages. Comparison between RT-LoRa (configurations A and B) and Industrial LoRa.

Fig. 3 also shows that the messages sent by the SNs that are furthest from the sink experienced the highest PLR. This can be explained as follows. For a CAP transmission, each end node randomly selects an SF among the recommended values in the list that is updated according to the multiple beacons received during the current superframe. The SNs closest to the sink can choose among multiple SF values, whereas it is likely that the furthest ones can only select SF9. As a consequence, the probability of collision, i.e., that two nodes randomly select the same set of transmission parameters in the CAP (i.e., the same timeslot identifier, channel, and SF), is lower for the devices closer to the sink than for the ones that are further away. The lower probability of collision determines the lower PLR measured for the closest SNs. About an MNs, the measured PLR depends on their distance from the sink, which changes according to the ChiangMobility model.

The aperiodic messages in configuration (B) experienced a lower PLR than in configuration (A). This is due to the longer duration of the CAP in configuration (B). In fact, using a larger CAP the end nodes can choose the timeslot to use for transmission among a larger set of timeslot identifiers, thus lowering the collision probability.

Finally, Fig. 3 shows that the slotted ALOHA-based MAC layer (used for transmissions in the CAP) of RT-LoRa outperforms the pure ALOHA-based MAC layer used by Industrial LoRa [10].

D. Simulation Results—Periodic Transmissions

The measured PLR values for the periodic N-flow messages using the configurations (A) and (B) were 1.44% and 2.28%, respectively, and are tolerable for the considered scenario. All the other periodic messages were received by the sink and confirmed in the CFP Ack section. The loss of some N-flow messages is because each message of an N-flow is transmitted in the timeslot with the lowest recommended SF value among those in the dynamic list maintained by the source node. If the source node is mobile, sometimes this list can be no longer reliable, because it is updated on the basis of the beacons received every superframe. As a consequence, a message might be sent using an SF that is no longer reliable, thus being lost or corrupted.

TABLE IX
MAXIMUM E2E DELAY OF REAL-TIME FLOWS

Nodes (QoS class of the flows)	Max e2e delay *Configuration (A)*	Max e2e delay *Configuration (B)*
Stationary nodes	20.884s	28.966s
Mobile nodes (N-flows)	21.695s	29.764s
Mobile nodes (R-flows)	20.874s	28.964s
Mobile nodes (R+-flows)	20.887s	28.960s

TABLE X
MAIN CHARACTERISTICS OF LORAWAN, INDUSTRIAL LORA, AND RT-LORA

	LoRaWAN	Industrial LoRa	RT-LoRa
Topology	Star (*)	Star	Star
Synchronization	Not supported	Beacon-based	Multiple beacons based
Transmission of real-time periodic flows	Not supported	Supported	Supported
Smart selection of Spreading Factors	Supported through ADR	Not supported	Supported through multiple beacons
MAC strategy for aperiodic transmission	Pure ALOHA	Pure ALOHA	Slotted ALOHA
MAC strategy for periodic transmission	Pure ALOHA	Multi-CH and Multi-SF TDMA	Multi-CH and Multi-SF TDMA
QoS classes	Downlink only (3 device classes provided)	Not supported	Uplink only (3 QoS classes provided)
Support for retransmission	Uplink only	Not supported	Not supported
Frequency rotation	Pseudo-random channel hopping	Not supported	Supported

(*) star-of-stars when multiple gateways are used

The maximum e2e delays measured for the periodic messages transmitted in the CFP section are shown in Table IX. As expected, the e2e delays (measured at the application layer for all the periodic real-time messages) are always lower than the deadline $D = 30$ s. Hence, they are upper-bounded. Moreover, as discussed in Section V and according to (5), the measured e2e delays are always lower than (or equal to) $(\mathcal{T}^{cfgA}_{\text{supfrm}} + \sigma^{(k)}) = 21.695$ s for configuration (A) and than $(\mathcal{T}^{cfgB}_{\text{supfrm}} + \sigma^{(k)}) = 29.775$ s for configuration (B), thus demonstrating the effectiveness of the proposed approach.

E. Summary and Discussion

Table X recaps the features of the LoRaWAN [8], Industrial LoRa [10], and RT-LoRa MAC protocols.

Both Industrial LoRa and RT-LoRa organize the network time in superframes that cyclically repeat. Moreover, both protocols support a mechanism for superframe synchronization that is based on beacon messages sent broadcast over the network. Conversely, LoRaWAN does not support any superframe structure or synchronization mechanism. Both Industrial LoRa and RT-LoRa are tailored for industrial applications. In fact, they support the transmission of periodic real-time flows in the CFP of the superframe, whereas LoRaWAN does not provide any support for real-time messages, due to the nondeterminism of its ALOHA-based medium access strategy.

Both LoRaWAN and RT-LoRa, for different purposes, use smart mechanisms to select the SF values eligible for transmission. LoRaWAN supports the so-called adaptive data rate (ADR) scheme to manage data rate (i.e., the SF) and RF output in order to maximize the battery life of the node. RT-LoRa uses a novel mechanism based on the transmission of multiple beacon frames that allow the end nodes to be aware of the SF values currently recommended for a transmission, thus ensuring higher reliability. We recall that the communication reliability is, together with timeliness, very important in Industrial IoT scenarios. RT-LoRa performs aperiodic transmissions in the CAP of the superframe using a custom slotted ALOHA strategy. Conversely, LoRaWAN and Industrial LoRa use a pure ALOHA strategy for the transmission of aperiodic flows. As a consequence, RT-LoRa outperforms both LoRaWAN and Industrial LoRa, as the smart version of slotted ALOHA determines a lower probability of collision between two transmitting nodes.

RT-LoRa supports three QoS classes (named N, R, and $R+$) that allow to achieve a tradeoff between reliability of uplink communications and energy consumption. LoRaWAN instead provides three device classes (named A, B, and C) in order to achieve a tradeoff between downlink communication latency and battery lifetime. Industrial LoRa does not provide any QoS support.

In LoRaWAN, Industrial LoRa, and RT-LoRa the transmissions sent by nodes (uplink) can be confirmed using the ACK messages. However, RT-LoRa and Industrial LoRa do not support the retransmission of the unconfirmed messages. Conversely, LoRaWAN supports retransmissions.

The RT-LoRa frequency-rotation procedure allows for a better utilization of the available bandwidth comparing with Industrial LoRa, while remaining compliant with the duty-cycle restrictions stated by the ETSI regulations. LoRaWAN, instead, implements only a pseudorandom channel hopping.

VII. CONCLUSION

This article proposed RT-LoRa, a medium access strategy working over LoRa that guarantees bounded e2e transmission delays to real-time flows in LPWAN-based Industrial IoT applications. The focus of this article is on enabling real-time communications on LoRa-based networks and, for this reason, here we do not explicitly deal with power consumption aspects. This article discussed RT-LoRa performance through simulation and analysis. The implementation of RT-LoRa on COTS devices that use the SX1272 LoRa transceiver has already started and it will addressed in future work. To implement RT-LoRa, the LoRaWAN protocol of the transceiver has to be disabled, thus we are implementing the RT-LoRa MAC layer from scratch. RT-LoRa uses a star topology and assumes that the transmissions performed using different SFs are orthogonal. Future work will go further in both directions. First, suitable approaches to enable the coexistence of multiple star networks running RT-LoRa operating in the same frequency bands, such as synchronization mechanisms for the sinks, will be addressed. Second, the quasi-orthogonality will be implemented in the simulator to assess how much it affects the performance of RT-LoRa.

REFERENCES

[1] M. Luvisotto, F. Tramarin, L. Vangelista, and S. Vitturi, "On the use of LoRaWAN for indoor Industrial IoT applications," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–11, May 2018.

[2] L. Lo Bello, O. Mirabella, and N. Torrisi, "Modelling and evaluating traceability systems in food manufacturing chains," in *Proc. WETICE*, 2004, pp. 173–179.

[3] L. Beltramelli, A. Mahmood, M. Gidlund, P. Österberg, and U. Jennehag, "Interference modelling in a multi-cell LoRa system," in *Proc. WiMoB*, 2018, pp. 1–8.

[4] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, "Long-range communications in unlicensed bands: The rising stars in the IoT and smart city scenarios," *IEEE Wireless Commun.*, vol. 23, no. 5, pp. 60–67, Oct. 2016.

[5] M. Bor, J. Vidler, and U. Roedig, "LoRa for the Internet of Things," in *Proc. EWSN*, 2016, pp. 361–366.

[6] M. Rizzi, P. Ferrari, A. Flammini, and E. Sisinni, "Evaluation of the IoT LoRaWAN solution for distributed measurement applications," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 12, pp. 3340–3349, Dec. 2017.

[7] L. Tessaro, C. Raffaldi, M. Rossi, and D. Brunelli, "Lightweight synchronization algorithm with self-calibration for industrial LoRa sensor networks," in *Proc. Metrol. Ind. 4.0 IoT*, Apr. 2018, pp. 259–263.

[8] *LoRaWAN® 1.1 Specification*, LoRa Alliance, Fremont, CA, USA, Oct. 2017.

[9] *LoRaWAN Regional Parameters*, LoRa Alliance, Fremont, CA, USA, Jan. 2018.

[10] L. Leonardi, F. Battaglia, G. Patti, and L. Lo Bello, "Industrial LoRa: A novel medium access strategy for LoRa in industry 4.0 applications," in *Proc. 44th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Washington, DC, USA, Oct. 2018, pp. 4141–4146.

[11] A. Augustin, J. Yi, T. Clausen, and W. M. Townsley, "A study of LoRa: Long range & low power networks for the Internet of Things," *Sensors*, vol. 16, no. 9, p. 1466, 2016.

[12] R. Rondón, M. Gidlund, and K. Landernäs, "Evaluating Bluetooth low energy suitability for time-critical Industrial IoT applications," *Int. J. Wireless Inf. Netw.*, vol. 24, no. 3, pp. 278–290, Sep. 2017.

[13] G. Patti, L. Leonardi, and L. Lo Bello, "A Bluetooth low energy real-time protocol for industrial wireless mesh networks," in *Proc. IECON*, Florence, Italy, Oct. 2016, pp. 4627–4632.

[14] S. M. Darroudi and C. Gomez, "Bluetooth low energy mesh networks: A survey," *Sensors*, vol. 17, no. 7, p. 1467, 2017.

[15] L. Leonardi, G. Patti, and L. Lo Bello, "Multi-hop real-time communications over Bluetooth low energy industrial wireless mesh networks," *IEEE Access*, vol. 6, pp. 26505–26519, 2018.

[16] I. A. Ismaili, A. Azyat, N. Raissouni, N. Ben Achhab, A. Chahboun, and M. Lahraoua, "Comparative study of ZigBee and 6LoWPAN protocols: Review," in *Proc. ICCWCS*, 2019, pp. 1–9.

[17] E. Toscano and L. Lo Bello, "Comparative assessments of IEEE 802.15.4/ZigBee and 6LoWPAN for low-power industrial WSNs in realistic scenarios," in *Proc. WFCS*, May 2012, pp. 115–124.

[18] L. F. Schrickte, C. Montez, R. D. Oliveira, and A. R. Pinto, "Integration of wireless sensor networks to the Internet of Things using a 6LoWPAN gateway," in *Proc. SBESC*, 2013, pp. 119–124.

[19] L. F. Del Carpio, P. Di Marco, P. Skillermark, R. Chirikov, and K. Lagergren, "Comparison of 802.11ah, BLE and 802.15.4 for a home automation use case," *Int. J. Wireless Inf. Netw.*, vol. 24, no. 3, pp. 243–253, Sep. 2017.

[20] M. Syafrudin, K. Lee, G. Alfian, J. Lee, and J. Rhee, "Application of Bluetooth low energy-based real-time location system for indoor environments," in *Proc. BDIOT*, 2018, pp. 167–171.

[21] G. Alderisi, G. Patti, O. Mirabella, and L. Lo Bello, "Simulative assessments of the IEEE 802.15.4e DSME and TSCH in realistic process automation scenarios," in *Proc. INDIN*, Jul. 2015, pp. 948–955.

[22] W. Du, D. Navarro, and F. Mieyeville, "Performance evaluation of IEEE 802.15.4 sensor networks in industrial applications," *Int. J. Commun. Syst.*, vol. 28, no. 10, pp. 1657–1674, 2015.

[23] L. Leonardi, G. Patti, F. Battaglia, and L. Lo Bello, "Simulative assessments of the IEEE 802.15.4 CSMA/CA with priority channel access in structural health monitoring scenarios," in *Proc. INDIN*, 2017, pp. 375–380.

[24] U. Raza, P. Kulkarni, and M. Sooriyabandara, "Low power wide area networks: An overview," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 855–873, 2nd Quart., 2017.

[25] R. Piyare, A. L. Murphy, M. Magno, and L. Benini, "On-demand LoRa: Asynchronous TDMA for energy efficient and low latency communication in IoT," *Sensors*, vol. 18, no. 11, p. 3718, Nov. 2018.

[26] W. Ayoub, A. E. Samhat, F. Nouvel, M. Mroue, and J.-C. Prévotet, "Internet of mobile things: Overview of LoRaWAN, DASH7, and NB-IoT in LPWANs standards and supported mobility," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 2, pp. 1561–1581, 2nd Quart., 2019.

[27] 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.

[28] F. Battaglia and L. Lo Bello, "A novel JXTA-based architecture for implementing heterogenous networks of Things," *Comput. Commun.*, vol. 116, pp. 35–62, Jan. 2018.

[29] M. C. Bor, U. Roedig, T. Voigt, and J. M. Alonso, "Do LoRa low-power wide-area networks scale?" in *Proc. MSWiM*, 2016, pp. 59–67.

[30] M. Bor and U. Roedig, "LoRa transmission parameter selection," in *Proc. 13th Int. Conf. Distrib. Comput. Sensor Syst. (DCOSS)*, Jun. 2017, pp. 27–34.

[31] B. Reynders, Q. Wang, P. Tuset-Peiro, X. Vilajosana, and S. Pollin, "Improving reliability and scalability of LoRaWANs through lightweight scheduling," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1830–1842, Jun. 2018.

[32] T. Polonelli, D. Brunelli, and L. Benini, "Slotted ALOHA overlay on LoRaWAN—A distributed synchronization approach," in *Proc. IEEE 16th Int. Conf. Embedded Ubiquitous Comput. (EUC)*, Bucharest, Romania, Oct. 2018, pp. 129–132.

[33] J. Haxhibeqiri, I. Moerman, and J. Hoebeke, "Low overhead scheduling of LoRa transmissions for improved scalability," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3097–3109, Apr. 2019.

[34] M. Magno, F. A. Aoudia, M. Gautier, O. Berder, and L. Benini, "WULoRa: An energy efficient IoT end-node for energy harvesting and heterogeneous communication," in *Proc. DATE*, 2017, pp. 1528–1533.

[35] M. Rizzi, P. Ferrari, A. Flammini, E. Sisinni, and M. Gidlund, "Using LoRa for industrial wireless networks," in *Proc. WFCS*, May 2017, pp. 1–4.

[36] L. Tessaro, C. Raffaldi, M. Rossi, and D. Brunelli, "LoRa performance in short range industrial applications," in *Proc. SPEEDAM*, 2018, pp. 1089–1094.

[37] J. Haxhibeqiri, A. Karaağaç, F. Van den Abeele, W. Joseph, I. Moerman, and J. Hoebeke, "LoRa indoor coverage and performance in an industrial environment: Case study," in *Proc. ETFA*, 2017, pp. 1–8.

[38] *LoRa™ Modulation Basics*, Semtech, Camarillo, CA, USA, May 2015.

[39] O. Georgiou and U. Raza, "Low power wide area network analysis: Can LoRa scale?" *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 162–165, Apr. 2017.

[40] D. Croce, M. Gucciardo, S. Mangione, G. Santaromita, and I. Tinnirello, "Impact of LoRa imperfect orthogonality: Analysis of link-level performance," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 796–799, Apr. 2018.

[41] K. Mikhaylov, J. Petaejaejaervi, and T. Haenninen, "Analysis of capacity and scalability of the LoRa low power wide area network technology," in *Proc. 22nd Eur. Wireless Conf.*, May 2016, pp. 1–6.

[42] P. Ferrari, A. Flammini, M. Rizzi, E. Sisinni, and M. Gidlund, "On the evaluation of LoRaWAN virtual channels orthogonality for dense distributed systems," in *Proc. M&N*, Sep. 2017, pp. 1–6.

[43] M. Slabicki, G. Premsankar, and M. D. Francesco, "Adaptive configuration of LoRa networks for dense IoT deployments," in *Proc. NOMS*, 2018, pp. 1–9.

Luca Leonardi received the M.S. degree from the University of Catania, Catania, Italy, in 2016, where he is currently pursuing the Ph.D. degree with the Department of Electrical, Electronic and Computer Engineering.

Since 2015, he has been researching in the field of real-time industrial networks, Wireless Sensor and Actuator Networks, and Industrial Internet of Things.

Filippo Battaglia received the M.S. degree in electronics engineering from the University of Messina, Messina, Italy, in 2008, and the Ph.D. degree in information engineering from the University Mediterranea of Reggio Calabria, Reggio Calabria, Italy, in 2013.

He is currently a Research Collaborator with the University of Catania, Catania, Italy. His current research interests include computer vision, Internet of Things and machine-to-machine communication, and sensor networks.

Lucia Lo Bello (M'02–SM'09) received the M.S. degree in electronic engineering and the Ph.D. degree in computer engineering from the University of Catania, Catania, Italy, in 1994 and 1998, respectively.

She is a Tenured Associate Professor with the Department of Electrical, Electronic and Computer Engineering, University of Catania. She was also a Guest Professor with the University of Malardalen, Västerås, Sweden, in 2014. She has authored or coauthored over 160 technical papers in the area of industrial networks, automotive communications, real-time embedded systems, and wireless sensor networks.