A Thorough Study of LoRaWAN Performance Under Different Parameter Settings

Davide Magrin, Martina Capuzzo, and Andrea Zanella

***Abstract*—LoRaWAN is an emerging low-power wide-area network (LPWAN) technology, which is gaining momentum thanks to its flexibility and ease of deployment. Conversely to other LPWAN solutions, LoRaWAN indeed permits the configuration of several network parameters that affect different network performance indexes, such as energy efficiency, fairness, and capacity, in principle making it possible to adapt the network behavior to the specific requirements of the application scenario. Unfortunately, the complex and sometimes elusive interactions among the different network components make it rather difficult to predict the actual effect of a certain parameters setting, so that flexibility can turn into a stumbling block if not deeply understood. In this article, we shed light on such complex interactions for a single-gateway (GW) system by analyzing the effect of some built-in features and configurations, including the GW's limitations in terms of duty cycle and the number of parallel reception paths, the number of allowed retransmissions for confirmed traffic, and the preconfigured data rate used in downlink transmissions. The simulation-based analysis reveals various tradeoffs and highlights some inefficiencies in the design of the LoRaWAN standard. Furthermore, we show how significant performance gains can be obtained by wisely setting the system parameters, possibly in combination with some novel network management policies (e.g., enabling selective prioritization of downlink transmissions at the GW).**

***Index Terms*—Communication systems, communications technology, Internet of Things (IoT), vehicular and wireless technologies, wireless sensor networks.**

I. Introduction

IN THE last few years, the Internet of Things (IoT) paradigm has been attracting enormous interest from the scientific and industrial communities, thanks to the remarkable potential of the vision in which virtually every object can be remotely accessed and controlled through a connection to the Internet. Such a pervasive connectivity would enable various services in a wide array of scenarios. For example, cities could benefit from smart lighting control, a more efficient waste management, and continuous infrastructure monitoring [1]. In industrial scenarios, connected sensors can help to continuously monitor the production process, making it possible to quickly detect or even predict failures, while in the agricultural sector, the widespread collection of environmental data, such as temperature and soil moisture, can improve the quantity and quality of the soil production, while reducing costs. Health monitoring, home security, and home automation are yet other examples of possible application scenarios [2].

In general, the communication needs of such scenarios differ significantly from the classic high-throughput and low-delay requirements that have so-far driven the design of traditional communication systems. For example, long communication range and low-energy consumption are more important than high bitrates, and supporting sporadic transmission of short packets from a massive number of devices is more important than providing stable high-throughput connections to few users [3].

Numerous technologies are currently rushing to close the gap opened by this paradigm shift. One prominent solution is provided by LoRaWAN, an open standard promoted by the LoRa Alliance that defines medium access control (MAC) and network management protocols on top of the long-range (LoRa) physical (PHY) layer, which is instead proprietary of Semtech.[1]

The network topology is very simple: the radio signals transmitted by end devices (ED) are received by one (or multiple) gateways (GWs), which then forward the packets to a network server (NS) for further processing. The NS is indeed in charge of the network management, which can be performed through a set of configurable parameters. In standard mode, communication is initiated by the ED, but after each uplink transmission the devices can receive data packets or acknowledgments sent by the NS. The LoRa chipset is designed to be very energy efficient and it promises to enable up to ten years of energy autonomy for battery-powered devices. Furthermore, the transmission range has been proven to reach up to 1.5 km in urban scenarios and 15 km in rural areas [1]. The PHY layer is based on a chirp modulation and supports multiple spreading factor (SF), which make it possible to trade bitrate for range. In addition, signals transmitted with different SFs are almost orthogonal, thus potentially enabling multipacket reception at the GW. Another benefit is that the PHY operates on industrial, scientific, and medical (ISM) bands in the megahertz range, pushing down the deployment cost of the technology.

Manuscript received May 14, 2019; revised September 18, 2019; accepted October 5, 2019. Date of publication October 9, 2019; date of current version January 10, 2020. This work was supported in part by the POR FESR 2014–2020 Work Program of the Veneto Region (Action 1.1.4) through the project "Sistema domotico IoT integrato ad elevata sicurezza informatica per smart building" under Grant 10066183, and in part by MIUR (Italian Ministry for Education) through the initiative "Departments of Excellence" (Law 232/2016). *(Corresponding author: Davide Magrin.)*

D. Magrin and M. Capuzzo are with the Department of Information Engineering, University of Padova, 35131 Padua, Italy (e-mail: magrinda@dei.unipd.it; capuzzom@dei.unipd.it).

A. Zanella is with the Department of Information Engineering, University of Padova, 35131 Padova, Italy, and also with the Human Inspired Technology Center, University of Padova, 35131 Padua, Italy (e-mail: zanella@dei.unipd.it).

Digital Object Identifier 10.1109/JIOT.2019.2946487

[1] https://www.semtech.com/products/wireless-rf/lora-transceivers

The open MAC standard stimulates the creation of publicly available solutions for the back-end [4], and makes scrutiny from the scientific community possible. In addition, the LoRaWAN standard offers large flexibility in the network configuration, which is another attractive factor. Indeed, the NS can choose the SF used by the different nodes, the duration of the receive windows, the transmission/reception channels, the priority of acknowledgment and downlink data packets, and so on. By properly setting these parameters, it is hence possible to support reliable/bidirectional communications and to change the balance between communication reliability, delay, energy efficiency, and system capacity. However, while the effect of certain parameters settings can be predicted in simple scenarios, with a relatively low number of nodes, the interactions among the different mechanisms of the system become much more complex and less intuitive in the large-scale scenarios promised by the IoT paradigm.

In this article, we leverage realistic network simulations to gain insight on the real performance of the LoRaWAN technology in such scenarios, and show how even small adjustments in the MAC layer parameters can significantly affect the system performance (e.g., the packet success ratio). By doing so, we highlight some inherent issues raised by the duty cycle limitations in European ISM bands and propose some improvements to mitigate the impairments that LoRaWAN may experience at scale. Such simple ingenuities can help to increase the number of devices that can be served by a single GW, postponing the potential collapse of the network in overcrowded scenarios and reducing the network management costs created by inefficient network layouts.

The rest of this article is organized as follows. Section II provides a background on the technology, briefly describing the LoRa PHY modulation and LoRaWAN standard, pointing out their most important features and mechanisms. Section III provides an outline of the state-of-the-art on LoRaWAN simulations and performance investigations, while Section IV gives an overview of the simulation framework used in this article, describing the configurable parameters that have been considered and the simulation scenarios. In Section V, we identify bottlenecks and relevant system dynamics. Furthermore, we discuss possible improvements to the standard and show how they can significantly improve the performance. Finally, Section VI contains our conclusions and some possible avenues for future research.

II. Technology Overview

This section discusses the key features of the LoRa modulation and chipset and provides an overview of the LoRaWAN standard, highlighting the aspects that play a major role in determining the network performance.

A. LoRa Modulation

LoRa is a modulation scheme, patented by Semtech and based on chirp spread spectrum. Its design allows for long communication ranges, reaching 15 km in line-of-sight rural areas and 1.5 km in outdoor urban scenarios. The sensitivity (and, thus, the coverage) can be improved at the price of a lower bitrate, by changing the SF parameter that takes integer values from 7 to 12. Higher SF values correspond to lower transmission bitrates, but require a lower signal received power for correct reception, which turns into a longer coverage range. For each value of the SF parameter, Table I shows the associated data rate (DR) index, the nominal DR, the sensitivity level, and the transmission (TX) time.

TABLE I
Main Transmission Features for Different Values of the SF Parameters. Packet TX Time Is Specified for a PHY Payload of 32 Bytes, Explicit Header Mode, Code Rate Equal to 2, and 125-kHz Channel Bandwidth

SF	DR	Data rate [kbit/s]	Sensitivity [dBm]	TX time [s]
7	5	5.470	-130.0	0.740
8	4	3.125	-132.5	0.136
9	3	1.760	-135.0	0.247
10	2	0.980	-137.5	0.493
11	1	0.440	-140.0	0.888
12	0	0.250	-142.5	1.777

A key feature of the LoRa modulation is that packets employing different SFs are almost orthogonal: transmissions overlapping in time and frequency can still be correctly decoded by the receiver, provided that the power of the target signal is sufficiently larger than that of interferers [5], [6].

B. LoRaWAN Standard

The LoRaWAN standard [7] defines MAC and network management protocols for devices using the LoRa modulation. The network topology is a star-of-stars, formed by three kinds of devices.

1) *ED:* A peripheral node, typically a sensor or actuator that communicates only through the LoRa PHY.
2) *NS:* A centralized entity that controls the network parameters, forwards messages to applications, and sends replies to the EDs through the GW(s).
3) *GW:* An intermediate node that relays messages between EDs and NS.

EDs and GWs communicate using the LoRa modulation, while the connection between GWs and NS is realized using legacy IP technologies. Typically, the GWs are equipped with LoRa chipset that allow for the parallel reception of multiple signals. Commercial LoRa radio chipsets feature eight parallel reception paths (or chains), each of which can listen to a specific frequency and demodulate overlapping signals, exploiting the *quasi* orthogonality of the different SFs.

LoRaWAN devices operate in unlicensed ISM bands, in specific frequencies which vary based on the regional regulation: in this article, we will consider the European 868 MHz ISM frequency band, which the standard divides in four channels centered at 868.1, 868.3, 868.5, and 868.625 MHz. Bearing in mind this caveat regarding the number of channels and the different channel access regulations, the conclusions drawn in this article have however broader interest, being valid also for other regions.

As described in Table II, according to the LoRaWAN specifications the first three channels can be used both for uplink (UL) and downlink (DL) transmissions, while the

TABLE II
LoRaWAN Default Channels and DC Limitations in Europe

Frequency (MHz)	Direction	DC	Power limit (dBm)
868.1	DL, UL	1%	14
868.3	DL, UL	1%	14
868.5	DL, UL	1%	14
869.525	DL	10%	27

868.625 MHz channel is reserved for DL communications. Moreover, the channels are regulated by different limitations on transmission power and duty cycle (DC). In particular, the three bidirectional channels belong to the same sub-band and, hence, are collectively subject to a common DC limitation of 1%, so that a UL (resp. DL) transmission in any of such channels will consume the UL (resp. DL) DC budget of all three channels. Instead, the DL-only channel at 869.525 MHz belongs to a different sub-band that permits a DC of 10% and a larger transmission power. Note that the data rate (DR) parameter in Table I is generally used to indicate a pair of SF and bandwidth values. Limiting our attention to the 125 kHz wide channels, the DR is in one-to-one association with the SF, as shown in Table I.

The standard also defines three classes of EDs, namely, Class-A (*All*), Class-B (*Beacon)*, and Class-C (*Continuous*), which differ in the management of the DL transmissions: Class-A devices can receive DL packets only immediately after a UL transmission; Class-B can schedule so-called ping-interval during which they can receive DL packets; finally, Class-C can always receive packets, unless they are themselves transmitting.

In this article, we will focus on Class-A devices, which are the most popular and challenging, considering the more constrained communication capabilities. They are expected to be battery powered and can receive only during two reception windows (RX1 and RX2) that are opened, respectively, 1 s and 2 s after the end of each uplink transmission. The radio interface is then switched off until the next UL to save energy. RX2 is opened only if no DL message is successfully received during RX1. According to the standard setting, the NS can reply to a UL transmission by sending a DL packet in RX1, using the same channel and SF of the UL packet, or in RX2, using a dedicated channel (at 868.625 MHz in Europe) and SF 12 (i.e., the lowest bitrate) to maximize the coverage range. These default settings can be changed by the NS through appropriate MAC commands. Moreover, the LoRaWAN standard also provides an adaptive DR (ADR) mechanism through which the NS can control the transmission parameters of the ED to optimize the performance of either the device itself or the network as a whole.

UL transmissions can either be unconfirmed or confirmed. In the first case, a message is transmitted only once and is not expected to be acknowledged by the NS, while in the latter case, the messages are retransmitted until an acknowledgment (ACK) packet is returned by the receiver, for a maximum of m transmission attempts, with $m \in \{1, \ldots, 8\}$. Note that setting $m = 1$ coincides with the unconfirmed case (the total number of transmissions cannot exceed m in both cases), but in case of confirmed message, the receiver will be required to generate an ACK.[2]

III. Related Work

In recent years, the LoRaWAN technology has been the subject of many studies, which analyzed its performance and features with empirical measurements, mathematical analysis, and simulative tools.

Some seminal papers on LoRaWAN such as [8] and [9] test the coverage range and packet loss ratio by means of empirical measurements, but without investigating the impact of the parameters setting on the performance. Other works, such as [10], examine the impact of the modulation parameters on the single communication link between an ED and its GW, without considering more complex network configurations.

To obtain more general results, Li *et al.* [11] use a stochastic geometry model to jointly analyze interference in the time and frequency domains. It is observed that when implementing a packet repetition strategy, i.e., transmitting each message multiple times, the failure probability reduces, but clearly the average throughput decreases because of the introduced redundancy. Ferre [12] proposed the closed-form expressions for collision and packet loss probabilities and, under the assumption of perfect orthogonality between SFs, it is shown that the Poisson distributed process does not accurately model packet collisions in LoRaWAN. Network throughput, latency, and collision rate for uplink transmissions are analyzed in [13] that, using the queueing theory and considering the Aloha channel access protocol and the regulatory constraints in the use of the different sub-bands, points out the importance of a clever splitting of the traffic in the available sub-bands to improve the network performance. Bankov *et al.* [14] presented a mathematical model of the network performance, taking into account factors, such as the capture effect and a realistic distribution of SFs in the network. However, the model does not include some important network parameters, preventing the study of their effect on the network performance. A step further is made in [15], where Capuzzo *et al.* developed a model that makes it possible to consider various parameters configurations, such as the number of ACKs sent by the GW, the SF used for the downlink transmissions, and the DC constraints imposed by the regulations. In this article, however, multiple retransmissions have not been considered.

The study presented in [16] features a system-level analysis of LoRaWAN and gives significant insights on bottlenecks and network behavior in the presence of downlink traffic. However, besides pointing out some flaws in the design of the LoRaWAN medium access scheme, this article does not propose any way to improve the performance of the technology. In [17], the system-level simulations are again employed to assess the performance of confirmed and unconfirmed messages and show the detrimental impact of confirmation traffic on the overall network capacity and throughput. Here, the only proposed solution is the use of multiple GWs, without

[2]Note that from the GW perspective, ACK packets are not distinguishable from any other DL packet and, hence, are subject to the same rules and constraints.

deeply investigating the specificities of the LoRaWAN standard. In [18], a module for the ns-3 simulator is proposed and used for a similar scope, comparing the single-GW and multi-GW scenarios and the use of unconfirmed and confirmed messages. In this case, the authors correctly implement the GW's multiple reception paths, but do not take into account their association to a specific UL frequency, which usually occurs during network setup: indeed, the number of packets that can be received simultaneously on a given frequency cannot be greater than the number of reception paths that are listening on that frequency. Also in this case, the study only focuses on the performance analysis, without proposing any improvement.

Hauser and Hégr [19] and Slabicki *et al.* [20] targeted the original ADR algorithm proposed by [4], suggesting possible ameliorations. Generally, the modified algorithms yield an increase of network scalability, fairness among nodes, packet delivery ratio, and robustness to variable channel conditions. Reynders *et al.* [21] computed the optimal SFs distribution to minimize the collision probability and proposed a scheme to improve the fairness for nodes far from the station by optimally assigning SFs and transmit power values to the network nodes, in order to reduce the packet error rate.

In [22], it is shown how the use of a persistent-carrier sense multiple access (p-CSMA) MAC protocol when transmitting UL messages can improve the packet reception ratio. However, attention must be paid to the fact that having many EDs that defer their transmission because of a low value of p may lead to channel underutilization. Zucchetto and Zanella [23] investigated, via simulation, the impact of DC restrictions in low-power wide-area network scenarios, showing that rate adaptation capabilities are indeed pivotal to maintain a reasonable level of performance when the coverage range and the cell load increase. However, the effect of other parameters setting on the network performance is not considered.

In this article, we differ from the existing literature in that we target large networks with bidirectional traffic, a scenario that makes it possible to observe some unforeseen effects rising from the interaction of multiple nodes served by one single GW and NS. Furthermore, in our analysis, we examine one by one the role played by the configurable network parameters, as detailed in Section IV-A, thus highlighting some pitfalls that can affect the network performance. We then propose possible counteractions that require some small changes at the MAC layer, and we evaluate their effectiveness in some representative scenarios. As a side result, we enriched the ns-3 `lorawan` module with new functionalities.

IV. Simulation Setup and Scenarios

The analysis carried out in this article leverages the ns-3 `lorawan` module described in [6] and [24]. The module has been improved in order to support various network configurations and test the proposed ameliorations, with the aim of gaining a deeper understanding of the role played by each configurable parameter and identifying unforeseen behaviors. The module supports both confirmed and unconfirmed messages, permits the configuration of multiple network parameters, and implements a realistic model of the GW chip, accounting for the eight available parallel reception paths. In the following, we give more details on the available functionalities that will be leveraged to analyze the network.

A. Available Network Settings

Next, we give a brief introduction to the network configuration options that are available in the simulator and that make it possible to control the behavior and features of both the GW and the EDs.

1) *GW DC:* In the simulator, we have the opportunity of turning on or off the DC restriction at the GW to analyze its impact on the network performance.
2) *Transmission/Reception Priority:* LoRaWAN GWs cannot receive and transmit simultaneously. In commercial GWs, if there is the need to send an ACK, the reception of any incoming signal will be immediately interrupted to start the DL transmission. In this article, we call this standard behavior "transmission (TX) prioritization." In addition to this mechanism, here, we also consider the RX prioritization option, according to which DL transmissions are performed only if the GW is not engaged in any packet reception. In case, the GW is receiving a UL message at the time when it should be sending an ACK, it will continue the reception procedure and postpone the DL transmission to the second receive window, if available. If a reception is also occurring when the second receive window opens, the ACK is definitively dropped. This feature could easily be implemented in future LoRa chips, only requiring minimal changes to the behavior of the NS and GWs.
3) *Sub-Band Prioritization:* The LoRaWAN standard requires that RX1 is opened on the same channel where the corresponding UL was received, while RX2 is opened on a dedicated DL channel, which in Europe also features more lenient DC restrictions (10% instead of the 1% allowed on the other channels). In the simulator, we have enabled a mode that switches this setting, making it possible to open RX1 on the dedicated DL channel, and RX2 on the channel used for the UL communications. The effect of this trick will be illustrated in Section V-D.
4) *ACK DR:* The LoRaWAN specifications recommend that ACKs transmitted on RX1 should use the same SF for the UL transmission, while the transmissions on RX2 use the lowest available DR (SF = 12). To explore other options, the simulation module has been modified to enable the use of higher DRs on both the reception windows. This setting involves a tradeoff between the robustness and efficient use of the available DC and time resources. Note that such an option can actually be implemented in LoRaWAN through a dedicated MAC command.
5) *Number of Transmission Attempts:* For confirmed traffic, the maximum number m of transmission attempts for the

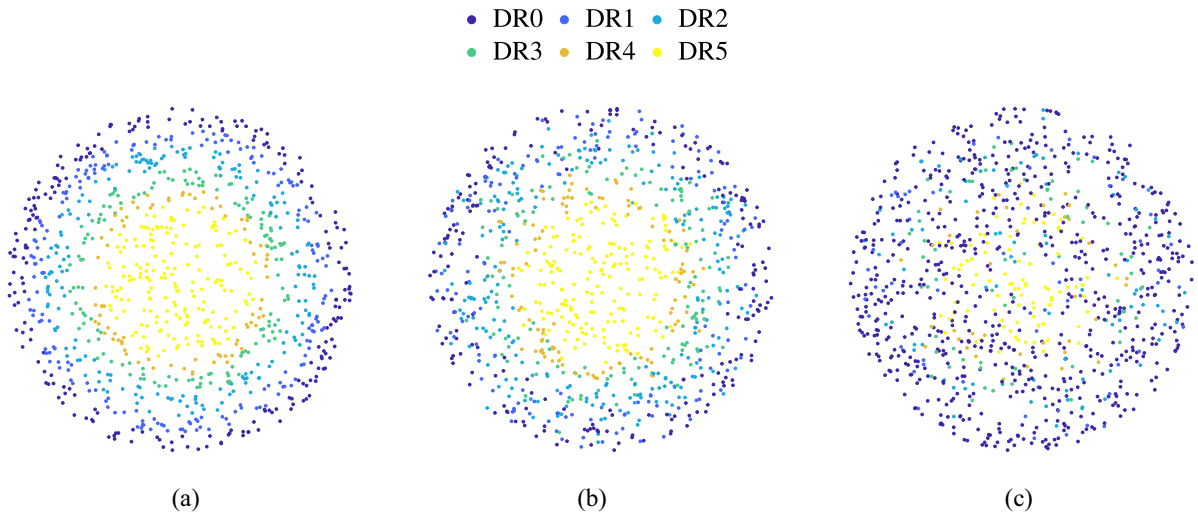


Fig. 1. Distribution of DRs for different channel models. (a) Baseline. (b) Shadowing. (c) Shadowing and buildings.

same message is configurable, and can take values in the set {1, 2, 4, 6, 8}.

6) *Full-Duplex GW (FDGW):* As mentioned, currently GWs cannot transmit and receive simultaneously. However, it might be interesting to investigate the potential performance gain obtained by implementing an FDGW. This functionality may be realized by co-locating two GWs or combining a GW with a simple LoRa chipset to be used for transmissions only, leaving the GW free to receive messages. Note that because of this, whenever the FDGW configuration is employed, the distinction between RX and TX prioritization loses meaning. In order to test this functionality, we added a new mode to the `lorawan` module in the ns-3 simulator that allows for ideal full-duplex communication.
7) *Number of Reception Paths:* The number r of parallel reception paths in the GW is a parameter that can be toggled in the simulator. Beside the standard value $r = 8$, we also considered the values $r = 3$ and $r = 16$ to study how the parallel reception capabilities of the GW can affect the overall system performance.

B. Reference Scenarios

We considered two main simulation scenarios. Since we are interested on the optimization of the MAC layer parameters, we assume a single GW serving multiple EDs, which generate packets periodically, with equal period but random phases. Furthermore, the traffic generated by the devices can be either confirmed, unconfirmed, or mixed, i.e., with half of the devices requiring ACKs and the other half sending unconfirmed packets.

In the first scenario, we assume that EDs are randomly distributed within the coverage range of the GW, and we only consider path loss.

The second scenario consists of a more *realistic urban deployment*, where EDs are randomly located outside or inside buildings having different height and wall width, following the model described in [25]. Here, the channel propagation is affected by path loss, spatially correlated shadowing, and attenuation due to the presence of buildings, as described in [6]. To obtain a realistic setup, we consider the traffic model described in the mobile autonomous reporting (MAR) reports [25], according to which the devices send packets with periods that vary from 30 min to 24 h, as described in Table III. The number of devices is also varied to estimate the capacity (in terms of the number of active devices) that can be supported by a GW in a realistic scenario.

TABLE III
INTERARRIVAL TIMES IN REALISTIC SIMULATIONS

Inter-arrival time	% of devices
1 day	40%
2 hours	40%
1 hour	15%
30 minutes	5%

To ease the interpretation of the results, we neglect short-term fading phenomena that may affect the received signal power, also considering that the chirp modulation is rather robust to multipath fading.

The effects of the channel model on the distribution of the SFs (and, thus, of DRs) can be observed in Fig. 1, where the dots show the position of randomly placed EDs around the central GW, while the colors are used to represent the bitrate of each device, i.e., its DR value (see Table I). The bitrate is the highest permitted by the signal received power at the GW, according to the sensitivity thresholds in Table I. Note that the rate distribution becomes more erratic in the presence of long-term shadowing and wall attenuation factors that affect the propagation.

C. Performance Metrics

A packet transmission at the PHY layer can have five possible outcomes.

1) *Success (S):* The packet is correctly received by the GW.

2) *Lost because under sensitivity (U):* The packet arrives at the GW with power lower than the sensitivity, and the GW cannot lock on it.
3) *Lost because of interference (I):* The packet is correctly locked-on by the GW, but its reception fails because of the interference from overlapping packets with enough power to disrupt signals orthogonality.
4) *Lost because of saturated receiver (R):* As mentioned, a GW usually has multiple reception paths, each configured to listen to a specific channel frequency. A packet transmitted on a certain channel is lost because of saturated receiver if it gets disregarded by the GW because all the reception paths for that channel are already locked receiving some other signals. Note that such a packet would be correctly received, if not for the saturation of the receiver. In other words, this error event would never occur if the GW had a sufficient number of reception paths for each frequency channel.
5) *Lost because of GW transmission (T):* The packet reception gets disrupted by the transmission of a DL packet (which could either be ongoing at the packet arrival time or started during the packet reception, in case, the GW gives priority to transmission).

In the case of unconfirmed traffic, we label a packet as *successful* when it is successfully received at the GW that, in turn, forwards it to the NS through a reliable connection. For confirmed traffic, we distinguish two cases, depending on whether the DL packets carry information (e.g., the UL packet is a query to the NS, and the corresponding DL packet is the reply), or are just an ACK used to stop retransmissions of the UL packets. In the first case, the transmission is successful when both the UL and the successive DL packet are successfully received by the NS and ED, respectively, within the available transmission attempts. In the second case, instead, we assume that a transmission is successful if at least one of the generated UL packets is delivered to the NS, irrespective of whether the ACK is received by the device.

Accordingly, we define two performance metrics.

1) *Confirmed packet success ratio (CPSR):* Probability that both the confirmed UL packet and the corresponding DL packet are correctly received in one of the available transmission attempts.
2) *Uplink packet delivery ratio (UL-PDR):* Probability that a UL packet is correctly received (whether or not the ACK is requested), in at least one of the available transmission attempts (which are just 1 for unconfirmed traffic, and *m* for confirmed traffic).

V. Performance Evaluation

In this section, we first provide the baseline for our performance analysis considering the default settings, which reveals some issues with the current LoRaWAN standard. Then, we study the impact of the configurable parameters, and finally validate the effectiveness of the proposed improvements using the simulator described in Section IV.

TABLE IV
Default Parameter Settings

Parameter	Value
GW DC	On
TX/RX priority	TX priority
Sub-band prioritization	Off
RX2 data rate	Lowest (SF 12)
Number of TX attempts m	8
Full-duplex GW	No
Number of reception paths r	8

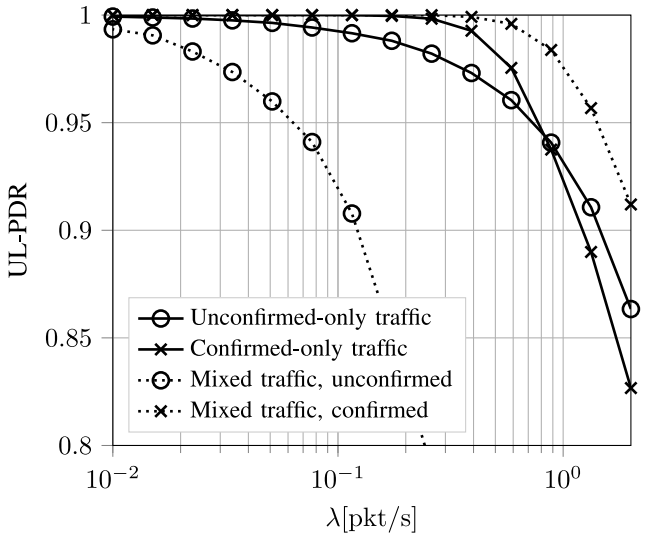


Fig. 2. Baseline UL-PDR performance for different kinds of traffic.

A. Baseline Performance Analysis

To begin with, it is interesting to compare the performance attained by confirmed/unconfirmed traffic in the mixed and homogeneous scenarios, for the same offered traffic at the application level. For these results, we used the default settings, as described in Table IV. The solid lines in Fig. 2 show the UL-PDR for the confirmed-only and unconfirmed-only cases (crossed and circle markers, respectively), while the dashed lines refer to the performance experienced by the two types of traffic sources in the mixed scenario. It is apparent that the mixture of confirmed and unconfirmed traffic sources favor the first class of sources, but penalizes much more severely the latter, with respect to the corresponding homogeneous traffic cases.

Focusing on the homogeneous traffic scenarios, we can see that the use of confirmed traffic maximizes the UL-PDR index up to an aggregate traffic load of almost $\lambda = 0.8$ pkt/s at the application layer (not including retransmissions). Beyond this point, it is more convenient to use unconfirmed-only communications. The reason behind this behavior becomes apparent in Fig. 3, which reports the fraction of packet losses caused by different events (I, R, and T; see Section IV-C) for the two homogeneous scenarios. The results are obtained for an offered traffic of $\lambda = 0.8$ pkt/s, for which the UL-PDR is the same for both the homogeneous scenarios. We can observe that with only unconfirmed traffic packet losses are mainly due to the interference (I) produced by multiple UL transmissions.

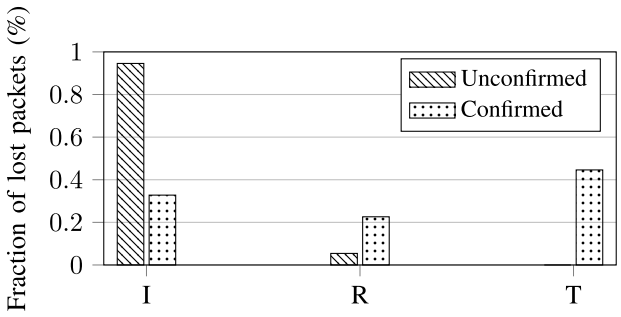


Fig. 3. PHY outcomes for traffic achieving the same UL-PDR.

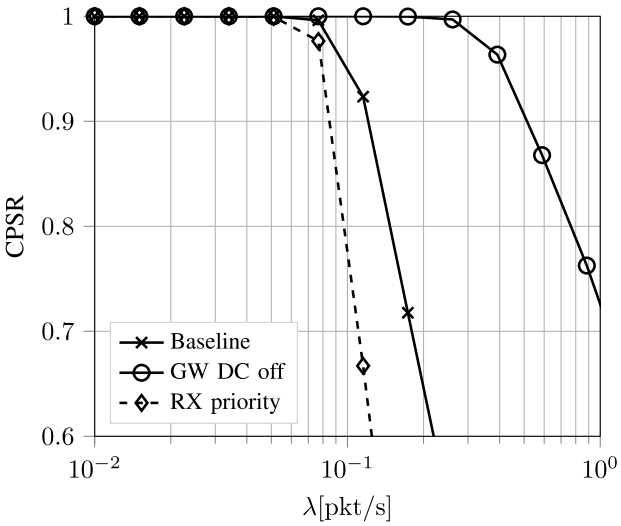


Fig. 4. CPSR of a network with only confirmed traffic sources.

Instead, confirmed traffic (with $m = 8$), in addition to losses caused by interference, also suffers from other impairments, such as the saturation of the GW's reception paths (R), and collision with ACKs (T), which plays a major role among the causes of failure. Therefore, confirmed traffic may enhance the data collection capabilities of the network as long as the overall load is light, but it can yield significant degradation of the PHY layer performance for higher loads, which in turn impairs scalability.

In the remainder of this section, we will investigate the impact that the parameters introduced in Section IV-A can have on the performance metrics and explore some simple precautions that can significantly improve both performance and fairness in LoRaWAN.

B. Gateway DC

The impact of the DC restriction at the GW is visible only when confirmed traffic is required by the EDs. The solid line with cross markers in Fig. 4 shows the baseline CPSR performance obtained in the case of only confirmed traffic. The solid line with circle markers, instead, gives the CPSR that can be obtained by removing the DC constraint at the GW. Compared with the two curves, it is clear that the DC restriction at the GW represents a severe bottleneck in terms of CPSR since successfully received UL packets may not be acknowledged by the NS in the due time because of the DC limitations of the GW. Furthermore, the missed ACKs

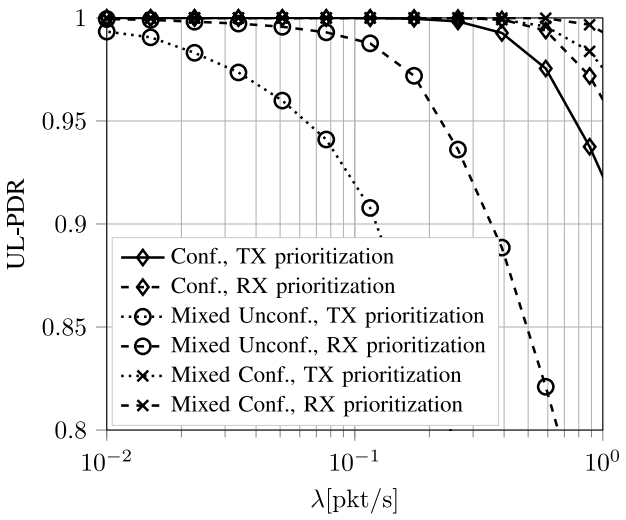


Fig. 5. UL-PDR performance for confirmed and mixed traffic when RX or TX is prioritized.

exacerbate the UL traffic load, triggering retransmissions of otherwise successfully delivered UL packets.

C. Priority of Transmission Over Reception

The effects of reception (RX) prioritization at the GW have been investigated both in terms of CPSR (Fig. 4) and UL-PDR (Fig. 5). It is worth to observe that RX prioritization can be implemented at the GW by simply avoiding the transmissions of DL packets if at least one of the eight parallel receive chains is occupied.

Fig. 4 shows that giving priority to RX yields a CPSR loss. In fact, as λ increases, the number of UL packets that are successfully received by the GW increases more rapidly than in the default case where TX is prioritized, and the probability that the GW is in the reception state quickly approaches 1, thus preventing the GW from transmitting ACKs. This, in turn, triggers packet retransmissions from the devices. On the other hand, as shown in Fig. 5, in mixed traffic scenarios, the RX prioritization improves the performance of both confirmed and unconfirmed traffic sources in terms of UL-PDR. In summary, giving priority to RX at the GW makes it possible to receive more UL packets, but this can yield to congestion in the DL channel.

More generally, DL packets could be marked by the NS based on their importance for the ED (which can either be explicitly signaled through a MAC header bit or inferred by the NS based on the application that is generating the data flow). If ACKs are required, the DL packet could be marked as prioritized over reception, and immediately sent by the GW. If, on the other hand, confirmation is merely used to stop retransmissions and the ED is only interested in maximizing its UL-PDR, then ACKs could be marked as low priority, and the GW would send them only if idle.

D. ACK Variations

This section analyzes the effect of two variations to the standard ACK mechanisms, named sub-band swapping and ACK DR, that try to alleviate the bottleneck due to the DC

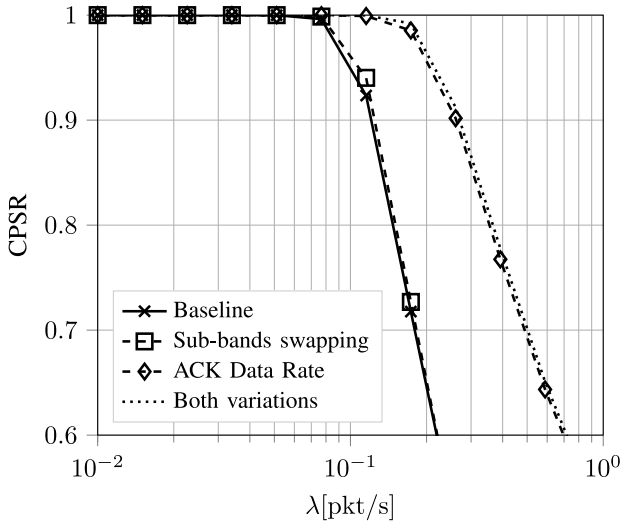


Fig. 6. Effect of improvements on CPSR.

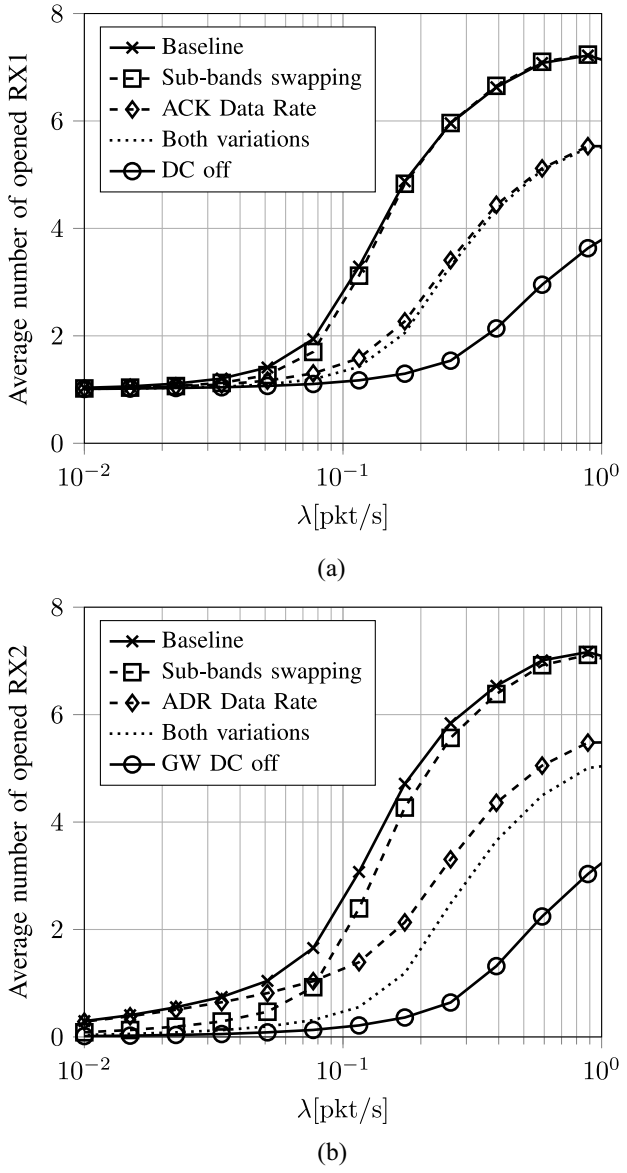


Fig. 7. Effects of the proposed ACK improvements on the average number of opened RX1 and RX2 windows.

restrictions at the GW and improve the system performance in terms of throughput and energy efficiency.

1) Sub-Band Swapping: As mentioned before, RX1 is always opened on the same sub-band used for the UL transmission, while RX2 is opened on a sub-band reserved to DL transmission, whose DC is 10%. Therefore, ACKs sent in RX1 will compete with other UL transmissions, generating and suffering interference, and can rapidly consume the 1% DC of that sub-band. We hence explored whether any benefit could come from swapping the sub-bands used for RX1 and RX2: we hence implemented a sub-band swapping scheme, according to which RX1 opens in the DL-reserved sub-band, while RX2 opens in the shared sub-band used for the UL transmissions.

2) ACK Data Rate: By default, LoRaWAN devices use the highest available SF (and thus the lowest DR) in RX2, in order to increase the probability that the downlink packet is received correctly. However, this can be detrimental, since longer transmission times of ACKs will rapidly consume the DC budget at the GW. To study which effect is dominant, we implemented the "ACK DR scheme," where all DL transmissions are always performed at the same DR used for the corresponding UL transmission.

In Fig. 6, we report the CPSR achieved by using the default setting (solid line with cross markers), each one of the ACK improvement schemes (dotted lines with square and diamond markers, respectively), and both the improvements together (dashed line without markers). We can observe that the sub-band swapping has a very marginal (yet positive) impact in terms of CPSR, which implies that the interference produced by UL transmissions on DL reception is not very significant. Conversely, the use of the same DR in all receive windows brings a significant gain in terms of CPSR over the baseline. We can hence conclude that the use of the lowest DR in RX2 can severely limit the performance of the system, in particular, when the missed reception in RX1 is not due to channel impairments, but rather to DC limitations of the GW in that sub-band.

A better strategy to provide efficient and reliable DL transmissions is hence to implement independent rate-adaptation strategies on all DL sub-bands, rather than following the very conservative policy of retransmitting at the basic rate to increase robustness, but at the cost of lower spectral efficiency.

The two ACK improvement schemes also have a positive impact on the energy consumption of the EDs. Indeed, the sub-band swapping mechanism makes it possible to return a larger number of ACKs in RX1, thanks to the looser DC constraint of the DL-reserved sub-band, thus avoiding the need to open RX2. This effect can be observed in Fig. 7, which shows the average number of times RX1 (above) and RX2 (below) is opened by the EDs, with the max number of retransmissions set to $m = 8$. The gain, however, tends to vanish as the traffic increases, since both sub-bands will then be used to return ACKs. We can also notice that by using the same DR in both receive windows, we significantly reduce the average number of opened receive windows per

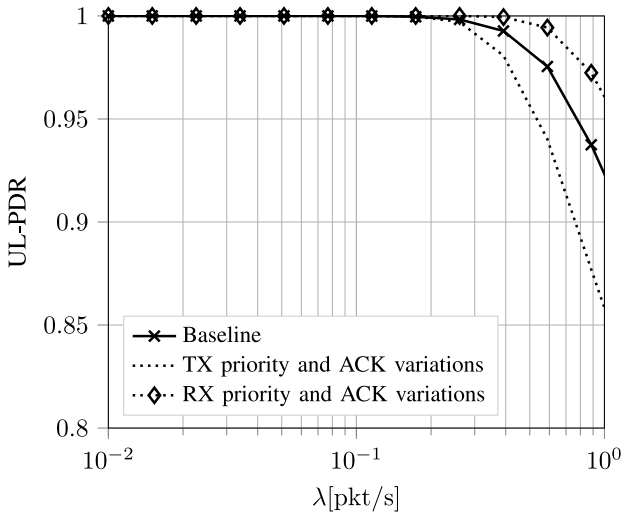


Fig. 8. UL-PDR performance in case of only confirmed traffic, when ACK improvements and RX priority are applied.

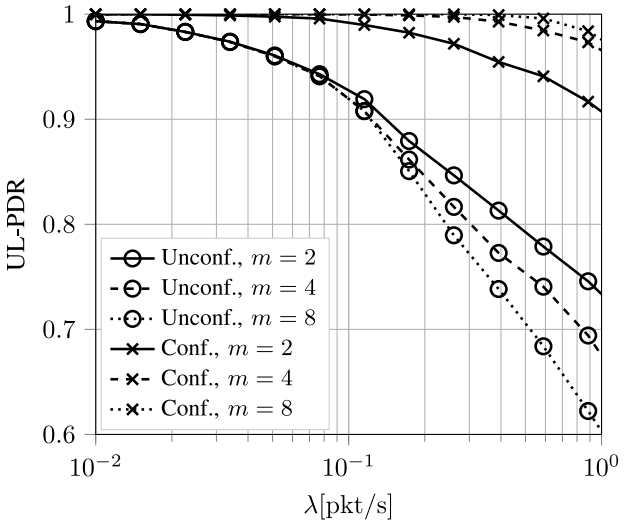


Fig. 9. UL-PDR for mixed traffic, different values of *m*.

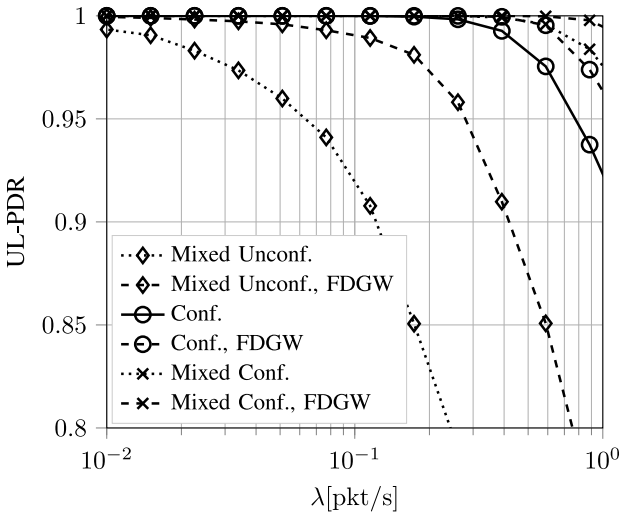


Fig. 10. Effects of full-duplex GW on UL-PDR.

transmission, also for a relatively high traffic. Indeed, transmitting DL packets at a higher rate contributes to alleviate the DC impairments, allowing the GW to serve more devices. In turn, this reduces the number of retransmissions and, consequently, the number of RX1 and RX2 that need to be opened.

The effect of the proposed ACK variations on the UL-PDR metric is depicted in Fig. 8 for a network of only confirmed traffic sources. In this case, both sub-band swapping and ACK DR mechanisms yield worse performance, when the GW adopts the standard TX prioritization policy. This is easily explained if we consider the type of DL traffic that a saturated network (i.e., one where the ACKs queues are always full) will generate when the proposed improvements are turned on and off: in the default case, long DL transmissions using low DRs will be followed by long waiting times due to the DC. During these silence periods, the GW will be forced to listen to the network, resulting in an improved UL-PDR performance. If, on the other hand, the GW sends short DL packets, it can do so more frequently, and in turn lose more UL packets because of DL transmissions. This behavior, however, can be counteracted by prioritizing RX over TX: as Fig. 8 shows, with this configuration we get the best of both worlds, achieving simultaneously the improvements on UL-PDR and the energy-saving benefits obtained with the sub-band swapping and use of the "ACK DR" schemes.

One final consideration regards networks in which some devices are interested in the UL-PDR metric, while others need to maximize their CPSR: in this case, the considerable improvement in CPSR brought by the proposed ACK variations would yield a slight loss in UL-PDR, which could be further reduced by implementing the dynamic transmission prioritization scheme proposed in Section V-C.

E. Number of Transmission Attempts

Our results showed that increasing the maximum number *m* of transmission attempts improve the CPSR by 5%–10% (though with sharply diminishing returns as *m* grows larger). On the other hand, as we can see in Fig. 9, smaller values of *m* can slightly improve the fairness in terms of UL-PDR in mixed traffic scenarios. In particular, at $\lambda = 1$ pkt/s, choosing $m = 4$ instead of $m = 8$ does not change significantly the UL-PDR for confirmed traffic, but yields an improvement in the UL-PDR of unconfirmed traffic, proving the sensitivity of the network performance to the setting of this parameter.

F. Full-Duplex Gateway

The impact of an FDGW scheme described in Section IV-A is shown in Fig. 10, where UL-PDR performance is reported both for the standard GW configuration and for the FDGW. As expected, this solution achieves a rather marked gain in terms of UL-PDR performance.[3]

[3]Note that when FDGW is employed, packets that are being received by the GW are still lost if a transmission on that same channel is performed due to the strong interference.

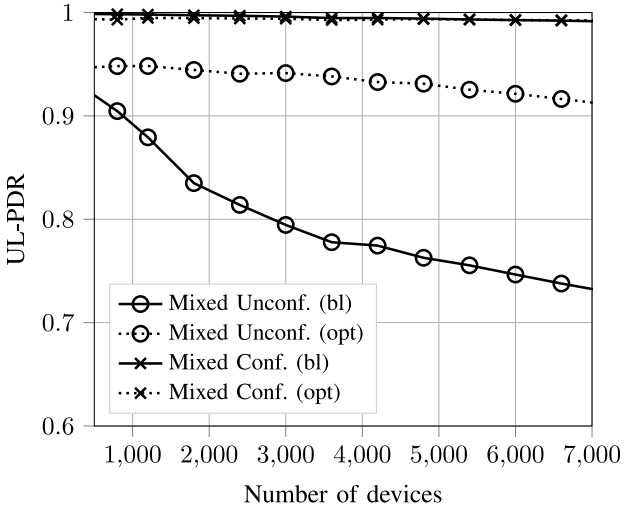
G. Number of Available Receive Paths

Our simulation results (not reported here due to space constraints) showed that UL-PDR performance increases with the number of receive paths in the GW, but with diminishing returns after eight reception paths, as interference still causes a relevant portion of locked-on packets to be lost. Having a chip with only three parallel receive paths, on the other hand, may enable cheaper GWs, and yield a slightly lower but still appreciable performance. The number of available receive paths, on the other hand, has no impact on the CPSR, for which the bottleneck is the DL channel due to the DC constraints: by the time the additional receive paths can make a difference in the reception probability of UL packets, DL channels at the GW will already be saturated, limiting the maximum achievable CPSR.

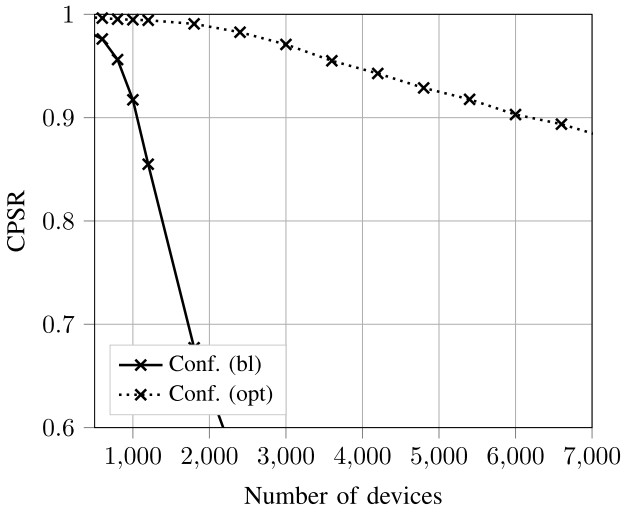
H. Best Configurations in Realistic Scenarios

A final simulation campaign had the objective of estimating the impact of the settings described in Section IV-A on the performance of a sensor network in a realistic urban scenario. Fig. 11(a) and (b) shows the UL-PDR and CPSR performance obtained in this scenario with baseline and "optimal" parameter settings. The optimal setting actually depends on the target performance index. More specifically, to optimize the UL-PDR, the GW was set in the RX prioritization mode and the maximum number of transmissions was set to $m = 4$, while to optimize the CPSR, we set $m = 8$ and the GW in TX prioritization mode. In both cases, the GW applies sub-band swapping and ACK DR mode. Note that the results have been plotted against the number of devices in the cell, in order to give an idea of the gain in network size that is achievable through a clever setting of the network's operational parameters.

Fig. 11(a) shows how the UL-PDR can improve by using the proposed settings configuration, and accommodate up to four times the number of unconfirmed devices that it would be possible to serve with standard settings. Similarly, Fig. 11(b) shows that the number of devices that can be provided a CPSR larger than 0.95 doubles when the proposed variations are applied.



(a)



(b)

Fig. 11. Simulation results for a realistic scenario. (a) UL-PDR performance, baseline (bl), and optimal (opt) parameter configuration. (b) CPSR performance, baseline (bl), and optimal (opt) parameter configuration.

TABLE V
SENSITIVITY COMPARISON

SF	GW Sensitivity (dBm)	ED Sensitivity (dBm)
7	-130.0	-124.0
8	-132.5	-127.0
9	-135.0	-130.0
10	-137.5	-133.0
11	-140.0	-135.0
12	-142.5	-137.0

I. Additional Observations

1) EDs Locking on Uplink Packets: The LoRaWAN standard does not allow direct transmission between EDs. Nonetheless, the simulation outcomes revealed that when an ED opens its receive window to listen for DL packets, the device can actually lock onto a message sent in UL by a second ED. Experimental trials with real LoRa devices confirmed this incorrect behavior. This is due to the fact that the same preamble is used in both UL and DL transmissions, so that a receiver is not aware of the transmission source until the packet is completely received and inspected. At the same time, an ED can also lock on a DL message intended for another receiver, experiencing, thus, a waste of energy and time, as the packet will eventually be discarded. The problem of EDs locking on UL messages could be easily avoided by using different preambles for UL and DL transmissions: in this way, the receiver would completely avoid the reception of UL packets and could return to sleep mode for the remaining duration of the ED receive window.

2) Sensitivities Asymmetry: Table V shows the sensitivity requirements of GW and EDs. We can observe that the requirements for EDs are more relaxed, mainly to reduce the

TABLE VI
PROPOSED SETTING CONFIGURATIONS TO INCREASE NETWORK PERFORMANCE FOR DIFFERENT TYPES OF TRAFFIC

Parameter setting	Unconfirmed	Confirmed		Mixed	
	UL-PDR	UL-PDR	CPSR	UL-PDR	CPSR
GW prioritization	RX	RX	TX	RX	TX
Sub-band swapping	–	–	Yes	Yes	Yes
ACK data rate	–	Yes	Yes	Yes	Yes
Number of transmissions m	1	8	8	4	8

manufacturing cost. However, the gap between the capabilities of the two kinds of device causes an asymmetric coverage range between UL and DL transmissions, and it may happen that the SF used by an ED to reach the GW is not sufficient to correctly deliver the return packet to the same ED because of its worse sensitivity: as an example, a UL transmission that arrives at the GW with a power of -128 dBm and SF $= 7$ may generate a DL transmission that, assuming a symmetric channel, will arrive with the same power at the ED, and thus below its sensitivity. While such an asymmetry is not an issue when all nodes send unconfirmed UL traffic, it may become a problem in case of confirmed traffic, since some EDs could be prevented from receiving a DL packet in the first receive window, in which the NS uses the same SF and carrier frequency of the UL message, forcing the systematic opening of the second receive window.

This problem can be easily mitigated at the NS by checking that the reception power of the packets coming from one ED is not in an interval such that the situation described above can occur (assuming symmetric channel). In such a case, the NS can use the appropriate MAC commands to inform the device that future DL transmissions in RX1 will use a higher SF than that one used in the UL.[4]

VI. CONCLUSION

In this article, we presented a systematic analysis of the impact of the tunable parameters in LoRaWAN some of which proved to be particularly meaningful. We observed that with a standard configuration, the presence of confirmed traffic sources can significantly degrade the performance of unconfirmed traffic, due to the additional interference generated by DL (ACK) transmissions. Considering only confirmed traffic, instead, the most critical factor appeared to be the DC constraint of the GW, which throttles the DL channel that soon becomes the bottleneck of the system in the presence of bidirectional flows. On the other hand, giving priority to RX at the GW can bring benefits to the network when most of the traffic is in the UL direction, but can be detrimental when the GW needs to transmit frequently in DL. Furthermore, we have shown that a rate adaptation strategy and the swapping of the frequencies to be employed in the two DL opportunities can significantly reduce the time the EDs spend in the reception state, thus improving their energy efficiency. Conversely, simulations showed that other system parameters, such as the maximum number of transmission attempts and the number of parallel received paths, appear to be already well configured and dimensioned. Thanks to the analysis described above, we can identify the best settings configurations for various combinations of traffic type and metric of interest, as summarized in Table VI.

Note that this analysis has been carried out by considering a single GW. In a multi-GWs scenario, UL transmissions are successful when correctly received by at least one GW. Similarly, DL packets can be transmitted, in principle, by any GW in the coverage range of the target receiver. Therefore, we expect better performance for both UL and DL traffic. It is also expected that most of the observations made in the single GW scenario will hold for the multi-GW case. Incrementing the number of GWs, and hence the number of data collecting devices in the network, however, is expected to diminish the gains brought by solutions that attempt at improving the GW's capacity in the UL, e.g., FDGW and RX priority. Finally, we note that the degrees of freedom in system configuration grows considerably with the number of GWs. Although not all configurations are meaningful, the optimization space becomes significantly larger, making the analysis of the possible interactions among different configurations even more complex. A discussion about proper parameter setting in a multi-GW case, therefore, requires an in-depth analysis, which we left for future work.

Overall, the analysis conducted in this article shows how the interplay among the different tunable parameters provided by the system is often subtle and difficult to predict, calling for the development of efficient system design and configuration tools, and motivating the investigation of possible side effects of new policies before deployment.

[4]The simulations, in this article, were performed by setting the ED's SF based on the ED sensitivities, as to avoid this asymmetry problem.

REFERENCES

[1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.

[2] D. Bandyopadhyay and J. Sen, "Internet of Things: Applications and challenges in technology and standardization," *Wireless Pers. Commun.*, vol. 58, no. 1, pp. 49–69, May 2011.

[3] M. Polese, M. Centenaro, A. Zanella, and M. Zorzi, "M2M massive access in LTE: RACH performance evaluation in a smart city scenario," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, 2016, pp. 1–6.

[4] *The Things Network*. Accessed: May 10, 2019. [Online]. Available: https://www.thethingsnetwork.org

[5] 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.

[6] D. Magrin, M. Centenaro, and L. Vangelista, "Performance evaluation of LoRa networks in a smart city scenario," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–7.

[7] *LoRaWAN™ 1.1 Specification*, LoRa Alliance, Grenoble, France, Oct. 2017.

[8] J. Petäjäjärvi, K. Mikhaylov, A. Roivainen, T. Hanninen, and M. Pettissalo, "On the coverage of LPWANs: Range evaluation and channel attenuation model for LoRa technology," in *Proc. Int. Conf. ITS Telecommun. (ITST)*, Copenhagen, Denmark, Dec. 2015, pp. 55–59.

[9] A. J. Wixted, P. Kinnaird, H. Larijani, A. Tait, A. Ahmadinia, and N. Strachan, "Evaluation of LoRa and LoRaWAN for wireless sensor networks," in *Proc. IEEE SENSORS*, Orlando, FL, USA, Oct./Nov. 2016, pp. 1–3.

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

[11] Z. Li, S. Zozor, J.-M. Brossier, N. Varsier, and Q. Lampin, "2D time-frequency interference modelling using stochastic geometry for performance evaluation in low-power wide-area networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–7.

[12] G. Ferre, "Collision and packet loss analysis in a LoRaWAN network," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Aug./Sep. 2017, pp. 2586–2590.

[13] R. B. Sørensen, D. M. Kim, J. J. Nielsen, and P. Popovski, "Analysis of latency and MAC-layer performance for class A LoRaWAN," *IEEE Wireless Commun. Lett.*, vol. 6, no. 5, pp. 566–569, Oct. 2017.

[14] D. Bankov, E. Khorov, and A. Lyakhov, "Mathematical model of LoRaWAN channel access with capture effect," in *Proc. Annu. Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Montreal, QC, Canada, Oct. 2017, pp. 1–5.

[15] M. Capuzzo, D. Magrin, and A. Zanella, "Mathematical modeling of LoRa WAN performance with bi-directional traffic," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018, pp. 206–212.

[16] A.-I. Pop, U. Raza, P. Kulkarni, and M. Sooriyabandara, "Does bidirectional traffic do more harm than good in LoRaWAN based LPWA networks?" in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

[17] 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.

[18] B. Reynders, Q. Wang, and S. Pollin, "A LoRaWAN module for ns-3: Implementation and evaluation," in *Proc. Workshop ns-3*, Jun. 2018, pp. 61–68.

[19] V. Hauser and T. Hégr, "Proposal of adaptive data rate algorithm for LoRaWAN-based infrastructure," in *Proc. Int. Conf. Future Internet Things Cloud (FiCloud)*, Prague, Czech Republic, Aug. 2017, pp. 85–90.

[20] M. Slabicki, G. Premsankar, and M. Di Francesco, "Adaptive configuration of LoRa networks for dense IoT deployments," in *Proc. 16th IEEE/IFIP Netw. Oper. Manag. Symp. (NOMS)*, Taipei, Taiwan, Apr. 2018, pp. 1–9.

[21] B. Reynders, W. Meert, and S. Pollin, "Power and spreading factor control in low power wide area networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, May 2017, pp. 1–6.

[22] N. Kouvelas, V. Rao, and R. R. V. Prasad, "Employing p-CSMA on a LoRa network simulator," *CoRR*, vol. abs/1805.12263, pp. 1–6, May 2018. [Online]. Available: http://arxiv.org/abs/1805.12263

[23] D. Zucchetto and A. Zanella, "Uncoordinated access schemes for the IoT: Approaches, regulations, and performance," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 48–54, Sep. 2017.

[24] M. Capuzzo, D. Magrin, and A. Zanella, "Confirmed traffic in LoRaWAN: Pitfalls and countermeasures," in *Proc. Annu. Mediterr. Ad Hoc Netw. Workshop (Med-Hoc-Net)*, Jun. 2018, pp. 1–7.

[25] "Cellular system support for ultra-low complexity and low throughput Internet of Things (CIoT), V13.1.0," 3GPP, Sophia Antipolis, France, Rep. 45.820, Nov. 2015.