Current LoRa networks including those following the LoRaWAN specification use the primitive ALOHA mechanism for media access control due to LoRa’s lack of carrier sense capability. From our extensive measurements, the channel activity detection feature that was recently introduced to LoRa for energy-efficiently detecting preamble chirps can also detect payload chirps reliably. This sheds light on an efficient carrier-sense multiple access protocol that we refer to as LMAC for LoRa networks. This article presents the designs of three advancing versions of LMAC that respectively implement carrier-sense multiple access, and balance the communication loads among the channels defined by frequencies and spreading factors based on the end nodes’ local information and then additionally the gateway’s global information. Experiments on a 50-node lab testbed and a 16-node university deployment show that, compared with ALOHA, LMAC brings up to 2.2× goodput improvement and 2.4× reduction of radio energy per successfully delivered frame. Thus, should LoRaWAN’s ALOHA be replaced with LMAC, network performance boosts can be realized.

CCS Concepts: • Networks → Wireless access points, base stations and infrastructure;

Additional Key Words and Phrases: LoRa, CSMA, LoRaWAN, LPWAN, chirp spread spectrum (CSS), ALOHA, wireless, IoT

ACM Reference format:

1 INTRODUCTION

Low-power wide-area networks (LPWANs) form an important class of wireless networks for geographically distributed Internet-of-Things (IoT) end devices. LPWANs’ long-range communication capabilities will increase the degree of connectivity of IoT and enable deep penetration of networked intelligence into urban territories (e.g., wide areas, buildings, and underground structures) that have challenged the existing low-power short-range wireless technologies. Among various LPWAN technologies (including NB-IoT and SigFox), LoRaWAN [12], an open data link layer
specification based on the LoRa physical (PHY) layer, offers the advantage of using license-free ISM bands, low-cost end devices, and the independence from managed cellular infrastructures.

Currently, the LoRa networks, including those organized in an ad hoc manner or by following the LoRaWAN specification, adopt the ALOHA media access control (MAC) mechanism for each logic communication channel defined by radio frequency and the spreading factor (SF) of the chirp spread spectrum (CCS) modulation. The primitive ALOHA, although enabling simple network implementation, is not competent for scaling with the communication demand. As such, even though each end device conforms to the channel access time requirement (e.g., 0.1% or 1% duty cycle in Europe [18]), the ALOHA-based LoRa networks will have degraded network performance due to massive collisions when the numbers of end devices grow sharply in this era of IoT [6, 7, 10, 20, 28, 29, 31, 39, 45, 46].

The absence of carrier sense capability on off-the-shelf LoRa end devices impeded studies and implementations of the more advanced carrier-sense multiple access (CSMA) schemes. However, a recently introduced feature called channel activity detection (CAD) sheds light on CSMA for LoRa networks. CAD, which is available on all of the latest LoRa radios (e.g., SX126x and SX127x), is designed for energy-efficient preamble detection. Although it is not for full-fledged carrier sense, our extensive measurement study shows that CAD can also detect the payload chirps with satisfactory performance. Specifically, it achieves more than 95% accuracy in detecting the occupancy of a logic channel due to an ongoing frame transmission.

Most existing CSMA designs from other wireless technologies are based on energy detection on the spectrum. For instance, they adopt received signal strength (RSS) as an indicator of the presence of an ongoing transmission [1, 2]. Simply porting them to LoRa networks may yield inefficiency. To achieve efficient CSMA for LoRa networks, the unique features and constraints of LoRa radios need to be considered. For instance, LoRa frame can traverse below the noise floor; concurrent transmissions with different SFs are encouraged in the same frequency channel. These features render the traditional energy detection based CSMA designs ineffective. Thus, a clean-slate CSMA redesign based on LoRa’s PHY layer characteristics is needed. To this end, we follow the progressive systems development methodology to design an efficient CSMA-based LoRa MAC (LMAC). The advancing versions of LMAC integrate various new designs and features based on the precedent code base to address additional and realistic networking problems. Their details are as follows.

**LMAC-1.** LMAC-1 implements the basic functionality of CSMA (i.e., to avoid collisions by refraining from transmission when the channel is busy). We implement the distributed inter-frame space (DIFS) mechanism that performs a fixed number of CADs and a random back-off (BO) mechanism. LMAC-1 shall outperform ALOHA in terms of packet reception ratio (PRR) and network goodput when the communication demand increases.

**LMAC-2.** LMAC-1 falls short of balancing the communication loads. As a result, the busier logic channels have higher frame loss rates. LMAC-2 aims to balance the loads among the logic channels by allowing LoRa end nodes to select the less crowded logic channels based on their local information. We devise an indirect channel probing approach for each end node to update its knowledge regarding the channels’ crowdedness based on the CAD results during DIFS and BO processes. Thus, LMAC-2 remains as a fully distributed MAC that aligns well with the LoRaWAN Class A specification. In addition, it does not introduce extra CAD overhead compared with LMAC-1. LMAC-2’s channel selection shall equalize the channel utilization and bring forth a performance improvement.

**LMAC-3.** In LMAC-2, the LoRa end nodes may develop biased local views regarding the channel loads. In LMAC-3, the gateway broadcasts global views of channel loads using periodic beacons.
The beaconing renders LMAC-3 consistent with the LoRaWAN Class B specification. Upon receiving the global view, the LoRa end nodes update their local views to guide the channel selection before transmission. With the gateway’s assistance, LMAC-3 shall further improve network performance with respect to LMAC-2.

To support system design and evaluation, we construct a testbed consisting of 50 LoRa end nodes and a LoRa gateway in a lab environment. We conduct extensive, comparative evaluation experiments with goodput, PRR, and radio energy consumption per successfully delivered frame as main performance metrics. The results show that, compared with ALOHA which yields sub-50% PRR, all LMAC versions maintain sub-90% PRR when the communication demand increases. With respect to ALOHA, LMAC-1, LMAC-2, and LMAC-3 respectively bring 1.5×, 1.9×, and 2.2× goodput improvements, and 2.08×, 2.37×, and 2.38× reductions in radio energy per successfully delivered frame. In addition, LMAC-3 achieves the most balanced channel loads. The experiments on a deployment of 16 LoRa end nodes in a university also shows significant performance improvements brought by LMAC in comparison to ALOHA.

The contributions of this article are summarized as follows:

- We experimentally quantify key limitations of RSS-based carrier sense for LoRa networks.
- We study LoRa CAD and show its full-fledged channel-selective carrier sense capability.
- We design advancing versions of LMAC to address LoRa’s unique features and various realistic networking problems. The fully distributed LMAC-2 and the gateway-assisted LMAC-3 are the final yields of this work that align well with the LoRaWAN Class A and Class B specifications, respectively.
- Our testbed experiments show that, compared with ALOHA, LMAC brings doubled goodput and halved radio energy consumption under saturated communication demands.
- We discuss how LMAC enhances low signal-to-noise ratio (SNR) capabilities of LoRa, resulting in improved quality of service (QoS) for network service providers.
- We discuss the extent of LMAC’s interoperability with ALOHA considering that LMAC would be gradually adopted across networks.
- We also discuss LMAC from a regulatory perspective. We identify possible challenges that may rise in some regions and propose solutions.
- In addition, we make available an LMAC-integrated LoRaWAN codebase (see [27]).

The rest of this article is organized as follows. Section 2 presents a LoRa primer. Section 3 investigates CAD. Section 4 describes the testbed. Section 5 designs LMAC. Sections 6 and 7 present testbed experiments and simulation results, respectively. Section 8 reviews related work. Section 9 concludes the article.

2 LORA PRIMER

In this section, we present a background on LoRa PHY layer, LoRaWAN specification, and CAD.

LoRa and LoRaWAN. A typical LoRa network consists of a number of geographically distributed end devices that transmit/receive data to/from one or more gateways. The radio frequency spectrum allotted to LoRa is also shared with a number of other PHY-layer protocols. The LoRa PHY layer divides the spectrum into multiple frequency channels (CHs), where each CH supports both uplink and downlink communications. LoRa employs CSS modulation, in which each chirp linearly sweeps a mandatory predefined bandwidth BW from the minimum frequency f_{min} to the maximum frequency f_{max} of the CH. CSS represents data by the initial frequency of a chirp.

A LoRa frame has three parts: a preamble of 8 up chirps, a start frame delimiter (SFD) of 2 \frac{1}{4} down chirps, and a payload of multiple data chirps. Each data chirp lasts exactly a duration of T_{sym} milliseconds. A data chirp starting from an arbitrary frequency f_{x} represents a radio frequency}
symbol out of $2^{SF1}$ total symbols where $SF$ is the spreading factor. The chirp starting from $f_x$ linearly increases in frequency over time, reaches $f_{max}$, wraps around, and resumes its journey from $f_{min}$ back to $f_x$. As such, each data chirp sweeps the entire bandwidth $BW$. Demodulation is performed at the gateway by relating the FFT bins of the multiplication of the payload with a series of down chirps.

The LoRa PHY layer allows concurrent transmissions in the same CH by using the distinct slopes of the concurrent chirps (i.e., SFs). This feature enhances a LoRa end device’s ability to co-exist among many end devices that use the same CH. LoRa supports six SFs, from SF7 to SF12. In the rest of this article, we use the notation CH/SF to denote a logic channel in which the simultaneous transmissions collide.

LoRaWAN is a data link layer specification based on LoRa. It defines three classes, namely Class A(all end devices), Class B(eacon), and Class C(ontinous listening). The three classes have distinct characteristics in response times and energy expenditures; however, they are all based on the primitive ALOHA MAC. Although the requirement on the channel access time (e.g., 0.1% and 1% in Europe) mitigates channel contention, the ALOHA-based LoRaWAN cannot sustain the growth of end devices. Besides MAC, LoRaWAN also includes other link-layer functions such as adaptive data rate and payload encryption.

Channel activity detection. LoRa recently included CAD as a power-optimized mechanism to detect preamble chirps preceding full-fledged payload demodulation. This allows a LoRa node to energy-efficiently detect an incoming frame without resolving to power-hungry continuous RX mode. Some details of CAD are as follows. Prior to responding to a CAD request, the LoRa radio should be set on the desired CH/SF. A single CAD lasts $T_{sym} + \frac{32}{BW}$ ms, during which the radio takes the following steps. First, a lock in the phase-locked loop is achieved to tune the radio to the desired CH. The radio then switches to the receiving mode for $T_{sym}$ ms. Second, the radio switches to processing mode to search for strong cross-correlation between the received I/Q samples and a locally generated up chirp of the same SF. The cross-correlation computation takes $\frac{32}{BW}$ ms. Last, a CadDone interrupt is generated. If the cross-correlation is high, a CadDetected interrupt is also generated.

3 ENABLING LORA CARRIER SENSE

Carrier sense capability is desirable for advancing LoRa. Two factors, however, challenge effective LoRa carrier sense. First, LoRaWAN encourages concurrent transmissions in a single CH using different SFs. Second, most CSS-based modulation methods including LoRa allow signals to traverse below the noise floor. For instance, a LoRa receiver’s sensitivity is $10^5$ times higher than that of Wi-Fi [14, 41]. The preceding two factors render the traditional RSS-based channel occupancy detection futile. In this section, we first highlight the limitations of traditional RSS-based carrier sensing for LoRa. We then extend the understanding of LoRa CAD beyond the technical documentation [4, 35] and evaluate the effectiveness of using CAD for the purpose of carrier sensing.

To use CAD for carrier sense, we need to consider three factors: (1) How reliable is CAD? (2) Could CAD be performed on targeted SFs while the CH is also used by transmissions in others SFs? (3) How frequently can consecutive CADs be performed? To answer these questions, we conduct an experiment with three LoRa nodes and a software-defined radio (SDR) setup. In the experiment, a node transmits an SF12 frame on CH1 (868.1 MHz, 125 kHz BW). Two other nodes tune to CH1 and continuously perform CAD in SF11 and SF12, respectively. Figure 1 shows a portion of the CH1 spectrogram recorded by the SDR and the two nodes’ CAD results. We have the following two observations. First, CAD can detect data symbols. Second, CAD strictly adheres to the specified SF and does not generate false positives for other transmissions of different SFs in the same channel. To develop a better understanding on the behavior of CAD, we conduct more
experiments on the cross-correlation process of CAD. Cross-correlation computes the extent of similarity between a received chirp and the locally generated as a function of time lags applied to the locally generated chirp. Specifically, we compute the cross-correlation using two SF9 and SF12 frames recorded by the SDR. First, we test the case in which the incoming SF12 preamble chirp aligns ideally with the locally generated SF12 up chirp. The cross-correlation result shown in Figure 2 indicates a clear spike signifying a detection. We then introduce a time offset of $\frac{T_{sym}}{2}$ ms, as shown in Figure 2 on the incoming SF12 preamble chirps. As a result, the locally generated up chirp covers portions of two consecutive incoming chirps. Their cross-correlation produces two clear spikes corresponding to the portions of chirps. Next, we compute cross-correlation for an ideally aligned incoming data chirp which again results in two spikes corresponding to each segment of the data symbol as shown in Figure 2. Afterward, we compute cross-correlation for a data chirp with a time offset of $\frac{T_{sym}}{2}$ ms and observe corresponding spikes as shown in Figure 2. Note that the preceding simulated chirp offsets can be naturally incurred when the CAD process starts at an arbitrary point of the incoming frame. This can be observed from Figure 1—that is, a successful CAD does not necessarily need alignment with the incoming symbol. We then investigate how cross-correlation behaves for chirps in a different SF. In Figure 2, CAD is performed between incoming SF9 chirps and locally generated SF12 chirps. It yields no peaks. Last, to predict the presence of a set SF, resulting peaks could be added and compared against a threshold.

Efficiency of CAD indoors and outdoors. LPWANs are typically deployed over large geographical areas. We also investigate the efficacy of CAD in an outdoor area. We deploy two LoRa gateways, one indoor and the other outdoor, and use a mobile LoRa node to perform CAD at multiple indoor and outdoor locations. At each location, we record the SNR reported by the LoRa node when successfully receiving a frame and the CAD efficiency that is the ratio of successful CADs among all CADs performed. Figure 3 illustrates the experiment scenarios of eight indoor locations across 800 m² of office spaces and 23 outdoor locations scattered across a university area of 3.5 km². Figure 4 shows the SNR and CAD efficiency at all visited locations. The median CAD efficiency values are 98% and 94% indoors and outdoors, respectively, with SNR ranging from −15 dB to +15 dB.

CAD energy footprint. Since CSMA introduces an additional energy overhead for each frame transmission, we profile the power consumption of CAD. To capture the power trace, we connect a power monitor [37] in series to the power supply pins of the SX1276 radio chip. Figure 5 depicts the incurred current of a single CAD in comparison with those during the idle and transmission states. CAD draws 0.03 W, which is 4.4× of the idle power and 11× less than the transmission power. A CAD in SF7 lasts for 1.2 ms only, consuming 0.029 mJ only. Under SF12, it lasts for 31.6 ms and consumes 1.1 mJ.

RSS-based carrier sensing. Received signal strength based Listen Before Talk (RSS-LBT) is a common strategy used by wireless technologies to minimize collisions. RSS-LBT avoids collisions by
sensing the instantaneous signal strength ($\text{RSS}_{\text{det}}$) of a set CH and a BW combination where transmissions are withheld when the detected instantaneous signal strength ($\text{RSS}_{\text{det}}$) is higher than the threshold set by the respective authority. Although the adoption of RSS-LBT within the ISM spectrum is optional in most regions (i.e., North America [5] and the EU [18]), it is mandatory in others.
Two example thresholds are −80dBm for Japan and −65dBm for South Korea. We conduct an experiment to estimate the sensing range of RSS-LBT and thereby estimate the protection range it can offer. The experiment recorded the RSS variation across a varying distance between two LoRa nodes. One node transmitted at a transmission power of 14dBm while the other, configured to the same CH and BW, consistently recorded the RSS ($RSS_{\text{det}}$) for each incremental distance. Two versions of the experiment were conducted, one indoors and the other outdoors. The outdoor experiment was conducted under LoS conditions. Figure 6 presents the RSS variation between the two nodes while also comparing them with the two RSS thresholds set by South Korea and Japan. It can be seen that the RSS threshold of South Korea (−65dBm) translates to a carrier sensing range of approximately 30 m, whereas that of Japan (−80dBm) translates 100 m under outdoor conditions. Figure 6 also suggests that this sensing range is even smaller for indoor conditions. This range may satisfy carrier sense requirements of short-range communication protocols such as Zigbee and WiFi. RSS-based carrier sense is well integrated to state-of-the-art communication protocols such as Zigbee and WiFi and performs well under those modulation schemes [2, 21]. However, the success of RSS-LBT as a carrier sense mechanism is reliant on the close communication range [13]. In contrast, a single LoRa gateway can cater to nodes operating under significantly larger areas. Therefore, using RSS as a carrier sense mechanism for LoRa could only provide a limited protection range among LoRa nodes. By comparing the ranges of RSS-based carrier sense (Figure 6) and CAD (Figure 4), it can be concluded that CAD-based detection enables order of magnitude longer detection distance.

Another disadvantage of RSS-LBT is its inability to differentiate concurrently transmitted orthogonal LoRa frames. As a result, a node may unnecessarily withhold transmissions until all on-air transmissions stop disregarding SF orthogonality. This leads to a significant loss of...
transmission opportunities and unnecessary wait times whenever nodes with multiple SFs inter-mingle. Consequently, these losses turn into inefficient use of time, spectrum, and energy. This limitation can be easily alleviated by replacing primitive RSS-LBT with a CAD-based carrier sensing mechanism which naturally offers SF selectivity.

Summary. The preceding experiments give the following key observations. First, CAD can be used as an effective and reliable carrier sense method. Second, CAD is SF-selective—that is, it can detect transmissions in a specified SF while remaining insensitive to other transmissions in different SFs. Third, CAD is an energy-efficient operation compared with the energy consumption of data transmissions. Fourth, CAD works indoors and outdoors with high detection efficiency. Fifth, CAD can cater to unique PHY-layer properties of LoRa modulation to enable the missing carrier sense capability in LoRaWAN networks.

4 LORA TESTBED DESIGN

To facilitate the design and evaluation of LMAC, we build a scalable LoRa testbed. This section presents the testbed design and experiment workflow that has been used for both indoor and outdoor experiments.

Testbed design. We aim to meet the following design requirements. First, the testbed should support a sparse deployment (i.e., across a university). Second, it should support convenient experiment configuration via a central controller. Third, it should provide a bidirectional backhaul link to disseminate configurations and receive experiment logs from the distributed LoRa nodes. Meeting these requirements imposes implementation challenges. For instance, it is non-trivial to synchronously control many LoRa nodes.

Connecting each LoRa node to the central controller is not feasible, as the LoRa nodes do not have wideband radios to support backhaul links. As such, we employ extensible units, each a cluster of LoRa nodes, and call them distributed controllers. The central controller and the distributed controllers are Raspberry Pi single-board computers that perform two coordinated roles. The distributed controllers execute and, if necessary, relay control commands to LoRa nodes, thereby creating wideband out-of-band control links between the central controller and LoRa nodes.

Each testbed LoRa node is assembled in house by integrating an SX1276 LoRa transceiver with an Arduino Atmega328p MCU via a bridge board having a capacitor to support transient power requirements unsupported by USB power delivery [30]. Communications between LoRa nodes and a distributed controller (e.g., protocol firmware downloading and experiment log uploading) happens via an emulated serial interface. Figure 7(a) illustrates the connectivity between each component of the testbed, Figure 7(b) shows a shot of testbed hardware, and Figure 7(c) shows the outdoor gateway. To address the requirement of supporting geographically distributed nodes, testbed devices communicate through a network tunnel [15] from their local Ethernets.

Experiment workflow. An experiment includes the following steps: establishing connectivity among all components, loading a MAC protocol into LoRa nodes, parsing experiment configurations, executing the experiments, and finally collating logs. In what follows, we provide some important details. At the beginning of an experiment, the central controller distributes a precompiled binary MCU program implementing a MAC protocol and two parameters (demand and experiment duration) to the distributed controller. The demand is the requested number of LoRa frames to be transmitted within the specified experiment duration. The extent that the requested demand is met reflects the network performance of the tested protocol. Varying the demand creates different levels of contention. After a mutual confirmation process to ensure that devices are ready, the central controller issues a start command to launch an experiment. Once the experiment duration expires, the distributed controllers report the following experiment logs back to the central controller: (1) LoRa node ID, (2) demand, (3) transmitted packets, (4) CADs performed on each CH/SF, and
Fig. 7. (a) Testbed architecture. (b) Indoor testbed devices. (c) Outdoor gateway.

Fig. 8. PRR versus demand under ALOHA. The error bar shows the max, min, and mean PRR among 50 nodes.

(5) the count of repeated experiments. In our experiments, all LoRa frames contain 16-byte payloads with the following fields to facilitate performance counting: (1) node ID, (2) packet counter, (3) demand, and (4) the number of padding bits.

5 DESIGN OF LMAC

This section presents the progressive design of LMAC to exploit CAD for CSMA-based MAC in LoRa networks. Under ALOHA, the end nodes access the medium abruptly regardless of the channel status, which results in excessive collisions when the communication demand is high. Figure 8 shows an experiment result in which the average PRR of ALOHA drops from 0.5 to 0.2, when the demand increases. The first version of LMAC—LMAC-1—uses CAD-based carrier sense to avoid the collisions.

5.1 LMAC-1

LMAC-1 is largely based on the distributed coordination function (DCF) that serves as one of the channel access methods of IEEE 802.11. However, as LoRa differs significantly from the 802.11 PHY layer, changes are required to adapt DCF to LoRa. A key difference between 802.11 and LoRa is that LoRa does not have a feedback channel and thus cannot support the acknowledgment mechanism of DCF.

In LMAC-1, the CH/SF is chosen randomly for each transmission. An LMAC-1 node transmits a frame after the selected CH/SF is confirmed to be idle. To this end, LMAC-1 adopts the DIFS
Fig. 9. Following LMAC-1, node B detects a collision when node A is transmitting in the same CH/SF.

mechanism from DCF with a fixed number of CADs per DIFS slot. When a DIFS slot completes with all CADs reporting idle channel, the MCU generates a random BO value ($N_{BO}$) and enters the BO phase immediately. Then, the node decrements $N_{BO}$ per each CAD reporting idle channel. The randomized initial value for $N_{BO}$ reduces the possibility where two or more frames collide should the DIFS processes happen to start at the same time. Upon $N_{BO}$ reaching zero, the node transmits the frame. Since a node continues checking the availability of the channel during the BO phase, an occupied channel will reset the node to the DIFS state. Should that occur, the node needs to wait for an idle channel again and resumes counting down $N_{BO}$. Note that the initial $N_{BO}$ value is not regenerated until the frame is transmitted. In the rest of this article, we refer to CADs that report a busy channel as failed CADs and vice versa. A failed CAD operation within DIFS or BO will render a failed DIFS or BO slot, respectively. Figure 9 illustrates the process in which two nodes contend for the same CH/SF under LMAC-1. Node B detects an ongoing transmission during DIFS and continues performing DIFS until the channel is idle. Upon detecting an idle channel, node B enters the BO state with $N_{BO} = 3$ and starts decrementing $N_{BO}$ with each successful CAD. Once $N_{BO}$ reaches zero, node B transmits the frame. In the implementation of LMAC-1, we set the range for the initial value of $N_{BO}$ to be from 4 to 64 and 12 CAD operations per DIFS slot.

LMAC-1 focuses on avoiding collisions. From the evaluation results that will be presented in Section 6, LMAC-1 improves the PRR significantly over ALOHA. In what follows, we present an experiment result that shows the imperfection of LMAC-1 and motivates LMAC-2. Figure 10 presents the numbers of frames transmitted by all of the end devices (labeled TX) and received by the gateway (labeled RX) in each CH/SF during an experiment. The difference between the TX and RX bars in a certain CH represents the number of lost frames. From the figure, an important observation is that the busy channels (i.e., those with high bars) have more lost frames. The cause of the frame losses is explained as follows. The gateway radio chip SX1301 incorporates eight decoders that can be configured to demodulate incoming frames concurrently. Each decoder must be configured to a single CH but listens to all SFs. However, all decoders share an I/Q buffer. Should the decoders lag behind in decoding incoming I/Q data, a buffer overflow will occur, resulting in frame loss. This is because a decoder is tied to a specific frequency CH and a large number of frames arriving in that CH result in that a decoder lags behind compared to others. We refer interested reader to the work of Liando et al. [26] for a thorough analysis of concurrent frame reception. The design of the SX1301 chip and the results shown in Figure 10 imply that if the loads of the logic channels are equalized, the communication performance can be further improved.

To this end, the nodes can choose underutilized or idle channels instead of contending for a highly utilized channel. LMAC-2 aims to implement this feature.

5.2 LMAC-2

To select an underutilized channel, a node needs to track the utilization levels of all channels. Since continuous channel probing is time consuming, LMAC-2 leverages information from failed...
Fig. 10. Numbers of frames transmitted by all of the nodes (TX) and received by the gateway (RX) under LMAC-1. More utilized channels with higher bars suffer from more frame losses (i.e., larger differences between the TX and RX bars).

DIFS and BO operations to enable indirect channel probing. Thus, each LMAC-2 node maintains historical information regarding the crowdedness of past sensed CH/SF combinations. Later, the LMAC-2 node consults this information to shift to a better CH/SF when a channel is sensed busy during DIFS or BO. Doing so leads to more equalized channel loads and thus, better network performance. In addition, from the perspective of a node, the average number of CADs per transmission and the related time/energy expenditure can be reduced since busy channels are avoided. The key differences between LMAC-1 and LMAC-2 are as follows. In LMAC-1, a node fixates to a random CH/SF until a frame is transmitted. In contrast, an LMAC-2 node uses historical information to dynamically select the best CH/SF. The details of LMAC-2 are presented as follows.

Each LMAC-2 node maintains a channel occupancy matrix ($\Gamma$) of size $N_{CH} \times N_{SF}$, where $N_{CH}$ and $N_{SF}$ respectively represent the numbers of CHs and SFs. Each element in $\Gamma$, denoted by $\gamma_{CH, SF}$, represents the historical utilization rate for a given CH/SF. The $\gamma_{CH, SF}$ is updated for the $t + 1^{th}$ time whenever a node has a failed CAD during DIFS or BO. We adopt exponential averaging for the updating: $\gamma(t + 1) = \alpha \left( \frac{CAD_{busy}}{CAD_{total}} \right) + (1 - \alpha) \gamma(t)$, where $\alpha$ is the learning rate, and $CAD_{busy}$ and $CAD_{total}$ are the number of CADs performed when the channel is busy and the total number of CADs performed, respectively. The preceding exponential averaging is used for adapting to channel dynamics whilst maintaining historical information. As LoRa channels are highly dynamic, in our experiments, we set $\alpha = 0.8$ to ensure that higher importance is placed on the latest sensed channels.

We now describe how an LMAC-2 node uses $\Gamma$ to select the next best CH/SF. A real-world LoRa deployment is subject to dynamic channel conditions. A channel that is seemingly busy may later become free. Therefore, simply choosing the least busy CH/SF according to the $\gamma$ value may not always reap optimal network performance. As such, we randomize the choice. Specifically, we choose the best, the second best, and the third best CH/SF in terms of $\gamma$ value with probabilities of 0.5, 0.3, and 0.2, respectively. This allows a forgiving approach for channels that were once crowded.

The next SF selected is set higher or equal to that configured during deployment to ensure the frame robustness.

We use an example shown in Figure 11 to illustrate the interaction between two LMAC-2 nodes that have selected the same CH/SF. Since node A completes BO earlier than node B, it proceeds with the transmission and node B reports a failed BO during node A’s transmission. Thus, node B updates the occupancy matrix element corresponding to the current CH/SF, and reconfigures the radio to the next CH/SF. Since the newly selected CH/SF is different from that used by node A, node B can complete DIFS and BO successfully and transmit the frame. If following LMAC-1, node B will wait until node A completes transmission.

As indirect channel probing is opportunistic, the channel occupancy matrices of LMAC-2 nodes are different. We illustrate this in Figure 12. Specifically, Figures 12(a), (b), and (c) visualize the $\Gamma$
In LMAC-3, a gateway transmits a periodic beacon that contains the global channel occupancy matrix ($\Psi$) to the nodes in the network. Promptly, each receiving node follows a merging process to supplement its local occupancy matrix $\Gamma$ with the gateway’s global occupancy matrix $\Psi$. The $\Psi$ is an $N_{CH} \times N_{SF}$ matrix where each element $\Psi_{CH,SF}$ represents the contention level in a given CH/SF.

However, several challenges need to be addressed for large-scale LoRa deployments to benefit from the global views of gateways. First, the process of disseminating a gateway’s global view across the network should not be detrimental to the gateway’s reception capability. Second, a LoRa node may receive overlapping coverage from two or more gateways. Due to network dynamics, a given CH/SF will face dissimilar contention across gateways.
Most commercial LoRa gateways only equip a single antenna and do not support duplex communications. As such, responding to each uplink frame with the $\Psi$ matrix will reduce reception performance. Thus, to disseminate a gateway’s global view, LMAC-3 employs time-distributed beacons via a predefined feedback channel similar to LoRaWAN Class B. Figure 13 shows an example of this process where a gateway disseminates its $\Psi$ to two nodes. Under LMAC-3, a LoRa gateway mostly stays in receiving mode and only switches to transmitting mode when it broadcasts beacons containing $\Psi$. Note that the nodes may miss a single beacon due to clock drifts. Thus, the beacon is transmitted twice on the feedback channel to increase the probability of reception.

LMAC-3 does not require tight clock synchronization across nodes but coarse synchronization. Upon receiving $\Psi$, the nodes will perform the matrix merging that will be explained shortly and continue to transmit pending frames while awaiting the next beacon period. The $\Psi$ matrix and the gateway ID is carried by a frame payload of 49 bytes where an element $\Psi_{CH, SF}$ is represented by a 8-bit non-negative integer with 254 representing the highest load.

The $\Psi_{CH, SF}$ is formally defined as the ratio of the chirps that utilized the channel ($S_{CH, SF}^RX$) over the maximum possible during a beaconing period ($T_{beacon}$). The maximum number of chirps during a $T_{beacon}$ is the multiplication of the chirp rate ($S_{CH, SF}$) and $T_{beacon}$. The chirp rate is defined as the reciprocal of the symbol duration in the considered SF. Following the preceding,

$$\Psi_{CH, SF} = \frac{S_{CH, SF}^RX}{S_{CH, SF} \times T_{beacon}}.$$  

To compute $S_{CH, SF}^RX$ for a given CH/SF, we count the number of chirps in the totally $rx$ received frames in the CH/SF as follows: $S_{CH, SF}^RX = \sum_{i=0}^{rx}(n_{sfd} + n_{pre}^i + n_{pl}^i)$, where $n_{sfd}$ is the number of the SFD chirps in a frame, and $n_{pre}^i$ and $n_{pl}^i$ are the numbers of preamble chirps and payload data chirps in the $i^{th}$ received frame [3].

We now discuss how to merge $\Gamma$ and $\Psi$. The simplistic approach of using $\Psi$ solely to guide the channel selection is problematic, because the nodes will likely choose the same channel causing channel occupancy imbalance and contention. Our merging process adopts an element-wise weighted sum of $\Gamma$ and $\Psi$, with a large weight for $\Gamma$ and a small weight for $\Psi$. Thus, the merging result preserves the diversity of the nodes’ observations while also capturing the global observations by the gateway.

To address the multiple gateway scenario, a gateway beacon integrates a single byte ID that aids in identifying different gateways. The $\Psi_{CH, SF}$ that has normalized channel contentions across gateways allows for simple merging of multiple $\Psi$ matrices at node level. The preceding enables LMAC-3 to be scalable.

Based on our empirical results under different environment conditions for varying contentions, we choose 0.8 and 0.4 respectively as weights for $\Gamma$ and $\Psi$. The experiment results presented in Section 6 suggest that the determined weights provide high performance across various conditions and therefore can be used as is by new networks.
6 EVALUATION

We conduct indoor and outdoor experiments using the testbed described in Section 4 to evaluate LMAC. We use PRR, goodput, and average energy consumption per successful reception as the performance metrics. We also consider protocol fairness in terms of the number of CADs performed by each node and distribution of channel loads.

6.1 Indoor Experiments

We conduct indoor experiments with co-located nodes in a lab environment, which allows us to better control the process and cover an extensive range of parameter settings. The experiments use 50 LoRa nodes. There are total of 16 CH/SF combinations used in the experiments (8 CHs and 2 SFs). In the experiments, we adopt the following default settings: 16-byte payload, 10-chirp preamble, and 2-dB transmission power. Each experiment lasts 1 minute and is conducted in different times of the days over 2 months.

**Overall network performance.** We evaluate the goodput, PRR, and node energy consumption when network demand varies under the four MAC protocols. Network goodput is the amount of meaningful data delivered per second. From Figure 14(a), we can see that compared with ALOHA, LMAC-1/2/3 improve the goodput by 1.52×, 1.87×, and 2.21×, respectively. Thus, LMACs can deliver data efficiently owing to the collision avoidance. Moreover, the performance increments of LMACs suggest the effectiveness of utilizing channel diversity and CH/SF selection. ALOHA goodput rises as the network demand increases but reduces once network demand is beyond 2,500 B/s. This is consistent with the existing understanding of ALOHA performance. The goodputs of LMACs saturate and remain flat once the demand is beyond 2,000 B/s. The saturation of LMACs goodput suggests the contention in channel utilization.

From Figure 14(b), the PRRs of LMACs are maintained above 90%, whereas that of ALOHA decreases with demand. The decline is due to increasing collisions. Differently, LMACs maintain high PRRs through avoiding collisions and busy CH/SF.

Figure 14(c) shows the average energy consumed by a node per successfully delivered frame under the four MAC protocols. Compared with ALOHA, LMAC-1/2/3 reduce the energy by 2.08×, 2.37×, and 2.38×, respectively. Figure 14(c) also shows that LMAC-2 and LMAC-3 consume similarly low energy, whereas LMAC-1 consumes slightly higher energy than LMAC-2/3. This is because LMAC-1 nodes continuously perform DIFS to contend for the channel. Differently, LMAC-2/3 perform less CADs due to wise CH/SF selection. In the experiment logs, we find that LMAC-3 performs 1.92% less CADs than LMAC-2, whereas LMAC-2 performs 7.81% less CADs than LMAC-1. The goodput gain of LMAC-3 over LMAC-2 is due to the time saved from less CADs performed. Specifically, LMAC-3 nodes can utilize the saved time for frame transmissions, whereas the PRRs of LMAC-2/3 are similar.
Fig. 15. CDFs of the per-node network performance under ALOHA, and LMAC-1/2/3 when the network demand is 2,600 B/s. Most LMAC nodes outperform ALOHA nodes except for a small set of ALOHA nodes. The steeper CDFs of LMACs suggest that the performance metrics are fairer among the nodes.

Overall, LMAC outperforms ALOHA through collision avoidance and achieves up to $2.21 \times$ goodput improvement and $2.38 \times$ reduction in nodes energy consumption, while maintaining above 90% PRR.

**Per-node performance.** A closer look at the results from a single experiment will provide more insights into the advantages of LMAC over ALOHA. Figure 15 shows the cumulative distribution functions (CDFs) of the per-node network performance metrics. From Figure 15(a), the majority of ALOHA nodes suffer low goodput that are smaller than 5 B/s, whereas the remaining ALOHA nodes achieve goodputs scattered from 28 to 74 B/s. The standard deviation of ALOHA’s per-node goodput is 26.79 B/s. In contrast, under LMACs, the CDFs do not exhibit undesirable long tails; the per-node goodputs are mostly within the range of 20 to 60 B/s, with a standard deviation of 14.89 B/s. The preceding results suggest that, compared with ALOHA, LMACs achieve better balance among the nodes in utilizing the communication medium.

Figure 15(b) shows the CDFs of the per-node PRRs. LMAC-2/3 achieve at least 80% PRR. The mean PRR values of all LMACs are above 90%. In contrast, about 40% of ALOHA nodes suffer zero PRR, as the frames are lost due to collisions. The remaining nodes have scattered PRRs from 0% to 95%. Experiment logs suggest that ALOHA nodes with high PRRs dominate the CH/SF and hinder the reception of other nodes.

Figure 15(c) shows the CDFs of the per-node energy consumption. Up to 30% of ALOHA nodes cannot deliver a single packet. Thus, they are not accounted for in the CDF. Another 40% of ALOHA nodes consume excessive per-frame energy due to collisions. LMAC nodes are rarely affected by frame losses. The total energy consumed for channel probing trades for the exemption of futile energy consumption for transmitting lost frames. Overall, all LMAC nodes outperform 85% of ALOHA nodes in terms of energy consumption.

In summary, a closer look into individual node performance reveals significant reduction in energy consumption per node, which is hidden in the earlier results due to averaging. Moreover, under LMACs, the goodput and PRRs of each node are more evenly distributed among nodes, whereas ALOHA nodes are clearly divided into the two extremes with a small number of exceptional nodes.

**Balance and fairness.** We evaluate the loads of all frequency channels under various MAC protocols. Since each channel supports multiple SFs, the frequency channel load ($\Psi_{CH}$) needs to be aggregated properly from the logic channels loads ($\Psi_{CH,SF}$) that are estimated by the gateway in the experiments. By averaging from SF7 to SF12 in each channel, the channel load for each frequency channel is obtained.

Figure 16 shows the scatter plot of PRR versus channel load. The area of PRR $\in [0.8, 1]$, $\Psi_{CH} \in [0, 0.1]$ is zoomed in to present a detailed view of LMAC performances. PRR of ALOHA drops quickly with the frequency channel load. The result of ALOHA shown in Figure 16 reinforces our
understanding on the disadvantages of ALOHA’s completely uncoordinated transmissions. In contrast, the frequency channel loads of LMAC-1 are maintained within 0.1. This low load helps maintain satisfactory PRRs. This shows that LMAC-1 effectively avoids frequency channel saturation. More importantly, the channel selection mechanisms make LMAC-2 and LMAC-3 operating within the region with frequency channel load smaller than 0.1 and PRR within [80%, 100%].

Figure 17 shows the sorted average numbers of CADs per frame performed by LMAC-2 and LMAC-3. LMAC-2 performs more CADs and yields a steeper slope, which means higher variation. The additional channel information provided by the gateway helps LMAC-3 nodes avoid known congested channels and perform less CADs. Many CADs are performed for every frame due to the DIFS design. Recall that each DIFS slot contains 12 CADs. Moreover, CADs are also performed during the decrement process of $N_{BO}$.

With the preceding experiments, LMACs yield a controlled and a fairer frequency channel utilization. This allows node performances to be more evenly distributed and ensure data delivery for each node. Moreover, the overhead of LMAC-2/3 is shown to be spread evenly among nodes where most nodes perform less than 250 CADs or 21 DIFS per transmission within a high contention network. The fewer CADs with high reception probability save radio energy for each successful frame delivery.

The latency introduced by LMAC is largely bound to the network’s contention levels. LMACs, under high contention levels, will incur higher latency due to the wait for free channels. However, ALOHA is not subject to this latency but suffers from much higher frame loss rate due to collisions, which translates to much higher delay if retransmissions are used to ensure the same level of PRR.

### 6.2 Outdoor Experiments

To investigate the feasibility and performance of LMAC in an outdoor environment, a set of experiments were conducted in our university area. The experiment makes use of four distributed controllers and 16 end nodes. To exhibit the advantages of LMAC in a high-density network with a low number of end nodes, the outdoor experiments use six CH/SF combinations. This set of experiments focus on LMAC-2 and LMAC-3 as the final yields of this work. We collect 30 minutes of experiment traces with 10 minutes for all protocols including ALOHA. Different from the indoor experiments, the outdoor experiments were conducted between 10AM to 10PM spanning across 7 days.

The results of the experiments are presented in Figure 18. Each point in the figure represents a node with the averaged SNR when the gateway received the frames. The standard deviation of SNR is between 0 and 3 dB. The PRR (Figure 18(a)) and goodput (Figure 18(b)) of LMAC-2/3
spread across vertically and begin to concentrate downward as SNR reduces. This shows that the nodes with higher SNRs (i.e., well placed or nearby nodes) can deliver their frames more reliably. However, nodes with lower SNRs (i.e., poorly placed or obstructed nodes) most likely cannot reach the gateway and fail to deliver their frames. In contrast, ALOHA nodes are scattered in the plots with no clear pattern. Although there are nodes that achieve very high PRRs and goodputs, the majority of nodes have low PRRs and goodputs. The average PRR and goodput of ALOHA are 0.22 and 15.98 B/s, respectively, whereas LMAC-2/3 reports 0.49 and 0.63 average PRRs with 25.17 and 34.83 B/s average goodputs, respectively. Notice that ALOHA has less points than LMAC-2/3. This is because more than 40% of ALOHA nodes’ frames are lost at the gateway, resulting in no SNR values registered for those nodes.

To understand the actual traffic of the network for each protocol, Figure 19 presents each frame contribution to the effective traffic of the respective protocol. The results show that high SNR frames are more likely to be received by the gateway. Notice that the effective traffic of ALOHA only starts to increase when SNR reaches 4 dB, whereas LMAC-2 and LMAC-3 effective traffic starts to increase at an earlier SNR. The late increase of ALOHA shows that frames with lower SNR are often drowned by higher SNR frames. Differently, LMAC-2/3 allow low-SNR nodes to transmit only when the channel is assessed to have low collision probability. Moreover, the total effective traffic of ALOHA is lower than those of LMAC-2/3. This coincides with the earlier observation where more than 40% nodes are not accounted for due to no frame reception.

The preceding results suggest that LMAC can be affected by environment dynamics but provides effective traffic control even for low-SNR nodes. Contrarily, ALOHA nodes are divided into two extremes where one yields good performance while the other suffers from low PRR hence low goodput.
6.3 Energy Overhead

LMAC energy overhead is mainly attributed to CADs performed under DIFS and BO windows. We derive the percentage energy overhead of LMAC in comparison with a conventional ALOHA transmission.

We review the indoor experiment results based on 50 nodes with 10 different traffic demands in Section 6.1 and derive the average number of CADs a node needs to perform to send out a frame with the different network contentions, denoted by $N_{CADs}$. $N_{CADs}$ accounts to average CADs performed during DIFS and BO windows across multiple SFs. From that, the average energy overhead per frame of LMAC, denoted by $E_{LMAC}$, can be computed as $E_{LMAC} = E_{CAD} \times N_{CADs}$, where $E_{CAD}$ denotes the average energy consumption per CAD across all SFs. Therefore, the per-frame percentage energy overhead of LMAC is estimated as $E_{LMAC} = E_{CAD} \times 100\%$, where $E_{frame}$ denotes the transmission energy for a single LoRa frame averaged across all SFs. We estimate through energy measurements that $E_{CAD}$ and $E_{frame}$ correspond to 0.356 mJ and 0.31 J, respectively. We also find through measurements that $N_{CADs}$ is similar across LMAC-2 and LMAC-3, which is also signified by the per frame energy presented in Figure 14(c). Therefore, we compute the energy overhead of LMAC as a single entity for both LMAC-2 and LMAC-3. We then extend the percentage energy overheads of LMAC for varied network demands in Figure 20. We see an energy overhead of about 20% to 25%, which does not significantly grow with the increase of network demands (thus higher contentions).

Simply computing LMAC overhead in comparison with ALOHA frame energy $E_{frame}$ does not fairly represent a real LPWAN network’s performance that suffers from frame loss. We factor in the frame loss when considering the energy overhead of goodput and derive the PRR calibrated energy overhead as expressed in Equation (1):

$$\left[\frac{E_{LMAC} + E_{frame}(LMAC)}{E_{frame}(ALOHA)} - 1\right] \times 100\%,$$

where $E_{frame}(LMAC)$ and $E_{frame}(ALOHA)$ correspond to the PRR calibrated per frame energy—that is, $\frac{E_{frame}}{PRR(LMAC)}$ and $\frac{E_{frame}}{PRR(ALOHA)}$—respectively.
Table 1. Minimum SNR Requirements of LoRa SFs for Successful Reception at the Gateway [24, 36]

<table>
<thead>
<tr>
<th>SF</th>
<th>SF7</th>
<th>SF8</th>
<th>SF9</th>
<th>SF10</th>
<th>SF11</th>
<th>SF12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Required SNR</strong></td>
<td>-6.5 dB</td>
<td>-8.5 dB</td>
<td>-11 dB</td>
<td>-13.5 dB</td>
<td>-18.5 dB</td>
<td>-21 dB</td>
</tr>
</tbody>
</table>

A negative SNR implies that LoRa frames can traverse under the noise floor.

The two quantities characterize the energy costs per successfully transmitted frame with LMAC and ALOHA. We present the PRR calibrated energy overhead of LMAC atop ALOHA in Figure 20 as well. The results show negative overhead and attest that significant energy gains (40% to over 80% energy savings) are achieved with LMAC, especially under high network contentions where ALOHA is subject to much higher frame loss. Figure 20 suggests that at a light energy cost in CADs, LMAC is able to greatly improve the PRR, which gains energy savings in achieving the same amount of goodput.

### 6.4 QoS and Interoperability

This section discusses the benefits of adopting LMAC from a network service providers’ perspective. First, we experimentally show how the adoption of LMAC could incentivize network service providers by adding to their QoS. Second, as a single geographical region likely receives overlapping coverage from multiple network service providers, we give consideration to the fact that MAC-layer changes will be adopted gradually across different network providers. Therefore, from an interoperability perspective, both LMAC and ALOHA nodes are bound to co-exist until all ALOHA nodes are phased out. As such, we discuss the implications of gradually introducing LMAC to ongoing deployments. Third, the LoRaWAN standard must conform to global regulatory policies. We study regulations from multiple regions and identify possible restrictions that may limit the default use of LMAC. We propose simple modifications to LMACs to ensure full compliance under said regions.

For all evaluations in this section, we choose to compare the state-of-the-art ALOHA directly with LMAC-2. The choice of LMAC-2 over other LMACs is motivated by the fact that LMAC-2 allows for the easiest global regulatory compliance (discussed later) and the least changes to existing infrastructure (e.g., gateway functionality).

**LMAC from a QoS perspective.** A network’s ability to best serve all nodes is a qualitative indicator of the service provided by a network service provider. Under typical LoRa deployment conditions, frames from the farthest nodes arrive at the gateway with the weakest instantaneous signal power compared to those from nearby nodes. Therefore, to increase reliability, nodes farthest from the gateway are configured to utilize lowest data rates (e.g., SF12), whereas the closest utilize faster data rates (e.g., SF7) to save energy. However, not all node positions can be optimized for highest reception at the gateway. Often, a part of the network is placed farthest from the gateway or naturally receives higher environmental obstructions. Frames from these nodes undergo significant path loss, and we refer to those as weak nodes. The instantaneous signal strength of frames from the weak nodes is much lower compared to that from nearby nodes. A key problem with weak nodes is that their frames are most vulnerable to arbitrary inference. Therefore, improving the reception capabilities of the overall network including for those weak nodes is a key challenge faced by network service providers. Today, it is deemed the last mile problem in improving QoS for all LPWANs.

Improving QoS for weak nodes is non-trivial. Although a naive solution would be to augment network reception by simply adding more gateways, it is challenging to consistently keep up with moving nodes [8, 38] and changing network dynamics [40, 44]. However, Table 1 shows the minimum SNRs required for demodulation by a LoRa gateway for data rates ranging from SF7 to SF12.
It can be seen from Table 1 that SF12 offers an additional link budget of 14.5 dB compared to SF7. SF12 frames are the most resilient of all SFs and are therefore chosen to be the most suitable data rate for weak nodes. Given that weak nodes mostly utilize slowest data rates, another question that arises is to what degree could LoRa’s orthogonality protect frames from weak nodes? One would expect that weak node transmissions will not be interfered by other concurrent orthogonal LoRa frames. However, this scenario has limitations. LoRa’s orthogonality works such that for a given concurrent transmission and a select SF, chirps from other SFs only contribute toward increasing noise power during the process of demodulation. Recall that frames from weak nodes are the most vulnerable to interferences. This results in lowered SNRs for all concurrent frames, but this effect mostly impacts frames from weak nodes. Therefore, nearby concurrent transmissions, despite orthogonality, could impair the reception of weak nodes by increasing the noise floor. A study conducted by Semtech [36] based on network traces collected from network providers estimates that as much as 18% of SF12 frames can be affected by close proximity SF7 frames.

Given this phenomena, we conduct an experiment to evaluate the impact of using LMAC on the QoS for weak nodes. We set up an outdoor experiment with 16 nearby nodes and 1 weak node. The 16 SF7 nodes were placed within a radius of 300 m from the gateway to emulate interference from nearby nodes. A weak node was emulated by placing an SF12 node at a distance of 1,000 m from the gateway. Nodes within the close proximity utilized a fixed transmission power of 10 dBm and a faster data rate of SF7. The weak node communicated using a slower data rate of SF12 but at a higher transmission power of 15 dBm. All nodes utilized the same channel and none were positioned line of sight to the gateway, to ensure realistic conditions. During the experiment, the nearby nodes transmitted at varying demands under ALOHA and LMAC-2 while the weak node transmitted a known number of frames for each setting. At the gateway, the PRR and SNR of received weak frames (for SF12) were recorded.

Figure 21 presents the SNR and PRR variations of received weak frames. As shown Figure 21(a), the SNRs of weak frames under LMAC-2 is less impacted under increasing demands of nearby nodes. In contrast, when nearby nodes follow ALOHA, Figure 21(b) shows that the SNR of weak frames are decreased at higher demands. The SNR gains of weak nodes by adopting LMAC-2 translate directly to PRR gains as shown in Figure 21(c), as higher SNR frames are more likely to be received by the gateway.

The SNR improvements of weak nodes under LMAC-2 are a result of reduced collisions across nearby transmissions which in turn lead to reduced instantaneous interference to weak frames. Figure 22 further illustrates this phenomena in terms of a spectrogram that represents interference to a weak frame under two scenarios. CH1 shows interference to a frame from a weak node when nearby nodes adopt LMAC-2, whereas CH2 represents interference to a frame from a weak node when nearby nodes adopt ALOHA. Recall that LMAC-2 nodes perform carrier sens-
Fig. 22. Interference to an SF12 frame under two scenarios are showcased. (a) Under LMAC-2, CH1 experiences less instantaneous interference to the SF12 frame, as nearby SF7 transmissions are less likely to collide with each other. (b) CH2 shows a scenario under ALOHA where the absence of carrier sensing has caused two nearby SF7 frames to collide, resulting in higher instantaneous interference to the SF12 frame.

...and therefore spread potentially colliding frames across time. As such, the likelihood of same SF collisions across a channel is significantly reduced under LMAC, whereas ALOHA takes no effort to minimize collisions. Therefore, it can be observed that a weak SF12 frame is less likely to be interfered with two or more SF7 frames under LMAC-2. This results in less instantaneous interference on the weak SF12 frame by nearby SF7 frames. In comparison, under ALOHA, two SF7 frames colliding on top of a weak SF12 frame causes double the instantaneous interference.

*Hidden terminals.* CSMA relies on the ability of two nodes to sense each other’s transmissions. Hidden terminals occur when nodes are unable to sense each other’s transmissions but are likely to be received by the gateway. If two hidden terminals transmit simultaneously, a collision occurs at the gateway. Note that hidden terminals and ALOHA nodes behave indifferently as far as carrier sensing is concerned. LMAC, as a CSMA protocol, is also subject to the hidden terminal problem. In addition, the ISM band is not regulated to perform carrier sensing in all regions of the world. This allows the existence of some nodes that do not perform carrier sensing among nodes that choose to perform carrier sense. At the same time, even though some countries enforce strict Listen Before Talk, their regulations mandate the use of RSS-based sensing, which is not an efficient carrier sense mechanism for LoRa [5, 18, 32]. Furthermore, the adoption of LMAC is also gradual, and more time is needed to transform an entire network to utilize CSMA capabilities. Therefore, at any given time, networks are likely to contain a combination of ALOHA and LMAC nodes. As a result, networks are forced to co-exist with both types of nodes.

To evaluate the impact on the performance of an LMAC-2 network when co-existing with hidden terminals, the quantity of hidden terminal needs to be controlled. However, emulating and controlling large numbers of naturally occurring hidden terminals is challenging. Therefore, for the sake of simplicity, we emulate hidden terminals with nodes that do not perform any type of carrier sensing (i.e., ALOHA). This enables experiments to be conducted such that the impact on LMAC-2 can be evaluated under different combinations of hidden terminals.

Figure 23 presents the performance of LMAC-2 under different hidden terminal combinations. Results suggest that even with 70% of hidden terminals, an extra 20% of throughput gains and 26% of PRR gains can be realized as compared to running pure ALOHA. In addition, with 40% of hidden terminals, an extra 40% of throughput gains and 71% of PRR gains can be realized as compared to a network that only conforms to ALOHA.

*LMAC from a regulatory perspective.* The choice of LMAC should not solely be driven by the need for higher network performance but also should ensure regulatory compliance.
Although not mandated by most, some regions may mandate equal use of channels for devices operating within the ISM spectrum, such as the Federal Communications Commission (FCC) [5]. The FCC requires technologies that communicate using multiple PHY channels (e.g., LoRaWAN) to utilize each channel equally. The LoRaWAN standard under ALOHA achieves this by making a pseudo random choice for the transmission channel while regularly checking with local logs to ensure that no channels are overutilized. In contrast, LMAC-2 and LMAC-3 are designed with more intelligence into the transmitter and therefore make informed decisions for the transmission channel to minimize potential collisions. However, minor modifications to protocol are possible for full compliance. They are explained in the following.

As far as LMAC-1 is concerned, it simply controls a node’s behavior on time. LMAC-1’s channel utilization is no different from the existing practice of LoRaWAN. Specifically, instead of directly transmitting a frame, an LMAC-1 node first senses the channel and waits until the channel is free before attempting further transmissions. Therefore, LMAC-1 will be FCC compliant by simply conforming to existing LoRaWAN practices of channel utilization. However, LMAC-2 makes an effort to choose a less crowded channel. As such, the choice of channel in LMAC-2 is not pseudo random. However, recall from Section 5.2 that LMAC-2 still makes a random choice from a list of potentially free channels. Therefore, to comply with content presented in reference [5], a node should memorize previous choices of channels (similar to current practice) and remove overutilized channels from the list of channels made available for future random selection. This allows all channels to be equally used on average by each transmitter in the long run. LMAC-3 can follow a similar approach to that proposed for LMAC-2 to ensure FCC compliance.

7 LORAWAN IMPLICATIONS

The testbed experiments in Section 6 have shown that LMACs improve the performance of LoRa networks. In this section, we apply LMAC to LoRaWAN, which is the prevalent data link layer protocol primarily based on ALOHA. As LoRaWANs are constrained by the channel access time requirements (<1% duty cycle ratio for each end node, etc.), it is difficult to experimentally show the full advantage of LMAC with limited numbers of devices. Instead, we conduct trace-driven simulations to exhibit the LMAC advantage for a scaled LoRaWAN network. We replace the ALOHA MAC in LoRaWAN Class A with LMAC-2 and LoRaWAN Class B with LMAC-3. LMAC-2 does not require information from the gateway, similar to the autonomy of LoRaWAN Class A; LMAC-3 requires gateway feedback that is provided by the beacon slot of LoRaWAN Class B.
7.1 Data Rates of LMAC-Based LoRaWANs

The data traces used are collected from 30 experiments each spanning 1 minute for each of the three protocols. Thus, there is a 10-minute trace in total for each protocol. Trace frames are added with LoRaWAN headers and mandatory delays (guard time, RX slots, etc.). With the addition of LoRaWAN overheads, trace durations are recalculated. Once trace duration has been updated, frames received and lost are replayed in the simulation as per trace records. The throughput and goodput of the replayed simulation are then recorded and presented in Figure 24.

Despite the low data rate, LMAC could still bring 3.06× and 5.93× performance improvements on LoRaWAN Class A with LMAC-2 and LoRaWAN Class B with LMAC-3, respectively. The simulation results suggest that the performance of LoRaWAN can be improved by utilizing LMAC as the default MAC. A noticeable difference from the results in Section 6.1 is that the data rate of LoRaWAN decreases due to protocol overheads. As most of the overheads are due to the delays imposed for the two RX slots following a TX, the data rate achieved without the imposed delays will be much higher. However, our simulations still impose such delays to be consistent with the current specification of LoRaWAN.

Since LMAC does not alter any part of the existing LoRaWAN protocol except the MAC mechanism, the integration of LMAC to the LoRaWAN protocol stack could be readily achieved. Moreover, as LoRaWAN imposes channel access restrictions for fairness and LMAC can improve the channel access fairness by avoiding occupied or congested channels, the channel access restrictions can be relaxed or even removed from LoRaWAN should LMAC be used.

7.2 Network Capacity

Understanding the maximum number of nodes that can be supported by the LMAC-based LoRaWAN is important. This section adopts a numerical analysis method to estimate the number of frames that the LMAC-based LoRaWAN can receive in a given period. Such number can be translated to the maximum number of supported end nodes once the per-node demand is given.

Using the results in previous experiments, the total number of chirps observed by the gateway in a channel \( (\mathcal{S}_{CH}^{RX}) \) can be extracted within an observation duration \( (T_{obs}) \). With the preceding data, the observable chirps per second in the channel \( (\mathcal{S}_{CH}) \) can be obtained through \( \frac{\mathcal{S}_{CH}^{RX}}{T_{obs}} \). Having
Fig. 25. Number of successfully received frames in 1 minute versus various parameters varying payload size with SF7 and CR4/8, varying SF with 16-byte payload and CR4/8, and varying CR with 16-byte payload and SF.

$S_{CH}$, an estimated number of frames a channel can receive ($F_{RX}$) can be estimated by simulating different $S_{RX}$ values generated from varying frame parameters, such as payload size, time period, SF, and code rate (CR). With the simulated $S_{RX}$, $F_{RX}$ is computed by $(S_{CH} \times T_{sim}) \div S_{RX}$, where $T_{sim}$ is the simulated time period.

A network analysis using the preceding method to estimate the network capacity when three parameters (payload size, SF, and CR) vary is presented in Figure 25. Varying payload size exhibits an exponential decay in the number of frames received, showing that the payload size plays a vital role in network capacity. The network capacity declines exponentially as the payload size increases due to frame overhead. Frames with a small payload size carry a larger overhead to payload ratio than frames with a large payload size. Varying SF also exhibits similar decline as varying payload. Such decline is due to the non-linear increase in the number of chirps per frame as SF increases. In contrast, varying CR decline is linear. This linearity reflects the linear increase in each CR setting from CR4/5 to CR4/8 where each increase in setting is an additional bit of redundancy added to every four bits of the frame. The results also show that LMAC-2 and LMAC-3 outperform ALOHA by 70% and 83%, respectively, and the numbers of nodes that LMAC-2 and LMAC-3 could support would likely be similar.

8 RELATED WORKS

As the primitive ALOHA scales poorly with the number of LoRa nodes, several studies aimed at improving the resilience and throughput of LoRaWAN networks by PHY layer modification or new MAC designs. We now review these PHY and MAC layer techniques separately.

LoRa PHY. Choir [17] exploits the distinct frequency biases of end devices to disentangle the collided frames. As Choir uses a new PHY algorithm, the gateway of Choir has to use SDR, such as USRP, and employ software demodulation implementations. As such, Choir will lose the line speed of hardware gateways. NetScatter [23] reduces LoRa modulation to on-off keying to simplify the operation on backscatter devices. It can scale for concurrent demodulation of backscatter LoRa signals. However, the unique design of NetScatter’s on-off keying modulation does not conform to standard LoRa modulation. To address the significant attenuation of the signals transmitted from end devices deeply located within buildings, Charm [16] jointly processes the signals received by multiple gateways to increase received SNR and the network coverage. Although Charm does not require a full-fledged SDR device, dedicated programmable hardware design is needed to collect the I/Q data of the received signal. To address channel contention, Chime [19] analyzes a single frame from a node through multiple gateways. With gateway synchronization and optimal frequency estimation, a gateway advises the node on the optimal transmission channel. Chime requires
SDR to act as gateway to collect the I/Q data and strict synchronization between gateways. Voigt et al. [43] exploit the use of multiple LoRa gateways and directional antennae to aid in decoding interfered LoRa frames [43]. However, directional antennae impede a node’s ability to be heard by multiple gateways, hence coverage of a node.

LoRa MAC. Slotted ALOHA and the more well scheduled time-division multiple access can reduce collisions compared with ALOHA. Specifically, the end node will transmit in certain time slots to alleviate the collision issue caused by random access in ALOHA. Several studies [22, 34, 42] focus on the clock synchronization service needed by slotted ALOHA and time-division multiple access. However, due to LoRa’s limited communication bandwidth [20], the communications for clock synchronization present a burden for LoRaWAN networks. Thus, CSMA is desirable in LoRaWANs, as it does not require clock synchronization. Beltramelli et al. [9] present a stochastic geometry-based model [9] to analytically show that CSMA outperforms ALOHA in terms of reliability and energy efficiency. However, full-fledged implementation of CSMA in LoRaWANs is still unavailable due to the lack of hardware support for carrier sense. To achieve carrier sense, DeepSense [11] detects on-air transmission of LPWAN radios by storing spectrograms as an image, then performing signal identification on the image with artificial neural networks. However, an additional SDR device and machine learning accelerator are needed. To implement CSMA with off-the-shelf LoRa devices, a previous work [33] ports the 802.11 CSMA protocol to LoRa, in which the carrier sense is performed by a stand-alone node near the transmitter. This introduces additional monetary cost and system complexity. In this work, we leverage CAD for carrier sense on LoRa end devices and design LMAC to improve the goodput and PRR of LoRaWAN networks. LMAC requires no modifications to the LoRa PHY layer and the LoRaWAN specification. Thus, it can be readily deployed in the current LoRaWAN networks. To the best of our knowledge, this article presents the first systematic research and implementation of CSMA for LoRa networks.

9 CONCLUSION

This article showed that the CAD feature of LoRa radios for preamble chirp detection can be exploited for reliable carrier sense. Based on CAD, this article designed three versions of LMAC that respectively implement CSMA for LoRa networks, and balance the loads of the channels defined by frequencies and SFs by using the end nodes’ local information only and then additionally the gateway’s global information. Testbed experiments showed that, compared with ALOHA, LMAC brings significant performance improvements in terms of PRR and goodput, as well as radio energy saving per successful frame delivery.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers and shepherd for their valuable comments and helpful suggestions that improved the quality of this article. The code of this work is built with the LMiC (LoRaMAC-in-C) library ported by Matthijs Kooijman.

REFERENCES


[34] Rajeev Piyare, Amy L. Murphy, Michele Magno, and Luca Benini. 2018. On-demand LoRa: Asynchronous TDMA for energy efficient and low latency communication in IoT. Sensors 18, 11 (2018), 3718.


[38] Zehua Sun, Huanqi Yang, Kai Liu, Zhiyue Liu, and Weitao Xu. 2022. Recent advances in LoRa: A comprehensive survey. ACM Transactions on Sensor Networks. Accepted.


Received 7 July 2022; revised 1 September 2022; accepted 15 September 2022