Understanding collisions in a LoRaWAN

Andri Rahmadhani Fernando Kuipers
Delft University of Technology
F.A.Kuipers@tudelft.nl

ABSTRACT

Long Range (LoRa) technology has quickly gained popularity as a radio technology for IoT devices. LoRa utilizes the license-free ISM-bands and its network adheres to the LoRaWAN specification.

In this report, we investigate LoRaWAN frame collisions, particularly with respect to the capture effect, through small-scale measurements and an evaluation on a large-scale network.

1 INTRODUCTION

The Long-Range (LoRa) radio technology uses a chirp spread spectrum (CSS) modulation to improve robustness against noise and increase coverage, while maintaining low power consumption. LoRaWAN, a de-facto standard for LoRa-based networks, defines the communication protocol and system architecture for the network [8]. The LoRaWAN specification regulates the use of LoRa technology in the unlicensed Industrial, Scientific, and Medical (ISM) frequency bands.

The interest in developing and adopting LoRa technology is rapidly increasing. For instance, one of the largest network operators in the Netherlands, KPN, has announced full LoRa coverage across the country. The Things Network (TTN), an open community which crowdsources the LoRaWAN network, already has wide presence across the globe. Furthermore, the use of unlicensed ISM bands opens up the possibility for individuals to create private LoRaWAN networks. However, when the number of LoRa devices registered to a particular network provider grows significantly, aggravated by the use cases that require frequent packet transmission, the capacity of the network will quickly saturate, leading to performance degradation. This problem is also intensified by the fact that all gateways in the vicinity, regardless to which network provider they belong, will receive packets transmitted by any LoRa device, causing inter-network interference.

Inter-network interference leads to frame collisions and attributes to packet loss. In LoRa, frame collisions happen when two or more packets overlap in time and use the same LoRa parameters, i.e. the same spreading factor (SF), bandwidth (BW), and carrier frequency (CF). Thanks to capture effect found in LoRa modulation, a packet received with a higher power level, at least 6 dB stronger, can still be decoded during collision [3]. In this report we study the capture effect in more detail.

The probability of frame collisions also depends on the selection of LoRa parameters used for transmitting a packet. When most devices use the same configuration, the chance of getting frame collisions will be higher. Furthermore, the selection of LoRa spreading factor and transmission power will influence the coverage area, meaning that due to the signal attenuation over distance, some packets will not collide. The probability of frame collisions is also affected by traffic characteristics, particularly on the periodicity of transmission and payload size. Larger payload size and more frequent transmissions accumulate to higher time on air and channel occupancy.

According to the simulation in [19], using directional antennas at the device side can reduce inter-network interference, but not as effective as deploying more gateways [19]. By adding and spreading more gateways, the probability of obtaining 6 dB power difference will be higher due to the propagation loss over distance.

In this report, we characterize LoRaWAN frame collisions through practical measurements. We analyze several aspects, including coverage, traffic characteristics, packet loss, signal quality, and LoRa parameter distributions.

2 LORA AND LORAWAN OVERVIEW

2.1 LoRa

LoRa adopts Chirp Spread Spectrum (CSS) as its modulation scheme. LoRa CSS modulation encodes symbols into multiple signals of increasing (up-chirp) or decreasing frequencies (down-chirp) with a constant chirp rate. It utilizes Forward Error Correction (FEC) in combination with whitening and interleaving that gives longer range than typical Frequency Shift Keying (FSK) modulation [14].

A LoRa frame is initialized by a long constant-chirp preamble, which is used by the receiver to lock the LoRa signal. The preamble is followed by two reverse-chirps used as a synchronization word, indicating the end of the preamble. Since there is no disparity between the preamble from one LoRa transmitter and another transmitter, the receiver may listen to an unwanted signal. Therefore, the detection of a preamble is very crucial to the reception of LoRa frames.

LoRa transmission can be characterized by several key parameters:

- **Bandwidth (BW)**, defines the width of radio frequencies being used in the transmission band. On the one hand, an increase in signal BW allows the use of a higher effective data rate, thereby reducing the time on air at the cost of reduced receiver sensitivity due to noise accumulation [17]. On the other hand, a lower BW requires more accurate crystals to achieve a higher sensitivity at the expense of higher chip cost and lower data rate. LoRa radios support BW from 7.8 kHz to 500 kHz, but only BW of 125 kHz, 250 kHz, and 500 kHz are used in LoRaWAN.
- **Carrier Frequency (CF)**, defines the frequency of the medium used for LoRa transmission. Several LoRa chipsets support frequency bands from 137 MHz to 1020 Hz [17] and can be programmed with a resolution of 61.035 Hz.
- **Coding Rate (CR)**, defines the FEC rate used by the LoRa modem to protect against bursts of interference. It can be configured to either 4/5, 4/6, 4/7 or 4/8. The higher the

CR, the more protection can be offered, but it prolongs the time on air.

- **Spreading Factor (SF)**, represents the ratio between symbol rate and chirp rate. One symbol is encoded into 2^{SF} chirps. A higher spreading factor improves signal to noise ratio (SNR), resulting in greater sensitivity and range, but it also increases the airtime of the packet and power consumption. LoRa supports SF6 to SF12, but only SF7 to SF12 are used in LoRaWAN. Those SFs are orthogonal given the same bandwidth/frequency.
- **Transmission Power (PWR)**, defines the transmission power used for transmission. LoRa chipsets support PWR ranges from -4 dBm to 20 dBm. However, due to hardware limitations, PWR higher than 17 dBm is limited to a duty cycle of 1% [17].

LoRa supports explicit and implicit frame formats. In explicit format, a frame consists of a header, which includes information of payload length in bytes, CR, and whether the payload 16-bit payload CRC is used or not. In implicit format, the header and its CRC are not included and must be manually configured in the transmitter-receiver pair, thereby reducing transmission time [17]. The LoRaWAN specification utilizes the explicit frame format, both for uplink and downlink communications [10]. Figure 1 illustrates a LoRa frame format.

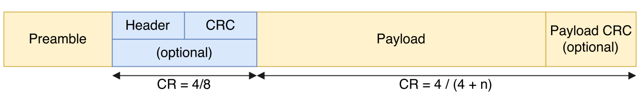


Figure 1: LoRa frame format with CR $n \in \{1..4\}$.

The preamble length can be configured from 6 to 65535 symbols. The LoRa modem then adds 4.25 symbols, representing a synchronization word. The header has a fixed FEC rate of 4/8, and the payload has a variable length ranging from 1 to 255 bytes. In LoRaWAN, 8 symbols are used for the preamble.

2.2 LoRaWAN

LoRaWAN is a specification of a Media Access Control (MAC) protocol for controlling low-powered devices in wide area networks. The first specification was released by the LoRa Alliance in January 2015. As of this writing, the latest specification is version 1.1 and was released in October 2017; however, the old version is still supported.

2.2.1 Network Architecture. The LoRaWAN network utilizes a star-shaped network topology, which is different from a mesh network used typically in wireless sensor networks. A typical LoRaWAN network architecture is illustrated in Figure 2 and it comprises:

- **End-device/node/mote**, a low-powered module (sometimes equipped with sensors) that communicates with gateways using LoRa's modulation scheme or FSK, as explained in the LoRaWAN specification.
- **Gateway/base station**, an intermediate device that receives messages from end-devices and forwards the messages to the network server over an IP-based backhaul, such as Ethernet, 3G, and 4G.
- **Network server**, a server that handles messages, including message duplication removal, decoding, and link configuration adjustment of end-devices.

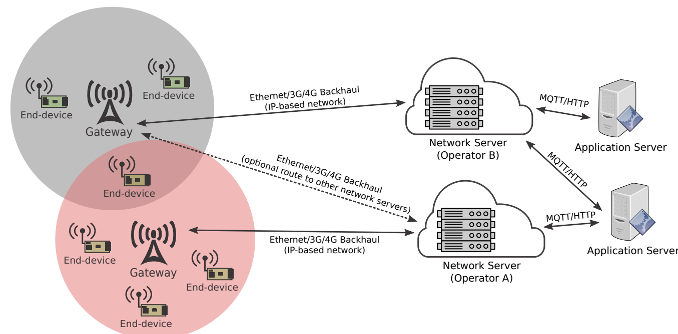


Figure 2: Typical LoRaWAN network architecture.

An end-device connects to the network directly through gateways in a one-hop manner, reducing protocol complexity and saving battery life. Also, it does not have a complex and power-hungry handover process as in cellular networks. However, a message transmitted from an end-device is received by all gateways in range, and each gateway forwards the message to the network server, resulting in message duplication. The network server then performs message deduplication. Transmission from a network server to an end-device is established through a gateway that is selected based on certain algorithms, typically selecting the best signal reception.

2.2.2 Operating Classes. Three different operating classes are introduced to address various applications:

- **Class A**: The end-device triggers a communication channel and transmits its message immediately based on its preference, similar to ALOHA-type protocols. Each uplink transmission is followed by two receive windows, namely RX1 and RX2. As a result, a downlink message from the server at any other time needs to wait for the next uplink message. Class A operation is mandatory for every LoRaWAN-compatible device, and it provides the lowest power consumption.
- **Class B**: End-devices open more receive windows at scheduled times in addition to Class A. To synchronize the timing, the gateway transmits beacon frames transmitted at regular intervals.
- **Class C**: End-devices continuously open receive windows. Class C operating class, in fact, consumes most power, but has the lowest latency.

In this report, we focus only on class A devices as it is mandatory for every LoRaWAN device. Figure 3 depicts uplink and downlink transmissions for LoRaWAN class A.

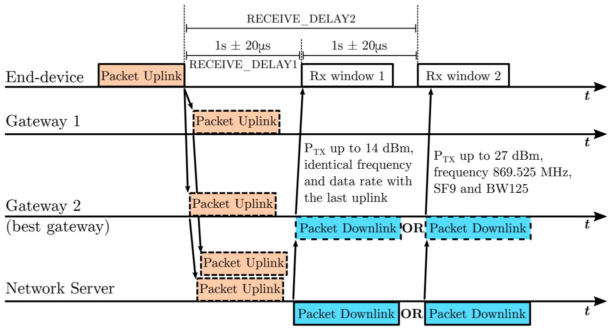


Figure 3: Uplink and downlink transmissions for LoRaWAN class A.

2.2.3 Device Activation. Devices can join a LoRaWAN network by using two methods:

- **Over-the-Air-Activation (OTAA)**: Devices initiate a join-procedure with the network by sending a join request message, which contains a globally unique device identifier (DevEUI), an application identifier (AppEUI), and a 128-bit AES encrypted application key (AppKey). The network then sends a join-accept message comprising a device address (DevAddr), a network session key (NwkSKey), and an application session key (AppSKey). These session data are then stored in the device's memory. Whenever an end-device loses its network session info, it has to do a new join-procedure. OTAA offers a fairly secure way to join a network as the network session info is dynamically assigned by the network.
- **Activation by Personalization (ABP)**: Devices use pre-programmed network session info (DevAddr, NwkSKey, AppSKey), and thus do not need to initiate a join-procedure, although this is less secure than OTAA. DevEUI and AppEUI are not required for ABP.

2.2.4 Frequency Regulation and Access Policy. LoRaWAN regulates operating frequency bands based on regions [9]. In Europe (EU), LoRaWAN allows the use of frequency bands 433 MHz (EU433) and 868 MHz (EU863-870). This paper will focus on the 868 MHz frequency band, given that most of LoRaWAN devices available in the market today operate on that frequency band.

The LoRaWAN specification defines three different 125 kHz channels for the 868 MHz band, namely 868.1, 868.3, and 868.5 MHz. These channels must be supported by all devices and networks as they are used for the join-procedure. The network can instruct the devices to add extra channels to their channel set, up to 16 channels for EU863-870, after joining. These channels are used for uplink and downlink communications. In the downlink, the first receive window RX1 uses the same frequency and data rate as the uplink. For the second receive window RX2, LoRaWAN uses a fixed parameter that can be pre-configured. The default RX2 parameter uses frequency 869.525 MHz, with SF12, and 125 kHz bandwidth.

The European frequency regulation (ETSI) imposes the use of either duty cycle or Listen Before Talk in combination with Adaptive Frequency Agility (LBT-AFA) access policy. The latest LoRaWAN specification exclusively uses the duty cycle access policy [9]. Duty cycle represents the percentage of time an end-device occupies a particular channel. For example, a device that uses a channel for 1 s should wait for 99 s for the next transmission when a 1% duty cycle is imposed. The waiting time or off-time can be expressed as:

$$T_{\text{off}} = \frac{\text{TimeOnAir}}{\text{DutyCycle}_{\text{subband}}} - \text{TimeOnAir} \quad (1)$$

Time on air denotes the time needed for transmitting a frame from node to gateway or vice versa. It can be derived from the symbol time T_S by solving the following set of equations [17]:

$$T_S = \frac{2^{SF}}{BW}\,, \quad (2)$$

$$n_{\text{payload}} = 8 + max\left(\left\lceil \frac{8PL-4SF+28+16CRC-20IH}{4\times(SF-2DE)} \right\rceil \times (CR+4)\,, 0\right)\,, \quad (3)$$

$$T_{\text{payload}} = n_{\text{payload}} \times T_S\,, \quad (4)$$

$$T_{\text{preamble}} = \left(n_{\text{preamble}} + 4.25\right) \times T_S\,, \quad (5)$$

$$T_{\text{packet}} = T_{\text{preamble}} + T_{\text{payload}}\,, \quad (6)$$

where n_{payload} denotes the number of symbols used for transmitting the PHY payload. n_{preamble} represents the programmed preamble length, i.e. 8 for LoRaWAN, to which is added a 4.25 symbols synchronization word. PL denotes the size of the PHY payload in bytes (1 to 255), and SF denotes the spreading factor. CRC represents whether 16-bit CRC is used at the end of the payload or not. IH denotes implicit header. $IH = 0$ when PHY header is enabled and $IH = 1$ when no header is present. DE represents low data rate optimization. DE is enabled when the symbol time is larger than 16 ms, e.g. for SF11 and SF12 with 125 kHz bandwidth. CR denotes coding rate, for which a value 1 corresponds to coding rate 4/5 and 4 to 4/8.

ETSI specifies two bands and five sub-bands for EU863-870 as listed in Table 1 [15]. The duty cycles should be respected by both end-devices and gateways, therefore making downlink messages relatively expensive.

Table 1: ETSI bands and sub-bands for EU863-870.

Band	Frequency range (MHz)	Power / Field	Spectrum Access
g(Note 7)	865-868	+6.2 dBm / 100 kHz	1% or LBT AFA
g(Note 7)	865-870	-0.8 dBm / 100 kHz	0.1% or LBT AFA
g1	868-868.6	14 dBm	1% or LBT AFA
g2	868.7-869.2	14 dBm	0.1% or LBT AFA
g3	869.4-869.65	27 dBm	10% or LBT AFA
g4	869.7-870	7 dBm	No requirement
g4	869.7-870	14 dBm	1 % or LBT AFA

LoRaWAN defines six transmission power options: 20, 14, 11, 8, 5, and 2 dBm. The default transmission power of devices for EU863-870 is 14 dBm and is limited to this value, except for the g3 sub-band that can go up to 27 dBm, which is typically used for downlink.

2.2.5 Adaptive Data Rate. Adaptive Data Rate (ADR) is a technique used for adjusting the actual device data rate to ensure performance, reliability, and capacity of the network. For instance, devices close to gateway will use a higher data rate, leading to a shorter time on air, and lower output power. Table 2 shows LoRa modulation data rates used in ADR.

Table 2: LoRa transmission data rate for EU863-870.

Data rate (DR)	Spreading Factor	Bandwidth (kHz)	Bit rate (bit/s)	Max. app. payload (bytes)
0	SF12	125	250	51
1	SF11	125	440	51
2	SF10	125	980	51
3	SF9	125	1,760	115
4	SF8	125	3,125	242
5	SF7	125	5,470	242
6	SF7	250	11,000	242

ADR can be initiated either by the network or the device. The device sends uplink messages with the ADR bit set. These ADR-enabled messages are collected in the network and are calculated based on certain algorithms. ADR messages typically consist of data rate, transmission power, usable channels, and the number of retransmissions. These settings are carried by the *LinkADRReq* command sent by the network. If the new parameters are applied successfully, the end-device will acknowledge them by sending *LinkADRAns*. It should be noted that the network server uses ADR only to instruct the device to switch to a higher data rate, whilst the device can switch to a lower data rate when the network server does not respond to the *ADRACKReq* message sent by the device.

The algorithm for ADR calculation is not explained in the LoRaWAN specification. It is up to the network operators to implement their own algorithms. As ADR is calculated and analyzed on the network back-end, it drastically reduces the complexity on the end-devices, but requires downlink transmission, which is relatively expensive due to the duty cycle limitations imposed on the gateways as well.

3 FRAME COLLISION CHARACTERIZATION

3.1 Small-scale Experiment

3.1.1 Single Gateway. We have performed small-scale experiments using two nodes and a single-channel gateway separated 5 m apart. Figure 4 depicts the experiment setup. Two scenarios were deployed: (1) equal received power (P_{RX}), and (2) different transmission powers (P_{TX}).

In the first scenario, both nodes used $P_{TX} = 2$ dBm. The node positions were adjusted to obtain similar RSSI ($P_{RX} \approx -70$dBm) values at the gateway. The experiment was performed using only one data rate, i.e. DR4 (SF8BW125), as a similar experiment was done in [3]. In the second scenario, node 1 (N1) and node 2 (N2) used $P_{TX} = 8$ dBm and $P_{TX} = 2$ dBm, respectively. The experiment used data rates DR0 to DR5, which are typically used in real-world LoRaWAN networks. For both scenarios, we used a fixed 26-bytes application payload that was sent with a regular period tailored to the duty cycle, which varies over the different data rates. The stronger node, N1, was delayed with increments of 1 ms. Each node sent 20 packets for each time offset. Both devices were connected to a PC via the serial port and were triggered to transmit packets simultaneously (using Python script). The transmissions used frequency 869.7 MHz, which is relatively free compared to the default LoRaWAN channels. The gateway, which was built using ESP8266 and a SX1276 LoRa module, sent the received frames to the PC via WLAN.

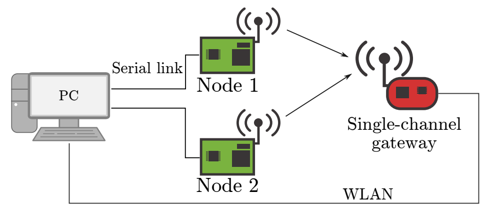
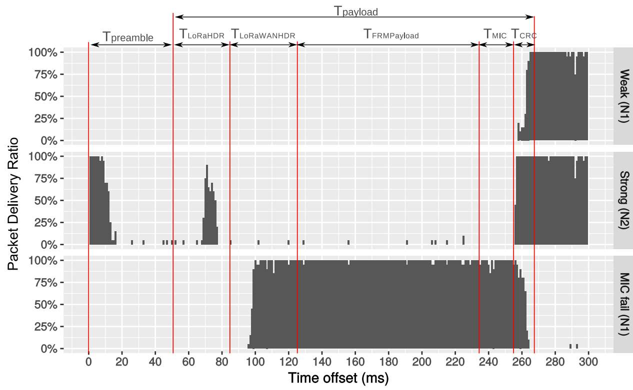


Figure 4: Frame collision experiment setup.

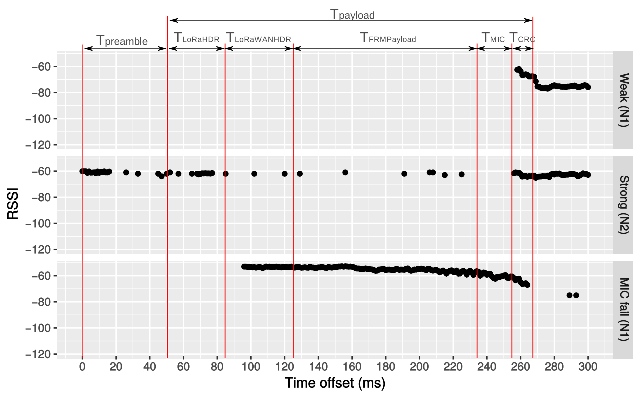
Similar to [3], the result for the first scenario, particularly in terms of packet reception rate (PRR), is difficult to interpret as the gateway indefinitely switches its state, either reading the signal from N1 or N2. However, we observe that the SNR values tend towards zero when most part of the frames overlap. It indicates that an unwanted LoRa signal having the same data rate, frequency, and power may lead to destructive interference. For the second scenario, we obtain a similar PRR, RSSI, and SNR pattern for all data rate combinations as shown in Figure 5. For the sake of clarity, we only display the result from the experiment with DR3. We also add timing information obtained by calculating the time on air using Equation 2 to 6. We measured the RSSI levels for both nodes and found the average RSSI difference to be approximately 15 dB, which is far above the 6 dB requirement for the stronger frame to be decoded correctly [6]. The corrupt frames belonging to the weak node have higher RSSI levels due to the superposition of both weak and stronger frames, and it is proportional to the length of the frame being interfered. The SNR values for the corrupted frame have similar patterns to the previous scenario.

Figure 6 illustrates several cases in which frame loss may occur. A clarification follows:

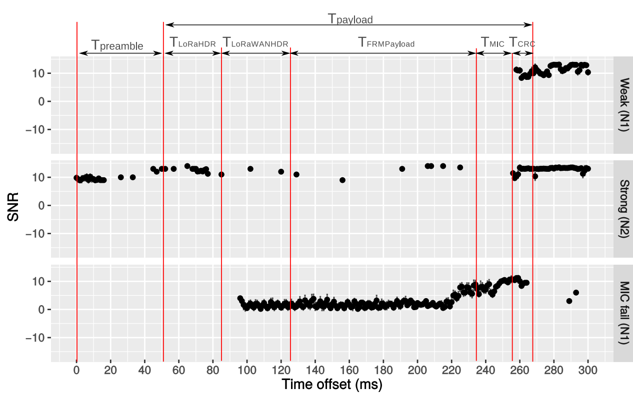
- **Case 1**. Both frames are considered lost when the stronger frame arrives later than the receiver locking time, i.e. 4 symbols. The receiver supposedly listens to the weaker frame signal, but it is suppressed by the stronger frame signal.
- **Case 2**. The stronger frame survives the collision when its arrival overlaps with the CRC header of the weaker frame. This will cause the receiver to release the lock on the weaker frame and start listening to the new frame.
- **Case 3**. Both frames are considered lost when the stronger frame arrives after the receiver finishes receiving the weak frame header, and the stronger frame overlaps with the LoRaWAN header of the weaker frame. This happens because the receiver keeps locking on the weak frame, whilst



(a) Packet reception rate.



(b) RSSI.

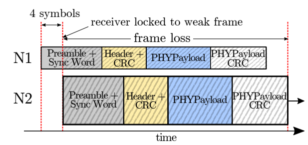


(c) SNR.

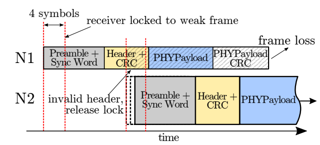
Figure 5: Frame collision result for different transmission powers using DR3 (SF9BW125). The time offset is calculated with respect to N1.

the data inside the LoRaWAN header, e.g. device address, becomes corrupt.

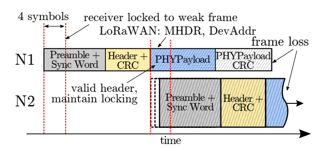
- **Case 4**. Both frames are considered lost when the stronger frame arrives after the receiver finishes receiving the LoRaWAN header of the weaker frame and slightly before the payload CRC of the weaker frame. In this condition, the weaker frame is successfully received, but the PHYPayload gets destroyed, resulting in a wrong message integrity check (MIC). Thanks to error correction techniques employed in the LoRa transceiver, the weaker frame can still be decoded whenever the stronger frame slightly overlaps with the payload CRC of the weaker frame.



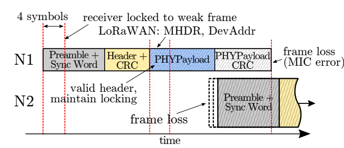
(a) Case 1.



(b) Case 2.



(c) Case 3.



(d) Case 4.

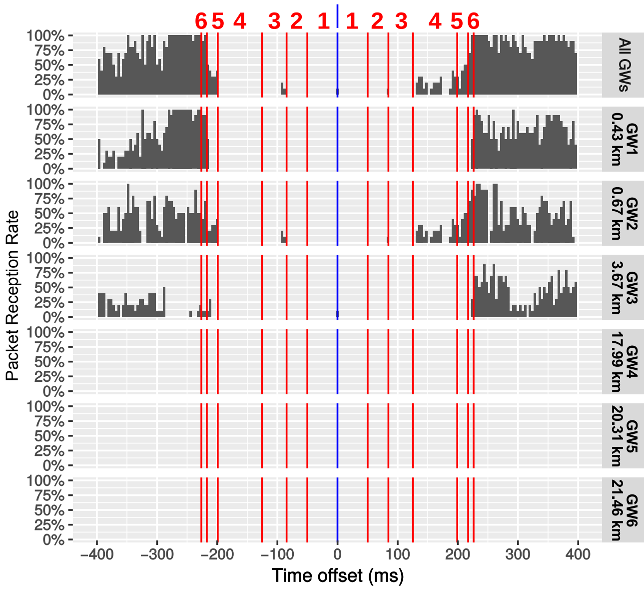
Figure 6: Conditions at which frame loss may occur.

The analysis of the experiment with different transmission power is in line with the results reported in [7]. Our findings deliver new insights into frame loss characteristics from the application level perspective, that is, regarding LoRaWAN header corruption and MIC failure.

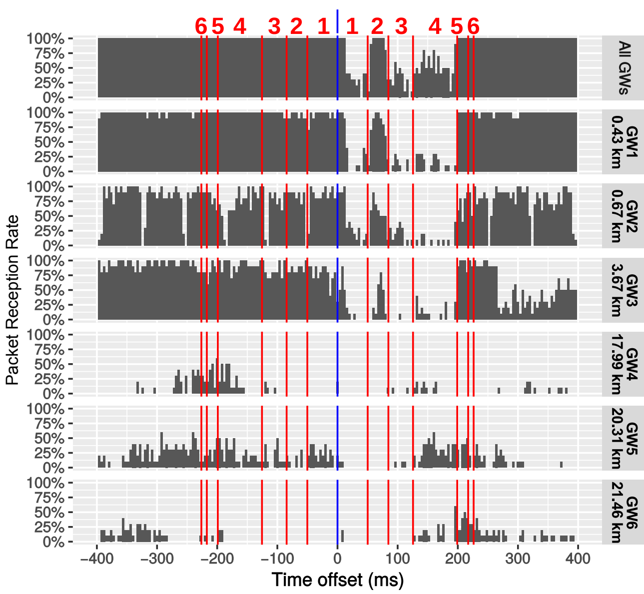
3.1.2 Multiple Gateways. We extend our experiment to observe the impact of adding more gateways. The setup is similar to the single gateway experiment, but at an altitude of 64.9 m above sea level. Both devices were registered to the TTN network using ABP without ADR enabled, and were configured to transmit with the same data rate (SF9BW125), coding rate (4/5), frequency (868.1 MHz), payload size (17 bytes), but with different transmission powers (14 dBm and 8 dBm). The weaker node (N1) was used as a reference for calculating the time offset. Both devices sent 10 frames per time offset. Those frames were received by six gateways at different distances. As the measurements were performed at a relatively high altitude in a fairly open area, the coverage extended up to 21.5 km.

Figure 7 shows PRR from the perspective of the application layer (all GWs), as well as per individual gateway. The number on top of the figure represent frame timing information: (1) preamble, (2) LoRa header, (3) LoRaWAN header, (4) frame payload, (5) MIC, and (6) payload CRC.

The result shows that the PRR is improved, especially for the delayed stronger frames. As the weaker frames could not reach the more distant gateways, i.e. GW4, GW5, and GW6, the gateways would not get a lock on the weaker frames, so that the stronger frames could be decoded properly. The closer gateways, i.e. GW1, GW2, and GW3, also contribute to the higher PRR level as they increase the probability of receiving correct frames even though the probability per gateway remains low, which is similar to the pattern found in the single gateway experiment. For the weaker



(a) Weak node (N1).

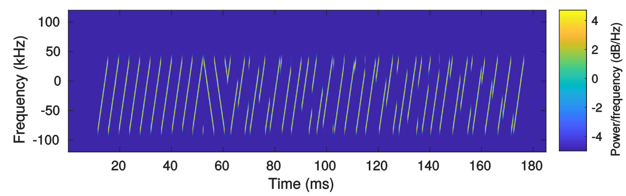


(b) Strong node (N2).

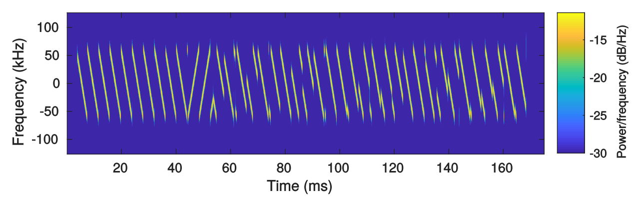
Figure 7: Packet reception rate (PRR) for the collision experiment with multiple gateways. The time offset is calculated with respect to the weak node.

node, the PRR is slightly improved when the stronger frame arrives during the reception of the weaker frame payload. The RSSI difference approaches zero as the distance increases. Considering LoRa signals can be decoded below the noise floor (negative SNR value), the average received power can exceed 6 dB, which is why the pattern is somewhat similar to the previous experiment. From this experiment, we can conclude that adding more gateways does improve the PRR of the stronger node, but it does not significantly increase the PRR of the weaker node.

3.1.3 Uplink and downlink frames. When the coverage of gateways overlap with each other, it is possible that the uplink transmission from one gateway collides with downlink transmissions from other gateways. However, as uplink and downlink frames are transmitted with different polarization by using I/Q inversion, both messages will not collide even though they overlap in time and have the same frequency and data rate [13]. We performed measurements to observe the realization of downlink signal inversion in the currently deployed LoRaWAN networks, for example, in TTN network. Figure 8 shows the spectrogram of LoRaWAN uplink and downlink frames received by a Software Defined Radio device (RTL-SDR).



(a) Uplink transmission.



(b) Downlink transmission.

Figure 8: Spectrogram of uplink and downlink LoRaWAN frames received by RTL-SDR.

Due to the limitation of hardware, we could not experimentally observe uplink and downlink frame collisions. Instead, we performed simulations and obtained that downlink transmissions will be regarded as noise by gateways and uplink transmission will be regarded as noise by the end-device. The possible source of interference is in the sync-pattern, where two inverse chirps are used to mark the end of the preamble.

4 REAL-WORLD LORAWAN NETWORKS

We investigated The Things Network (TTN).

4.1 Data Collection

4.1.1 TTN. The Things Network (TTN) dataset was collected from the TTN network using the TTN API for the month June 2016. We obtained the data from [2] and repeated some of the evaluations. All measurements were sent by Class A devices and were activated using ABP with the Semtech default keys. Only uplink frames originated from TTN-registered devices were available in this dataset. Table 3 shows basic information obtained from the TTN dataset.

Table 3: Basic information extracted from the TTN dataset.

Devices (IDs)	Gateways	Gateway status updates	Received frames	Unique frames
488	229	12,271,366	1,616,500	1,471,719

4.2 TTN Dataset Evaluation

4.3 Network Density

4.3.1 Gateways. From 229 TTN gateways worldwide, up to 111 gateways were located in the Netherlands and were concentrated in major urban areas, such as Amsterdam, The Hague, Rotterdam, and Utrecht. These gateways were placed in fixed locations during the given period. Our analysis will be focused on these areas in order to observe the effect of gateway density on the packet loss ratio.

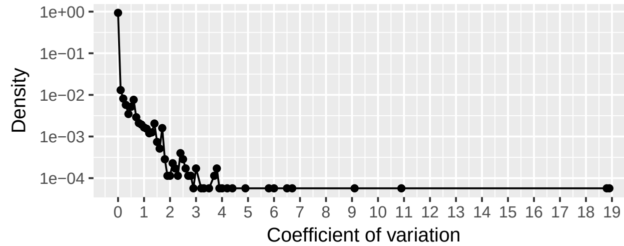
4.3.2 Devices. Some frames contain payloads that carry device locations in the form of GPS lat-lon coordinates, which is encoded either in a string or binary format. Such binary-encoded formats predominantly originated from LoRaMote devices [16]. Up to 27,614 frames carried GPS coordinates but only 6460 frames, sent from 12 different devices, were received by the gateways located in the Netherlands. Such frames were found frequently in the Emmeloord and Utrecht regions. We will use these frames to calculate the coverage area in different landscapes.

4.4 Traffic Characteristics

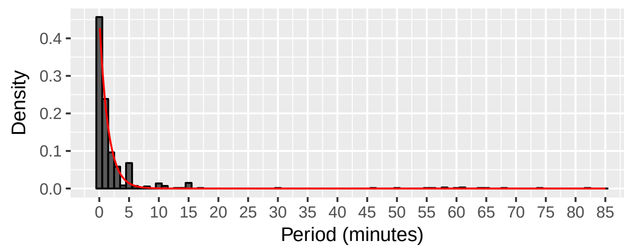
4.4.1 Periodicity. LoRaWAN devices might send packets using irregular or regular periods. We extracted the period for each device in the TTN dataset by calculating the time difference between two consecutive unique frames. The sequence of unique frames was based on the frame count (FCnt) found in the payload and was sorted in an ascending order. As there might be some packet loss in the dataset, the difference of FCnt could be larger than one, and thus we did not take into account such values.

The periods were grouped by device ID and the number of restarts as some devices might have restarted several times. By taking the average μ and standard deviation σ of those grouped periods, the coefficient of variation, $c_v = \frac{\sigma}{\mu}$, could be obtained to roughly determine periodicity and irregularity of the traffic. Figure 9(a) illustrates the histogram of c_v in which a value close to zero indicates stronger tendency towards periodic behaviour. Around 93.7% of the average periods have $c_v = 0.0$, 98.58% have $c_v < 1$, and only 1.49% have $c_v >= 1$. The results indicate that most traffic in the TTN network is periodic. From those transmissions, 46% have period less than 1 minute, 23% from 1 to 2 minutes, and the rest are spread up to 82 minutes. Only ten packets sent with period less than 1 s, which must have been in violation of the duty-cycle limitation. The histogram of different transmission periods is depicted in Figure 9(b), which can be fitted by an exponential model $n = a \cdot \exp(-b \cdot t)$, where $a = 0.46$ and $b = 0.69$.

4.4.2 Retransmission and Packet Loss. At least 5,302 retransmitted packets (0.36%) were obtained from the dataset of which 14.5% retransmitted every 5 s, and 9.92% every 8 s. We roughly estimate that around 40.4% of unique frames were received successfully and 60.6% were lost. This indicates that packet loss was relatively high in June 2016 for the TTN network deployment.



(a) Coefficient of variation.



(b) Transmission period.

Figure 9: Histogram of traffic periodicity for the TTN dataset.

4.4.3 Payload size. Payload size influences the time on air: a bigger payload requires a longer transmission time. Figure 10 depicts the histogram of payload size. It should be noted that the payload found in the TTN dataset represents application payload, not the PHY LoRa payload. Hence, when calculating the time on air using Equation 2, we need to add around 13 bytes of the LoRaWAN frame header to the *PL*.

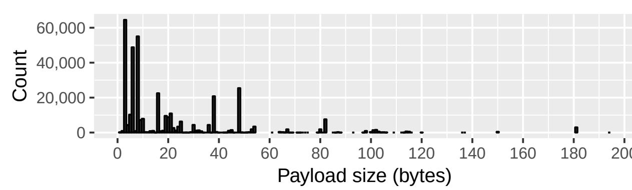


Figure 10: Histogram of payload size for the TTN dataset.

Up to 93.9% percent of the packets have payload less than 60 bytes of which 56.8% were less than or equal to 10 bytes. It indicates that the current usage of LoRaWAN network is either intended for applications that only require few bytes in each transmission, such as in environmental monitoring, or used by some users for testing the network.

4.4.4 Message Type. We observed up to 96.53% unique frames were unconfirmed uplink and 3.47% were confirmed uplink. It shows that the LoRaWAN network was mostly used by applications that did not require confirmations from the network server, which is on a par with the result of the payload analysis.

4.5 Signal Quality

Figure 11 reflects signal quality in terms of received signal strength indicator (RSSI) and SNR. We take $d_0 = 100$ and assume devices use transmission power (PWR) 14 dBm with typical antenna gain $G_{TX} = 2$ dBi from a pigtail antenna as most LoRaWAN devices are deployed with these default settings. We also assume antenna gains of gateways $G_{RX} = 2$ dBi.

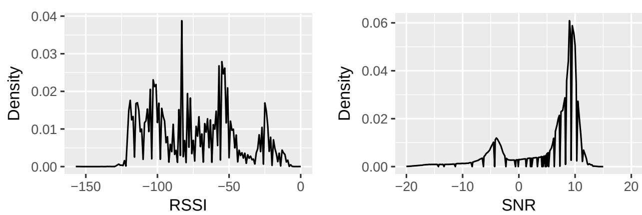


Figure 11: Received signal quality for the TTN dataset.

Most frames were received with positive SNR (> 10) and RSSI (> −100). The RSSI threshold relates to the receiver sensitivity, which, in this case, is limited to around -130 dBm. Higher RSSI and SNR values imply that the device was positioned close to the gateway. It is nevertheless not always the case, for example, if there are many interferers from other LoRa nodes or coexisting technologies, the RSSI may increase but the SNR value will drop drastically even if the node is located near the gateways. Knowing device locations will be useful for validating this result.

4.6 LoRa Parameters Utilization

The selection of parameters used for transmitting a packet can influence the probability of frame collisions, i.e. when multiple packets are sent with the same bandwidth, spreading factor, and carrier frequency. Figure 12 shows a histogram of LoRa parameter utilization from the TTN dataset.

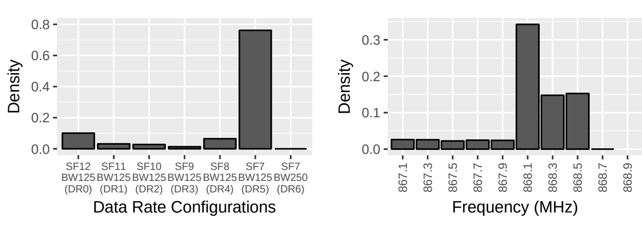


Figure 12: Histogram of LoRa parameter utilization for the TTN dataset.

Most packets were sent using DR5 (SF7BW125), which is likely to be the default setting of LoRaWAN devices. Some packets were also transmitted with the lowest data rate, i.e., DR0 (SF12BW125). Using a higher data rate decreases the time on air and reduces the probability of frame collisions, but when most devices use this setting, the frame collision probability remains high. It could be exacerbated when those packets are sent using the same frequency, which, in this dataset, were dominated by the three LoRaWAN default channels.

4.7 Coverage

The coverage of LoRaWAN device and gateway can contribute to the probability of frame collision due to the increased number of devices and, possibly, due to the interference between gateways. The first step is to calculate link budget:

$$P_{RX(dBm)} = P_{TX(dBm)} + G_{TX(dBi)} - L_{PL(dB)} + G_{RX(dBi)}\,, \quad (7)$$

where P_{RX} denotes the received power level at the receiver, P_{TX} denotes transmitter power, G_{TX} represents antenna gain of the transmitter, G_{RX} represents antenna gain of the receiver, and L_{PL} denotes path loss of the channel, which includes multipath and shadow fading. As LoRa signals can be decoded below the noise floor, we use the effective signal power (ESP) instead of RSSI. The received power P_{RX} then becomes:

$$P_{RX(dBm)} = RSSI_{(dBm)} + SNR_{(dB)} - 10\log_{10}\left(1 + 10^{0.1SNR_{(dB)}}\right) \quad (8)$$

Figure 13 shows a comparison between ESP, RSSI, and the sum of RSSI and SNR against different SNR values. For negative SNR, ESP fits with the summation of RSSI and SNR, but it differs for the positive SNR in which it becomes similar to RSSI. Since most frames were received with SNR ranging from -20 to 15, ESP is more appropriate to use.

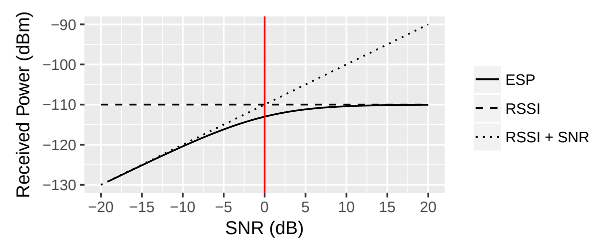
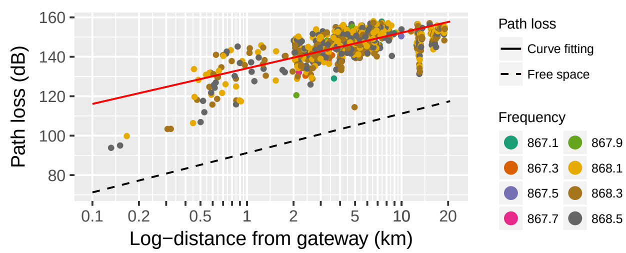


Figure 13: Comparison of RSSI and ESP over SNR.

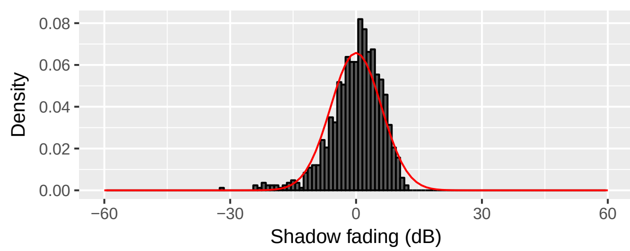
The next step is to calculate the path loss, which is mainly influenced by distance [5]. We use a log-distance path-loss model to estimate the path loss and its dependency on landscape type, such as urban and rural areas. The log-distance path-loss model is expressed as [3]:

$$L_{PL}(d)_{[dB]} = \overline{L_{PL}}(d_0)_{[dB]} + 10\gamma\log_{10}\left(\frac{d}{d_0}\right) + X_{\sigma[dB]}\,, \quad (9)$$

where d denotes the distance from the transmitter, d_0 denotes the reference distance (10-100 m for outdoors), $\overline{L_{PL}}(d_0)$ denotes the average path loss at reference distance d_0, γ denotes path loss exponent, and X_σ represents the log-normal shadow fading. We calculate the empirical path loss using Equation 7 and 8 with RSSI, SNR, and distance values taken from the TTN dataset. These values are then used for curve-fitting. Due to the limited number of frames having GPS location of the devices, we combine all frames and use the mean frequency for the curve-fitting as in [12]. Those frames were observed predominantly in the Emmeloord area. We assume all devices sent the frames with P_{TX} = 14 dBm and G_{TX} = 2 dBi, and the gateways have G_{RX} = 2 dBi. Figure 14 shows path loss curve-fitting and log-normal shadowing for the Emmeloord area.



(a) Path loss.



(b) Log-normal shadowing.

Figure 14: Path loss and log-normal shadowing distribution in Emmeloord (SF12BW125).

We also performed curve-fitting for the Utrecht urban area, but the number of frames was less than in the Emmloord area. The results can be seen in Table 5. The path loss exponent for the Emmloord area is slightly below the free-space path loss exponent (γ = 2) and has a lower variance, which is reasonable as the landscape is relatively open and flat, whilst the path loss exponent for the Utrecht area fits with a typical urban or suburban area and has a larger shadowing variance.

Table 4: Path loss exponents and shadow fading obtained from curve-fitting for Emmeloord and Utrecht regions.

	Emmeloord	Utrecht
Average frequency ($\overline{f_c}$)	868.2129 MHz	867.8026 MHz
Path loss exponent (γ)	1.806	2.6234
Mean path loss at d_0 ($\overline{L_{pl}}$)	116.0952	119.2509
Shadowing (σ_{SF})	6.064119	9.551812

Finally, the coverage area can be estimated by deriving the link budget formula:

$$R = d_0 \cdot 10^{\left(P_{TX}+G_{TX}+G_{RX}-S_{RX}-\overline{L_{PL}}(d_0),\right)/10\gamma} \tag{10}$$

where S_{RX} denotes sensitivity of the receiver, which depends on the spreading factor, bandwidth, and chipset type [18, 17]. It should be noted that we do not consider the shadowing effect in this formula so that the coverage area becomes a perfect circle. Table 6 shows the estimated coverage for the Emmeloord and Utrecht regions. The result shows that lowering the data rate, i.e. increasing SF value and reducing bandwidth, improves up to 40% coverage.

Table 5: Estimated coverage for different data rates and chipsets in the Emmeloord and Utrecht regions.

Data Rate	Distance (km)			
	Emmeloord		Utrecht	
	Gateway (SX1301)	Device (SX1276)	Gateway (SX1301)	Device (SX1276)
DR0	19.6	12.6	2.9	2.1
DR1	13.4	8.6	2.2	1.6
DR2	9.7	7.5	1.8	1.5
DR3	7.1	5.1	1.4	1.1
DR4	5.1	3.5	1.1	0.9
DR5	3.7	2.4	0.9	0.7
DR6	2.6	1.6	0.7	0.5

5 CONCLUSION

An extensive evaluation of LoRaWAN performance in small-scale and large-scale networks has been performed, in particular, on the characteristics of LoRaWAN frame collisions.

Spreading factor is one of the leading parameters that influence the time on air. Packets with SF7 and SF8 might suffer from packet loss due to the coverage outage, while SF12 might improve the coverage but increases the time on air and collision probability at several gateways, leading to packet loss. However, from the application layer standpoint, there will be no packet loss as long as one gateway receives the packet. It is recommended to use SF7 when devices are located close enough to a gateway, approximately below 700 meters in an urban area. SF9 should be sufficient for typical usage. Using higher coding rate also increases the time on air.

REFERENCES

[1] Bankov, D., Khorov, E., and Lyakhov, A. On the Limits of LoRaWAN Channel Access. In *2016 International Conference on Engineering and Telecommunication (EnT)* (11 2016), pp. 10–14.

[2] Blenn, N., and Kuipers, F. LoRaWAN in the Wild: Measurements from The Things Network.

[3] Bor, M., Roedig, U., Voigt, T., and Alonso, J. M. Do LoRa Low-Power Wide-Area Networks Scale? In *Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems* (2016), pp. 59–67.

[4] Georgiou, O., and Raza, U. Low Power Wide Area Network Analysis: Can LoRa Scale? *IEEE Wireless Communications Letters* (2017), 1–4.

[5] Goldsmith, A. *Wireless communications*. Cambridge university press, 2005.

[6] Goursad, C., and Gorce, J. M. Dedicated networks for IoT: PHY/MAC state of the art and challenges. *EAI Endorsed Transactions on the Internet of Things 1*, 1 (2015), 1–12.

[7] Haxhibeqiri, J., Van den Abeele, F., Hoebeke, J., and Ingrid, M. LoRa Scalability: A Simulation Model Based on Interference Measurements. *Sensors 17*, 6 (2017), 1193.

[8] LoRa Alliance. A technical overview of LoRa and LoRaWAN. Tech. Rep. November, 2015.

[9] LoRa Alliance. LoRaWAN Regional Parameters. Tech. rep., 2016.

[10] N. Sornin, M. Luis, T. Eirich, T. Kramp, and O.Hersent. LoRaWAN Specification v1.0.2. Tech. rep., 2015.

[11] Petäjäjärvi, J., Mikhaylov, K., Hamalainen, M., and Iinatti, J. Evaluation of LoRa LPWAN technology for remote health and wellbeing monitoring. In *International Symposium on Medical Information and Communication Technology, ISMICT* (2016), vol. 2016-June.
[12] Petäjäjärvi, J., Mikhaylov, K., Roivainen, A., Hänninen, T., and Pettissalo, M. On the coverage of LPWANs: Range evaluation and channel attenuation model for LoRa technology. In *2015 14th International Conference on ITS Telecommunications, ITST 2015* (2016), pp. 55–59.
[13] Pop, A.-I., Raza, U., Kulkarni, P., and Sooriyabandara, M. Does Bidirectional Traffic Do More Harm Than Good in LoRaWAN Based LPWA Networks?
[14] RevSpace. Decoding LoRa, 2016.
[15] Semtech. ETSI Compliance of the SX1272/3 LoRa Modem. Tech. Rep. July, 2013.
[16] Semtech. LoRaMote User Guide. Tech. Rep. July, 2014.
[17] Semtech. SX1276 datasheet. Tech. Rep. August, 2016.
[18] Semtech. SX1301 datasheet. Tech. Rep. May, 2017.
[19] Voigt, T., Bor, M., Roedig, U., and Alonso, J. Mitigating Inter-network Interference in LoRa Networks.