When LoRaWAN Frames Collide

Andri Rahmadhani  
Delft University of Technology  
Delft, The Netherlands  
andrewflash@gmail.com

Fernando Kuipers  
Delft University of Technology  
Delft, The Netherlands  
F.A.Kuipers@tudelft.nl

ABSTRACT
LoRa, an abbreviation of Long Range, is a Low-Power Wide Area Network (LPWAN) radio technology that has quickly gained popularity as a communications means for the Internet-of-Things (IoT). LoRa is typically used together with the MAC protocol LoRaWAN and operates in the license-free ISM-bands. As such, anyone is allowed to deploy their own LoRaWAN network, provided that they adhere to the LoRaWAN specification and ISM regulations. However, an uncoordinated deployment of LoRaWAN networks may cause neighboring networks to interfere and LoRaWAN frames to collide. In this paper, we present an in-depth investigation of LoRaWAN frame collisions – and the capture effect in particular – through various experiments. Contrary to previous research, we focus on correct reception of data at the application, instead of at the gateway, and we consider multi-gateway, multi-provider, and dense scenarios to obtain insight into collisions within actual networks.

CCS CONCEPTS
• Networks → Wireless access networks;

KEYWORDS
LoRaWAN; Capture Effect; IoT

1 INTRODUCTION
LoRa is a radio technology that enables long-range, low-power communication for Internet-of-Things (IoT) applications that only require low data rates. LoRaWAN is the de facto Low-Power Wide Area Network (LPWAN) MAC protocol for LoRa [8], and its specification regulates the use of LoRa technology in the unlicensed Industrial, Scientific, and Medical (ISM) frequency bands.

In the past couple of years, LoRa(WAN) has seen a steep adoption curve. The Things Network (TTN), an open community which crowd-sources the LoRaWAN network, has provided the back-end to connect thousands of gateways globally. Moreover, 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, the capacity of the network will quickly saturate, leading to performance degradation. This problem is aggravated by the fact that all gateways in the vicinity, regardless of which network provider they belong to, will receive packets transmitted by any LoRa device in range, causing inter-network interference.

In this paper, we present a characterization of LoRaWAN frame collision conditions. In particular:
• We perform extensive frame collision experiments using two nodes connected through a single gateway and then scale up the setup to use multiple gateways.
• We evaluate a dense scenario involving hundreds of nodes transmitting concurrently to a few gateways.
• We build a simulator to study the data extraction rate at the application.

2 BACKGROUND ON LORA AND LORAWAN
2.1 LoRa
LoRa is a Chirp Spread Spectrum (CSS) modulation scheme that encodes symbols into one or multiple signals of increasing (up-chirp) or decreasing (down-chirp) frequencies. A LoRa frame is initialized by a long constant-chirp preamble, which is used by the receiver to lock onto the LoRa signal. The preamble is followed by two down-chirps used as a synchronization word, indicating the end of the preamble. Since there is no disparity between the preambles from different
LoRa transmitters, the receiver may listen to unwanted signals. A LoRa transmission can be characterized by several key parameters:

- **Bandwidth (BW)** defines the width of the radio frequency being used.
- **Carrier Frequency (CF)** defines the medium frequency.
- **Coding Rate (CR)** defines the Forward Error Correction (FEC) rate used by the LoRa modem.
- **Spreading Factor (SF)** represents the ratio between symbol rate and chip rate. A higher spreading factor results in greater sensitivity and range, but it also increases power consumption and airtime. SF7 to SF12 are used in LoRaWAN. Transmissions on different spreading factors can be received concurrently at a gateway.
- **Transmission Power** defines the power used for transmission. LoRa chipsets support power from -4 dBm to 20 dBm.

The frame format used by LoRaWAN consists of a header, which includes information on the payload length in bytes, CR, and whether the 16-bit payload CRC is used or not. The preamble length can be configured from 6 to 65535 symbols. The LoRa modem then adds 4.25 symbols representing a synchronization word. The payload has a variable length ranging from 1 to 255 bytes.

### 2.2 LoRaWAN

LoRaWAN is a Media Access Control (MAC) protocol for controlling low-power devices in wide-area networks. Its first specification was released by the LoRa Alliance in 2015.

#### 2.2.1 Network Architecture

A LoRaWAN network is star-shaped and comprises:

- **End-device/node/mote**, a low-power module (sometimes equipped with sensors) that communicates with gateways using LoRa or FSK.
- **Gateway/base station**, an intermediate device that receives messages from end-devices and forwards them to the network server over an IP-based back-haul.
- **Network server**, a server that handles messages, e.g., duplicate message removal, decoding, and link configuration adjustment of end-devices.

#### 2.2.2 Classes

Three different classes (A, B, and C) are introduced to facilitate various applications. In this paper, we focus only on class A devices as that class is mandatory for every LoRaWAN device. In class A, an end-device transmits a message when it wants, as long as it obeys the regulations. After each uplink transmission two receive windows are opened, namely RX1 and RX2. As a result, a downlink message at any other time needs to wait for the next uplink message.

#### 2.2.3 Frequency Regulation and Access Policy

In Europe (EU), where our measurements were performed, 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 LoRaWAN devices operate in that frequency band. The default transmission power of devices for EU863-870 is 14 dBm and is limited to this value, except for the 869.4 - 869.65 sub-band (up to 27 dBm), which is typically used for downlink traffic.

The European frequency regulation (ETSI) imposes the use of either duty cycle or Listen Before Talk in combination with Adaptive Frequency Agility (LBT-AFA). Duty cycle, the option used by LoRaWAN, represents the percentage of time an end-device occupies a particular channel.

#### 2.2.4 Adaptive Data Rate

Adaptive Data Rate (ADR) is a technique used for adjusting the device’s data rate to control performance of the network. For instance, devices close to a gateway will use a higher data rate, leading to a shorter time on air, and lower output power.

ADR can be initiated either by the network or by 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 (not part of the specification).

### 3 COLLISION CHARACTERIZATION

In this section, we investigate under which conditions two colliding frames are both lost.

#### 3.1 Small-scale experiment

##### 3.1.1 Single Gateway

We start with small-scale collision experiments using two nodes and a single-channel gateway located 5 m from the nodes. A PC acts as the network server and receives the gateway logs via UDP. The gateway status updates are filtered out, keeping only the received frames. Afterwards, each frame is decoded to obtain the device address. If the address is not valid, i.e., it does not belong to either of the nodes, the frame is labeled as "unrecognized". If the address is valid, the Message Integrity Check (MIC) – a cryptographic integrity check on the MAC header and payload data, e.g., see [10] – of the frame is verified according to the device address. A frame with an invalid MIC is regarded as a corrupt frame. Such frame will still be forwarded to the network server when its payload CRC is valid. 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 (SF8BW125), similar to the experiment done in [3]. In the second scenario, node 1 (N1)
and node 2 (N2) used \( P_{TX} = 2 \) dBm and \( P_{TX} = 8 \) dBm, respectively. For both scenarios, we used a fixed 26-bytes application payload that was sent periodically (tailored to the duty cycle). The stronger node, N2, 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 at the right time. The transmissions used frequency 869.7 MHz, which was relatively free compared to the default LoRaWAN channels, and were done for all SFs. The gateway, which was built using ESP8266 and an SX1276 LoRa module, sent the received frames to the PC via WLAN.

The results for the first scenario are difficult to interpret as the gateway kept switching its state in reading the signal from N1 or N2. However, we have observed that the SNR values (not displayed here) tend toward zero when the frames largely overlap, which indicates that an unwanted LoRa signal having the same SF, frequency, and power may lead to destructive interference.

For the second scenario, we obtained similar Packet Delivery Ratio (PDR) – the fraction of frames with correct MIC over the total number of frames sent – patterns for all SFs, so we only display the results for SF11 in Figure 1. We also added timing information obtained by calculating the time on air using the equations from Semtech (e.g., see [2]). We measured the RSSI levels for both nodes and observed that the average RSSI difference was approximately 15 dB, which is far above the 6 dB threshold that reportedly is needed to be able to decode a stronger frame correctly [5].

**Figure 1: Packet delivery ratio (PDR) for the 2nd scenario.**

From our experiments with the second scenario, we infer several cases at which frame loss may occur. Figure 2 illustrates these cases:

- **Case 1:** Both frames are considered lost when the stronger frame arrives later than the receiver locking time, i.e. 4 symbols. In that case, the receiver attempts to listen to the weaker signal, but it is suppressed by the stronger signal.
- **Case 2:** The stronger frame survives the collision when its arrival overlaps with the header CRC of the weaker frame. This will cause the receiver to release the lock on the weaker frame and start listening to the new stronger frame.
- **Case 3:** Both frames get corrupted 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 onto the weak frame, whilst the data inside the LoRaWAN header, e.g. device address, becomes corrupt.
- **Case 4:** Both frames are dropped 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 case, the weaker frame is successfully received, but the payload gets destroyed, resulting in a wrong MIC. Thanks to error correction techniques employed in LoRa transceivers, the weaker frame might still be decoded whenever the stronger frame only slightly overlaps with the payload CRC of the weaker frame.

Contrary to previous results by [3] and [6], our findings deliver new insight into frame loss characteristics from the application-level perspective, that is, including LoRaWAN header corruption and MIC failure.

### 3.1.2 Multiple Gateways

We have extended our experiment to observe the impact of adding more gateways, which has not been investigated before: Both devices were registered to The Things Network (TTN) using Activation By Personalization (ABP) without ADR enabled, and they were configured to transmit with the same data rate (SF9BW125), coding rate (4/5), frequency (868.1 MHz), and payload size (17 bytes), but with different transmission powers (8 dBm and 14 dBm). The weaker node (N1) was again used as the reference for calculating the time offset. Both devices sent 10 frames per time offset. Those frames were received by six gateways at different distances given in Figure 3, which shows the Data Extraction Ratio (DER), i.e. the fraction of packets correctly received at the application layer. PDR and DER are similar: the only difference is that the PDR is obtained by checking the MIC at the network server, while the DER is calculated from the application (which receives packets that passed the MIC check by the network server). The number on top of the figure represents the frame timing information: (1) preamble,
Figure 2: Conditions at which frame loss may occur.

(a) Case 1
(b) Case 2
(c) Case 3
(d) Case 4

Figure 3: DER for the collision experiment with multiple gateways. The time offset is with respect to the weak node.

Our results reveal an improved DER w.r.t. the single gateway scenario, especially for the delayed stronger frames. As the weaker frames could not reach the more distant gateways (GW4, GW5, and GW6), the stronger frames could be decoded properly. The closer gateways (GW1, GW2, and GW3) also contributed to the higher DER levels, as having multiple gateways increases the probability of receiving correct frames.

We proceeded to connect the devices to different networks. One end-device was registered to TTN, while the other one was registered to KPN (a Dutch telco). Both devices had ADR disabled. The KPN device was used as an interferer. Both nodes used the same data rate (SF12BW125), coding rate
(4/5), frequency (868.1 MHz), and payload size (17 bytes), but with different transmission powers (either 14 dBm or 8 dBm) corresponding to two scenarios: (1) weak interferer, and (2) strong interferer.

The time offset was set using time steps of one symbol time and with respect to the TTN node. Both end-devices sent 7 frames per time offset. Those frames were received by the same gateways as in the previous experiment.

It turned out that the KPN device received two new frequencies (either 14 dBm or 8 dBm) through ADR, i.e., 867.7 and 867.9 MHz, even though the device had not called for it, whilst the TTN device remained to use the 868.1 MHz frequency channel. The KPN device used those three frequencies equally, i.e., 33.3% for each channel, meaning that the probability of collision reduced to 33.3%.

Figure 4 depicts the DER for both scenarios. In the weak interferer scenario, most TTN packets were delivered to the application layer even when the interferer arrived first, which is indicated by a negative time offset. This result is due to the interferer not being able to reach the more distant gateways and because of the differences in allocating the frequency channels.

In the strong interferer scenario, the TTN frames hardly reached the farthest gateway (GW6), due to the lower transmit power and the interference caused by the strong interferer. Surprisingly, no frames were received by GW2 for positive time offset, which suggests that the gateway was momentarily inactive (as the other gateways could still receive the frames).

3.2 High utilization experiment

We have analyzed data from an experiment conducted by Thomas Telkamp during the Electronics & Applications (E&A) event at the Jaarbeurs in Utrecht, the Netherlands, from May 30 until June 1st, 2017.

3.2.1 Experiment Setup. During the E&A event, attendees were handed a LoRaWAN device (KISS Lora). All devices were initially programmed to have the same device address. Each "anonymous" device then switched to a personal address after personalizing it at designated booths. The devices were programmed to send a fixed 23-bytes packet approximately every 20 seconds, in which each transmission used a random spreading factor with probability 1/2 for SF7, 1/4 for SF8, 1/8 for SF9, 1/16 for SF10, 1/32 for SF11, and 1/64 for SF12. The frequency used for transmission was distributed uniformly over 8 frequency channels. Traffic from other devices was also observed during the measurements.

We discovered 8 gateways around the Jaarbeurs area, but only 5 gateways had GPS. From those 5 gateways, 3 (GW1, GW2, GW3) were placed inside the Jaarbeurs as part of the experiment and the rest (GW4, GW5) were located outside, at distance 4.24 km (GW4) and 9 km (GW5). As the KISS LoRa devices were not equipped with a GPS module, we could only guess their position as a reference point to calculate their distance to the gateways.

3.2.2 Number of End-devices. The dataset was filtered to obtain data from KISS LoRa devices only. We found 1426 devices (hardware EUIs) of which 489 anonymous devices were not personalized, 845 anonymous devices were switched to personal devices, and 92 devices were personalized outside the event, e.g., at home.

Not all end-devices sent the same number of packets. Figure 5(a) shows the histogram of end-devices that sent $n$ packets. It shows that up to 158 devices only sent one packet and only one device sent 1940 packets in the specified time-frame. Assuming each device can send exactly one packet per 20 seconds, a device could have sent 3240 packets at most. This
indicates that those devices did not activate all the time and might have restarted.

Partitions of a 1-minute intervals were created and the number of distinct end-devices and total packets sent per partition were counted to observe the relationship between the number of end-devices and the total packets sent. The result is depicted in Figure 5(b). The data is fit with a linear model with a gradient of 2.3, meaning that, on average, an end-device sent a packet approximately every 26 seconds, which is close to the initial setup of 20 seconds.

$$y = 2.3x, \quad r^2 = 0.998$$

![Figure 5: Traffic characteristics: (a) histogram of packets sent vs. number of devices, (b) number of devices vs. estimated number of packets sent per minute.](image)

3.2.3 *LoRa parameters.* Figure 6 depicts the distribution of SF and CF based on the traffic from the KISS LoRa devices. The results show that the distribution of SF and CF match with the experiment setup, particularly for the gateways placed inside the Jaarbeurs (GW1, GW2, and GW3). For low SFs, some packets were unable to reach GW4 and GW5.

3.2.4 *Frame Error Rate.* From the gateway status updates, the number of frames lost due to a CRC error can also be extracted, as shown in Figure 7. The frame error rate is proportional to the number of received frames per minute and is approximately 5% for the gateways placed inside the Jaarbeurs, and up to 32% for the gateways placed outside the Jaarbeurs. This indicates that most frames hardly reached the more distant gateways (GW4 and GW5), which is possibly due to the low spreading factor used most of the time. Collisions can also aggravate the situation, especially for the frames that use high spreading factors and required longer time on air.

![Figure 6: Histogram of SF and CF of KISS LoRa devices obtained from the received frames per gateway.](image)

3.2.5 *Data Extraction Rate.* Based on the number of devices and transmitted packets within a measurement period from
the experiments, the average DER can be calculated. Assuming a linear relationship between nodes and generated frames (see Figure 5(b)), we can plot DER against the number of active devices per minute. The DER plots are depicted in Figure 8.

![Figure 8: Average DER for single and multiple gateways obtained from the dataset.](image)

The results show that for a single gateway, DER mostly depends on the distance between the devices and gateways. Even for the closer gateways, some variations in DER can still be observed. But in general, DER declines gradually as the number of devices increases. Adding more gateways can improve DER around 1% to 5% for approximately 140 active devices per minute.

We have developed a simulator that employs the log-distance path loss model used in [3] for simulating transmission propagation, as well as the additive interference model (AIM) [7] and capture threshold model (CTM) [4].

The path loss exponent ($\gamma = 2.6234$) and shadowing ($\sigma = 7.3871$) for the log-distance path loss model are obtained by analyzing the data from [2] for the region around Utrecht. As three gateways are placed indoor, the path loss exponent and shadowing variance are different. In this case, we use $\gamma = 3$ and $\sigma = 9$ for the distance below 100 m. We compare the CTM and AIM interference models with a conservative approach “Full” in which frames are marked lost when they occupy the same spreading factor, bandwidth, and frequency, and overlap in time even when the overlapping part is infinitesimally small.

The simulation assumes no interference between spreading factors and no packets surviving a collision when the stronger frame destroys the LoRaWAN header of the first-arrived weaker frame. In order to do so, one should track the collision time, which becomes complex when dealing with multiple collisions. In our case, a frame is successfully received as long as any interfering frames do not overlap with the current frame preamble. Algorithm 1 describes our simulation procedure.

**Algorithm 1 E&A Event Simulation Procedure**

1. Define a vector $N = \{n_1, n_2, \ldots, n_l\}$ that represent a sequence of number of active devices $n_i$.
2. Define the duration of measurement $T$.
3. Define a fixed transmission period $t$ per device.
4. Generate a vector of packets $P_k$ per device with a total of $(n_i \cdot t)$ packets.
5. Assign spreading factor and frequency channel for packet $P_k$, following the real experiment distribution.
6. Assign a random transmission start time $t_{start}$ to packet $P_k$.
7. Clone packet $P_k$ for $G$ number of gateways.
8. Calculate the distance to gateway $l$ and propagation loss for packet $P_k$ per gateway.
9. Check received power for packet $P_k$ at gateway $l$ and mark it as lost if the received power is below the receiver sensitivity.
10. Check for collision of packet $P_k$ with the prior and subsequent packets at gateway $l$ by observing the start time $t_{start}$, time on air $t_{onair}$, and the end time $t_{end}$ of the packets.
11. Calculate DER for gateway $l$ and the aggregated gateways.

The measurement duration is set to 1 hour for each number of devices $n_i$, ranging from 1 to 150 devices. The results are depicted in Figure 9. For the single gateway case, the trends are relatively similar to those of the experiment in Figure 8. When multiple gateways are deployed, the conservative method does not show any improvement whatsoever as it, eventually, depends solely on the arrival and end-of-reception time, which can be different for each gateway if the propagation time is added to the simulation. The most notable improvement appears in the CTM model, providing up to 0.5% DER improvement when adding more gateways. Comparing these results with the real experiment, we believe
Several frames (especially those received by GW1) survived a collision even when there was overlap after preamble detection. We therefore investigated the power difference in overlapping packets in terms of Estimated Signal Power (ESP) to identify whether such frames had at least 6 dB received power difference. The results are depicted in Figure 11. We only considered frames that overlapped more than 25% of the frame duration, and no frames received by GW4 and GW5 fulfilled this condition. It can be seen that the ESP differences for GW1 are shifted to the right so that more packets exceeded the 6 dB difference. However, even for power differences less than 6 dB, sometimes a frame could survive a collision. The exact cause is difficult to analyze since the accuracy of the signal strength measurements and timing at the gateways could not be verified.

4 RELATED WORK

Bor et al. [3] have developed LoRaSim to evaluate LoRaWAN scalability. Their results showed that LoRaWAN deployments do not scale well. Bankov et al. [1] studied the capacity limit of a LoRaWAN network using simulation, which utilized a Poisson process to generate messages and the Okumura-Hata model to calculate path loss. Their results show that a single LoRaWAN gateway can easily become congested.

Voigt et al. [9] have investigated the use of directional antennas and the addition of new LoRa base stations on the reduction of packet loss due to LoRa inter-network interference. LoRaSim was used for simulating the efficiency of the proposed solutions in terms of DER. The simulation was performed using the same settings for all nodes. No ADR, no channel hopping, and only uplink messages were considered. However, the capture effect was included in LoRaSim. They concluded that deploying multiple base stations gives better results than employing directional antennae on end-devices.

Blenn and Kuipers [2] have extensively measured and evaluated a real-world, large-scale LoRaWAN network (TTN) within a time-frame of 8 months. The evaluation covers frame payloads, signal quality, and spatio-temporal aspects to estimate the performance of LoRaWAN. Simulations were used to evaluate packet loss, as well as the effect of confirmed and downlink packets. The results show that not all channels provided by the network are used evenly, which leads to more packet loss.

Hashibeqiri et al. [6] have evaluated the scalability of a single-cell LoRaWAN network in terms of the number of nodes that can be served by using simulation based on real measurements. The experiment was conducted to observe the collision of two LoRa transmissions in an isolated lab setup. The results show that a frame can survive a collision whenever the last six symbols of the preamble and header do not collide.
5 CONCLUSION

In this paper, we have provided an in-depth investigation of LoRaWAN frame collisions. We have evaluated Packet Delivery Ratio (PDR) during collisions. The whole packet duration has been partitioned based on the LoRaWAN frame structure to investigate under which circumstances collisions lead to frame loss. We have classified the corrupted frames based on the device IDs and investigated the Message Integrity Check (MIC).

Furthermore, we have performed frame collision measurements in a multi-gateway and multi-provider environment and have studied the Data Extraction Rate (DER) through practical measurements and simulations.

ACKNOWLEDGEMENT

We thank Thomas Telkamp for having shared the E&A data with us.

REFERENCES