LoRa-where? Analysing signal pick-up from LoRaWAN sensors in the wild

XIANZHI WU, University of Twente, The Netherlands

LoRaWAN, as a wireless communication technology, has become increasingly pervasive. Where is LoRa? LoRa is in your study place, in the parking lot, and in your office. LoRa signals can propagate over a long range and be picked up by gateways within that range. The University of Twente has used LoRaWAN sensors all over campus to monitor parking spots, room occupancy, air quality, and more. By performing a longitudinal analysis on the data that was collected by LISA over a span of 4 years, we aim to explore the impact of the spreading factor on the signal propagation in the LoRaWAN network, as well as the impact of gateway location on the signal pick-up.

Additional Key Words and Phrases: LoRaWAN, signal propagation, data analysis, UT campus.

1 Introduction

The LoRa (**Lo**ng **Ra**ng) technology has gained popularity in the Internet of Things (IoT) application due to its low-power consumption, long-range transmission, and ease of deployment [1]. LoRa is a physical access layer specification employing Chirp Spread Spectrum (CSS) technology. Transmission modulated by LoRa is robust against channel noise and can travel longer distances compared to other technologies like WiFi and Bluetooth. LoRa is suitable for sensors that operate in low power mode and transmit small size data, such as temperature and humidity information, at low bit rates [2].

LoRaWAN is a Media Access Control (MAC) layer protocol built on top of the LoRa physical layer that enables communication between sensors and gateways [2]. Therefore, LoRaWAN is a wireless, low-power and wide-area technology specifically developed for use in battery-powered sensor-type applications. Transmitting in the unlicensed spectrum, LoRa signals can reach long distances of up to 15 kilometers or more.

The University of Twente (UT) has set up a wireless network specially for smart devices and IoT. Figure 1 illustrates the distribution of the known LoRaWAN sensors and gateways deployed on the UT campus. In this figure, the symbol 'S' represents the LoRaWAN sensor, while 'GW' represents the LoRaWAN gateway. It is worth noting that all the LoRaWAN sensors are indoor. The building Ravelijn has the highest number of installed sensors, a total of 144, with one indoor gateway as well as one on the rooftop. Following closely, the library building Vrijhof hosts the 52 sensors and one indoor gateway. Additionally, the building Spiegel has 3 sensors along with two gateways. One gateway is installed in the office for testing purposes, while the other is positioned on the rooftop. Lastly, the building Horst Complex accommodates 5 sensors.

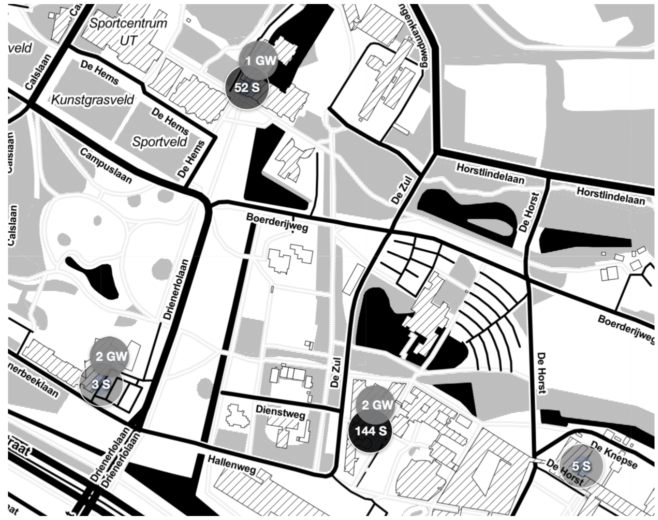

Fig. 1. Distribution of UT's known LoRaWAN sensors and gateways on Map. S: sensor. GW: gateway. Leaflet | Map tiles by Stamen Design, CC BY 3.0 — Map data © OpenStreetMap contributors

LoRaWAN, as compared to WiFi, Bluetooth and Cellular, is a relatively new technology, thus making it a worthy subject for research. In this research, our objective is to examine the performance of LoRaWAN on the UT campus. By conducting a longitudinal analysis of the LoRaWAN data, we aim to identify what factors can influence the signal propagation in the LoRaWAN network. Through this exploration, we seek to gain valuable insights into optimizing the LoRaWAN network performance.

LoRaWAN signal propagation and reception is influenced by factors such as the spreading factor, the distance between sensor and gateway, and buildings. We want to study how these factors affect performance on the UT campus. In order to do so, we formulate the following main research question.

What factors can affect the patterns in signal propagation in the LoRaWAN network and how can they be optimized to improve the network performance?

We break this main question into the following questions.

(1) How does signal propagation differ when comparing signals picked up by indoor and outdoor gateways?

(2) How can the configuration parameters of a LoRaWAN sensor such as spreading factor affect the patterns in signal propagation?

2 Related works

Previous studies have been conducted to identify patterns in LoRa signal propagation. The study [5] found that the stability of LoRa

TScIT 39, July 7, 2023, Enschede, The Netherlands

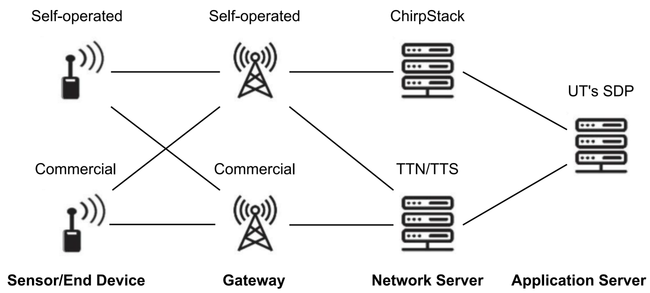

Fig. 2. UT's LoRaWAN architecture

signals heavily depends on the surrounding environment and it is more stable in suburban areas compared to high-density urban areas. A study employing loRadar as a research tool has been conducted to provide an valuable insight that a larger SF can make a signal transmit longer and less sensitive to noise such that it can achieve longer communication ranges [3]. Additionally, an experimental study [7] shows that with SF7, RSSI tends to be higher when compared to SF12. The study of signal propagation in LoRa is relatively limited, with only a few existing works. Given the scarcity of research in this area, we are going to investigate signal propagation in LoRa specifically within the UT campus environment.

Our research focuses on the impact of the spreading factor and the gateway's indoor or outdoor location on signal propagation on the UT campus. Before performing the research, it's essential to describe how the data is collected.

3 Data collection

The UT's LoRaWAN architecture is shown as Figure 2. LISA makes use of a combination of self-operated sensors and gateways, as well as commercial ones. The gateways receive data from the sensor and forward it to the network server providers ChirpStack and The Things Network (TTN) or The Things Stack (TTS) which is TTN v3. After receiving and processing the data, the network server providers send it to the sensor data platform of UT. The data collected by LISA not only records the data the sensor sent, but also records which gateways picked up the data signal. The data is stored in JSON format but with multiple different structures and contains information such as the time the data was received from a sensor, the spreading factor, the bandwidth, the gateway that picked up the signal and the signal strength, etc. The data used for this research spans from May 15th, 2019 to May 12th, 2023.

4 Methodology

We perform a longitudinal analysis on the data that was collected from the same group of sensors on the UT campus at different points of time after preprocessing the raw data.

4.1 Preprocess the data

Relevant information for this research is extracted from the raw data, including the measurement ID, the time the message was received by the gateway, the network server and the sensor data platform, the sensor's EUI, spreading factor, bandwidth, frequency, coding rate, the gateway's EUI, RSSI, channel, SNR and frame counter.

4.1.1 Data conversion Messages that were received by different network servers are stored in deeply-nested JSON data format with different structures. Messages are read one by one from the data set and then transformed into a standardized format. The messages information is stored in the tabular data format CSV. If a message was received by multiple gateways, it will be stored in multiple lines in the CSV file, with each line representing one copy of the message received by one gateway. Table 1 displays information related to sensors, while Table 2 presents information pertaining to gateways. The information is extracted from messages received by different network servers, each utilizing its own key name. For instance, the device EUI is extracted from the data packet using one of three key variations: 'devEUI', 'hardware_serial' or 'dev_eui', depending on the network server. Here are some additional explanations regarding the tables:

- 'NS' represents the network server that received the messages. It can be ChirpStack, TTN or TTS.
- 'data_rate' in the message received by TTN represents the data rate. For instance, 'SF7BW125' means that the spreading factor is 7 and the bandwidth is 125KHz.
- If the sensor EUI is base64 encoded, it will be decoded and then converted into a 64-bit hex string.
- In the data packet of TTS, gateway EUI can be extracted from 'eui' that is nested in 'gateway_ids' or 'forwarder_gateway_eui' that is nested in 'packet_broker'. It depends on whether the message was relayed using Packet Broker or not. If the EUI of the gateway is base64 encoded, it will be decoded and then converted into a 64-bit hex string. For instance, the base64 encoded gateway eui 'He4Dmqx1wwc=' can be converted into a 64-bit hex string '1DEE039AAC75C307'. Furthermore, this gateway was given the descriptive name of 'utwente-enschede-macrocell', if a gateway EUI is extracted as 'utwente-enschede-macrocell', it will be converted into '1DEE039AAC75C307'.
- In the data packet of TTN, one key 'time' is nested in 'gateways'. It is the time when the gateway received the data packet.
- 'Time on NS' in the header of Table 2 represents the time the message was received by the network server after being relayed by the gateway.

4.1.2 Data cleaning Check if there are any incorrect data values and remove them from the CSV file to make the data suitable for analysis. To accomplish this, the following steps are carried out:

(1) Remove messages that contain empty values for measurement ID, the time they were received by the gateway, the network server and the sensor data platform, the sensor's EUI, spreading factor, bandwidth, frequency, coding rate, the gateway's EUI, RSSI, channel, SNR or frame counter. Additionally, eliminate messages with negative RSSI values.
(2) Frame counter is used to prevent the end device from processing the same message twice, as well as being exploited by replay attacks [6]. Messages that have the same values of the time they were received by the gateway, the sensor's EUI, spreading factor, bandwidth, frequency, coding rate, the gateway's EUI,

NS	Sensor's EUI	Spreading Factor	Bandwidth	Frequency	Coding Rate	Frame Counter
ChirpStack	devEUI	spreadingFactor	bandwidth	frequency	codeRate	fCnt
TTN	hardware_serial	data_rate		frequency	coding_rate	counter
TTS	devEUI	spreadingFactor	bandwidth	frequency	codeRate	f_cnt

Table 1. Information related to sensor

NS	Gateway's EUI	RSSI	Channel	SNR	Time	Time on NS
ChirpStack	gatewayID	rssi	channel	loRaSNR	time	publishedAt
TTN	gtw_id	rssi	channel	snr	gateways.time	time
TTS	gateway_ids.eui packet_broker.forwarder_gateway_eui	rssi	channel_index	snr	time	received_at

Table 2. Information related to gateway

Gateway	Location	SF7	SF8	SF9	SF10	SF11	SF12	Total
UTwente Kerlink	Spiegel rooftop	4,891,098	413,646	2,878,378	52,413	15,102	3,046,054	11,296,691
UTwente Ravelijn rooftop	Ravelijn rooftop	5,242,024	338,484	1,707,334	50,960	10,483	2,969,500	10,318,785
UT Ravelijn indoor	Ravelijn indoor	3,090,046	4,020	2,734	2,843	1,482	1,782,099	4,883,224
dragino-ub	Vrijhof indoor	289,983	3,92	668,838	-	-	1,155,427	2,114,640
utwente-lg308-03	Spiegel indoor	5,743	78	142	178	199	36,619	42,959

Table 3. Number of messages received by each gateway with different spreading factors.

SF7	SF12	SF9	SF8	SF10	SF11
13,518,894	8,989,699	5,257,426	756,620	106,394	27,266

Table 4. Number of messages received by gateways with different spreading factors.

Gateway Count	1	2	3	4	5
Message Count	4,887,735	6,122,695	3,809,443	23,705	5

Table 5. Number of messages received simultaneously by the respective number of gateways.

868.1	868.3	868.5	867.1	867.3	867.5	867.7	867.9
0	1	2	3	4	5	6	7

Table 6. Radio frequencies and channels that sensors use on the UT campus.

RSSI, channel, SNR and frame counter can be seen as duplicated or invalid. Consequently, only one instance of such duplicated messages will be retained. There are 1,074,907 copies of such duplicated messages.

(3) Messages that contain duplicate gateway EUIs introduce uncertainty regarding their accuracy. To ensure data reliability, these messages are removed. There are 186,064 copies of such duplicated messages.

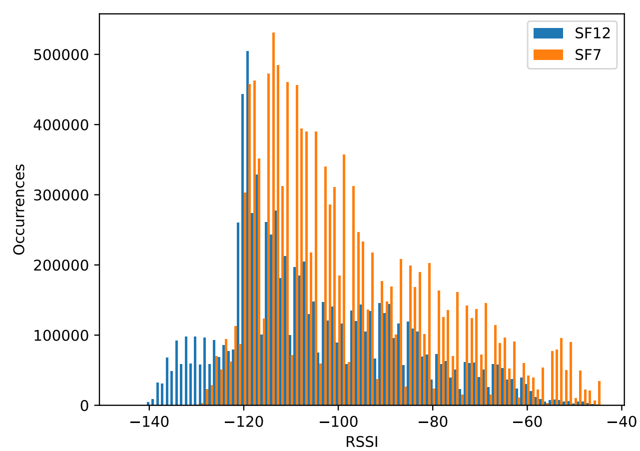

Fig. 3. Distribution of messages received by gateways from sensors using SF7 and SF12, categorized by RSSI.

4.2 Dataset demographics

In this research, we focus on analyzing the data received by the top 5 gateways that have received the highest number of messages. These gateways are installed on the UT campus as shown in Figure 1. The remaining gateways, which are located outside the UT campus, are not included in our analysis.

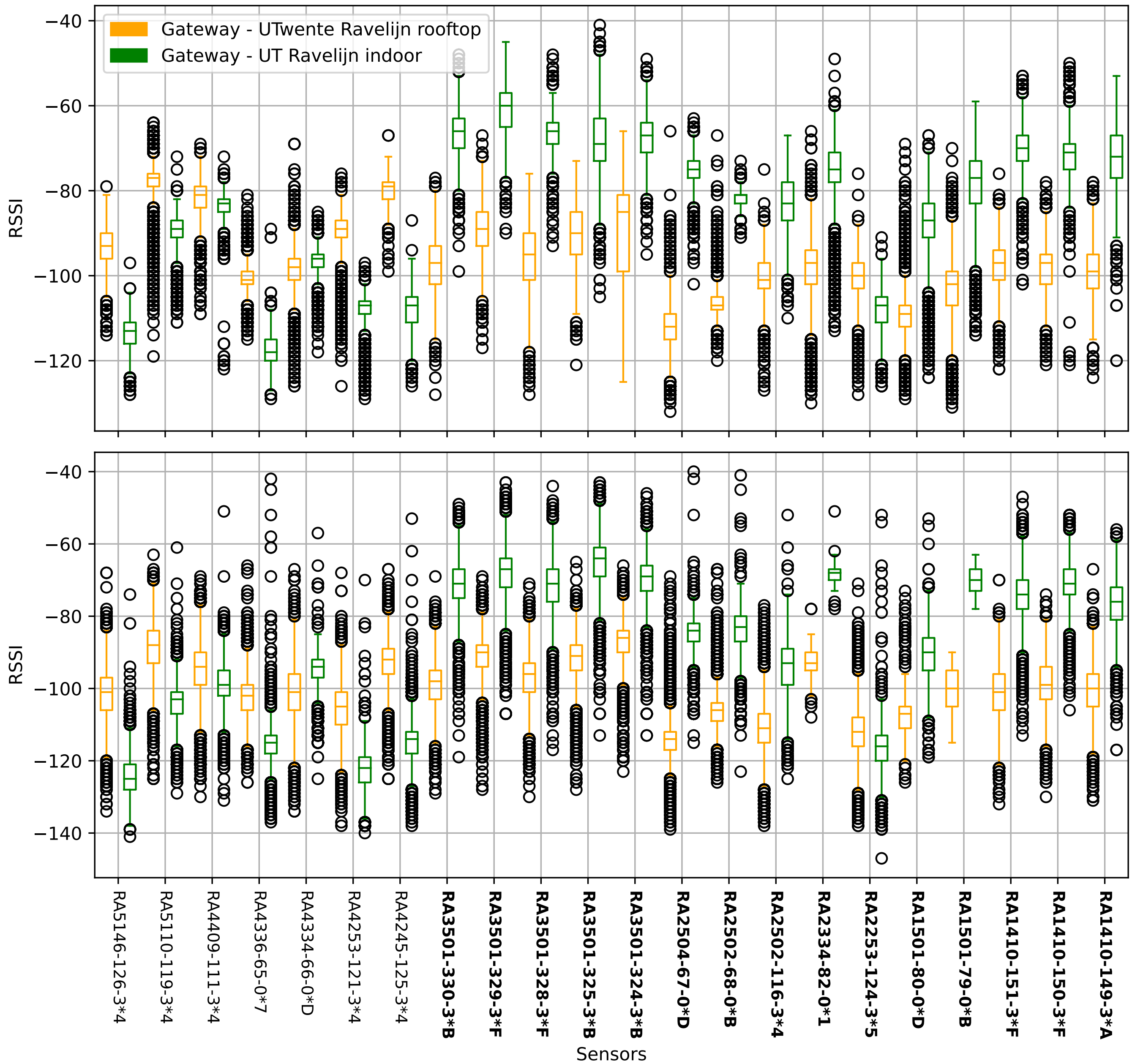

Fig. 4. Distributions of RSSI values of messages received by gateways from sensors using SF7 (top) and SF12 (bottom), categorized by gateway.

A total of 28,656,299 copies of valid messages are extracted from the datasets for this research. All these messages were sent by sensors using spreading factors (SF) ranging from 7 to 12 with a fixed bandwidth of 125kHz and a coding rate of 4/5, and then were successfully demodulated by gateways. Interestingly, they are not evenly distributed among the 6 spreading factors. The distribution, presented in a descending order, is shown as Table 4. To provide additional details, Table 3 gives a comprehensive overview of the number of messages received by each gateway, categorized by different spreading factors. Additionally, Table 5 shows the number of messages received by the respective number of gateways at the same time. Lastly, it's worth noticing that the radio frequencies that sensors use on the UT campus are: 868.1 MHz, 868.3 MHz, 868.5 MHz, 867.1 MHz, 867.3 MHz, 867.5 MHz, 867.7 MHz, and 867.9 MHz, which are summarized in Table 6.

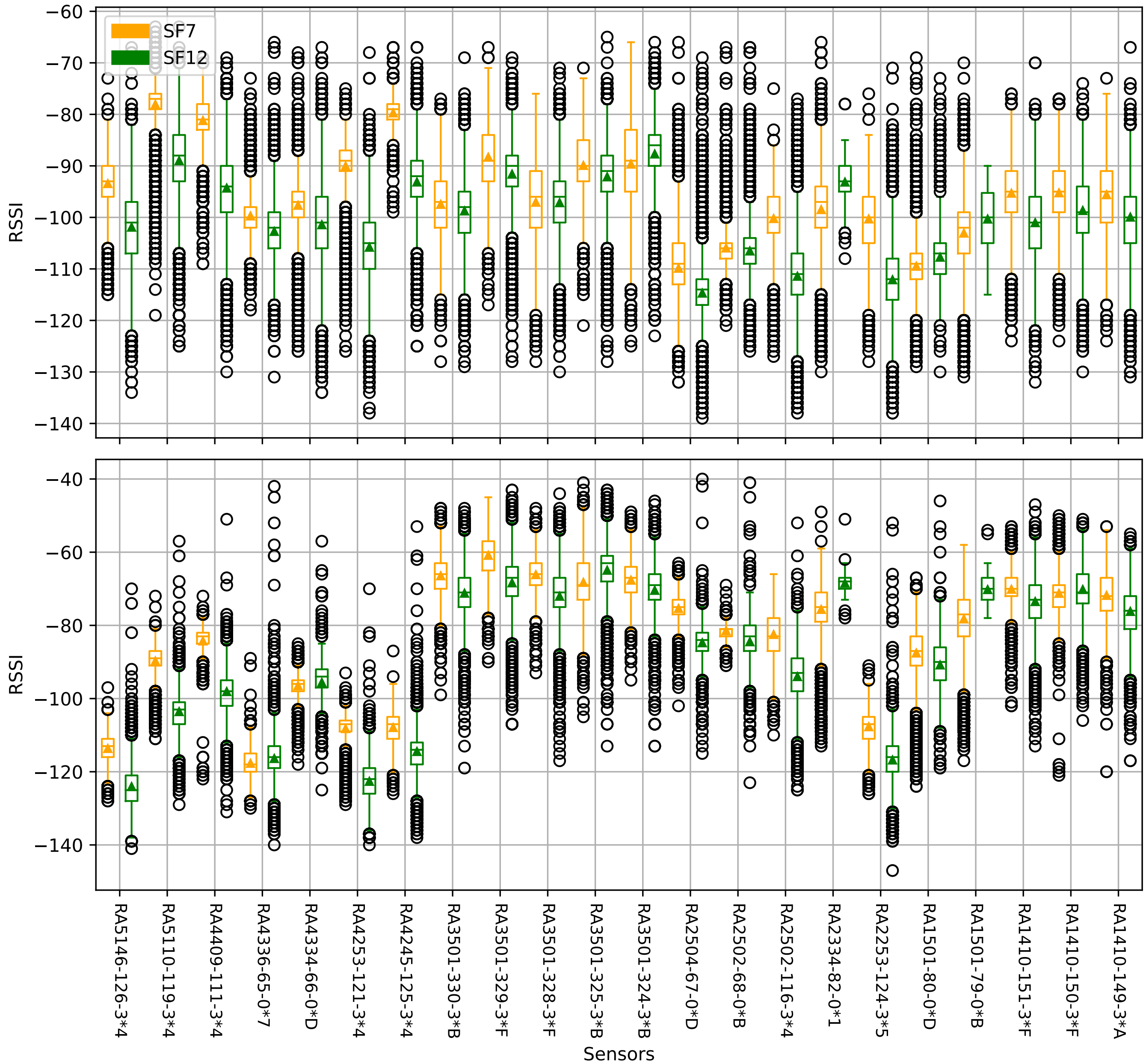

Fig. 5. Distributions of RSSI values of messages received by gateway 'UTwente Ravelijn rooftop' (top) and gateway 'UT Ravelijn indoor' (bottom) from sensors using SF7 and SF12, categorized by SF.

4.3 Visualize and analyse the data

To gain insights into the impact of SF on signal propagation, the first step is to plot the number of messages sent by sensors using SF7 and SF12 and demodulated successfully by gateways with Received Signal Strength Indicator (RSSI) values. As Figure 3 shows, the distributions of both SF7 and SF12 exhibit a similar data pattern, having a long right-skewed tail. At first glance, SF7 appears to provide a higher signal strength compared to SF12. However, the distribution of SF12 displays a clear difference on the left side, with messages exhibiting lower RSSI values. This discrepancy suggests that higher spreading factors could be more easily demodulated by gateways [4].

To go further, we do analysis on the messages received by two gateways installed in Ravelijn, namely 'UTwente Ravelijn rooftop' and 'UT Ravelijn indoor'. We extract messages that were sent by each sensor in Ravelijn and received by both gateways at the same time from the data set because this can exclude other factors that could affect the signal propagation.

There are a total of 148 sensors sending messages using SF7 and 192 sensors using SF12, which are received by the two gateways in Ravelijn at the same time. Among these sensors, 126 out of 148 and 120 out of 192 are specifically located within the building Ravelijn. However, for demonstration purposes, we have selected only 5 sensors from each of the floors 1, 2, 3, and 4 in Ravelijn. Out of these, only 2 sensors from floor 5 were received. Therefore, a total of 22 sensors are plotted. For the box plots depicting the distributions of RSSI values of messages received by the gateways from the sensors in Ravelijn, refer to Appendix A.

As Figure 4 shows, the x-axis represents the sensors, of which identity can be separated into three parts. The first part represents in which room they are installed, the second part represents its location ID on the UT campus, and the third part represents its 64-bit unique Extended Unique Identifier (EUI), in which only the first and last characters are retained, the remaining characters are replaced with '*' for anonymity. For instance, 'RA5146-146-3*4' indicates that a sensor with the EUI 3*4 is installed in room RA5146, and its location ID on the UT campus is 146.

Figure 4 illustrates the distributions of RSSI values of messages received by gateway 'UT Ravelijn indoor' and gateway 'UTwente Ravelijn rooftop' from sensors in Ravelijn using SF7 and SF12, categorized by gateway. The RSSI values of the messages transmitted by sensors on floors 5 and 4 using SF7 and received by gateway 'UTwente Ravelijn rooftop' are higher compared to those received by gateway 'UT Ravelijn indoor'. However, for floors 3, 2, and 1, the RSSI values exhibit the opposite trend. Furthermore, the RSSI values of the messages transmitted by sensors using SF12 also exhibit a similar pattern. These findings indicate that the closer the sensor is to the gateway, the stronger the signal strength of the message received by the gateway is.

We plot the distributions of RSSI values of messages received by two gateways from sensors installed in Ravelijn using SF7 and SF12. Figure 5 shows the distributions of the RSSI values of messages received by gateway 'UTwente Ravelijn rooftop' and gateway 'UT Ravelijn indoor' from sensors in Ravelijn using SF7 and SF12, categorized by SF. In this figure, the upper point of the triangle represents the mean of the RSSI values. The box plots of the RSSI values of messages received by gateway 'UTwente Ravelijn rooftop' show that the maximum RSSI value of messages sent by each of 17 out of 22 sensors using SF7 is higher compared to those using SF12. Additionally, the overall RSSI value of messages sent by each of 13 out of these 17 sensors using SF7 is higher compared to the same 13 sensors using SF12. Similarly, the box plots of the RSSI values of messages received by gateway 'UT Ravelijn indoor' also demonstrate a similar pattern with 17 sensors, although some of them are different. Each of 14 out of these 17 sensors exhibits a higher overall RSSI value using SF7 in this case. By analysing the plots, it becomes evident that the signal strength of the messages transmitted by sensors using SF7 is higher compared to those using SF12 when sensors are close to the gateways, overall.

5 Conclusion

In our research, we investigate the impact of spreading factor on signal propagation, as well as the impact of gateway location on signal strength. We accomplish this by doing an analysis on the distributions of RSSI values of messages received by gateways from sensors using SF7 and SF12 respectively.

Firstly, the analysis reveals that the signal strength of the received signals can be different when they are received by gateways located in different positions. The closer the sensor is to the gateway, the higher the RSSI value of its message received by the gateway is. By strategically placing the sensor close to the gateway or adjusting the position of the gateway itself, we can optimize the signal transmission in the LoRaWAN network.

Furthermore, it is evident from the data that sensors transmitting signals using SF7 exhibit higher signal strength than those using SF12 when sensors are close to the gateways, overall. By studying this effect, we can gain valuable insights on enhancing the performance of the LoRaWAN network. Therefore, it's advisable to use the lowest spreading factor, namely SF7, when the sensor is close to the gateway.

In conclusion, the results highlight the importance of gateway location and spreading factor selection to optimize LoRa signal transmission. These findings provide valuable insights into optimizing the LoRaWAN network performance.

References

[1] AMekki, Kais, Eddy Bajic, Frederic Chaxel, and Fernand Meyer. 'A Comparative Study of LPWAN Technologies for Large-Scale IoT Deployment'. ICT Express 5, no. 1 (1 March 2019): 1–7. https://doi.org/10.1016/j.icte.2017.12.005.

[2] The Things Network. 'What Are LoRa and LoRaWAN?' Accessed 20 June 2023. https://www.thethingsnetwork.org/docs/lorawan/what-is-lorawan/.

[3] Choi, Kwon Nung, Harini Kolamunna, Akila Uyanwatta, Kanchana Thilakarathna, Suranga Seneviratne, Ralph Holz, Mahbub Hassan, and Albert Y. Zomaya. 'LoRadar: LoRa Sensor Network Monitoring through Passive Packet Sniffing'. ACM SIGCOMM Computer Communication Review 50, no. 4 (26 October 2020): 10–24. https://doi.org/10.1145/3431832.3431835.

[4] The Things Network. 'RSSI and SNR'. Accessed 24 June 2023. https://www.thethingsnetwork.org/docs/lorawan/rssi-and-snr/.

[5] Ferreira, Ana Elisa, Fernando M. Ortiz, Luís Henrique M. K. Costa, Brandon Foubert, Ibrahim Amadou, and Nathalie Mitton. 'A Study of the LoRa Signal Propagation in Forest, Urban, and Suburban Environments'. Annals of Telecommunications 75, no. 7–8 (August 2020): 333–51. https://doi.org/10.1007/s12243-020-00789-w.

[6] 'Sending and Receiving Messages with LoRaWAN | DEVELOPER PORTAL'. Accessed 2 July 2023. https://lora-developers.semtech.com/documentation/tech-papers-and-guides/sending-and-receiving-messages-with-lorawan/sending-and-receiving-messages/.

[7] Torres, Ana Paula Alves, Claudio Bastos Da Silva, and Horácio Tertuliano Filho. 'An Experimental Study on the Use of LoRa Technology in Vehicle Communication'. IEEE Access 9 (2021): 26633–40. https://doi.org/10.1109/ACCESS.2021.3057602.

A Box plots

Link of the box plots depicting distributions of RSSI values of messages received by gateway 'UTwente Ravelijn rooftop' and gateway 'UT Ravelijn indoor' from sensors using SF7 and SF12: https://github.com/xianzhi-wu/lorawan/blob/main/RA-gateways-SF7-SF12-RSSI.pdf