Dataset: Environmental Impact on the Long-Term Connectivity and Link Quality of an Outdoor LoRa Network

Pei Tian[1,3], Fengxu Yang[1,6], Xiaoyuan Ma[2], Carlo Alberto Boano[5],
Xin Tian[1], Ye Liu[4], and Jianming Wei[1]

[1] Shanghai Advanced Research Institute, Chinese Academy of Sciences, China   [2] SKF Group, China
[3] University of Chinese Academy of Sciences, China   [4] Nanjing Agricultural University, China
[5] Institute of Technical Informatics, Graz University of Technology, Austria
[6] School of Information Science and Technology, ShanghaiTech University, China

{tianpei2018, tianx, wjm}@sari.ac.cn, yangfx@shanghaitech.edu.cn
ma.xiaoyuan.mail@gmail.com, cboano@tugraz.at, yeliu@njau.edu.cn

ABSTRACT

Recently, several datasets shedding light on connectivity aspects in real-world LoRa networks have been provided to the community. However, they typically only involve a limited number of nodes, deal with unidirectional communication only, or focus on very specific physical layer settings. More importantly, existing datasets typically lack fine-grained environmental information such as the temperature in the surroundings of each node, which is known to have a strong impact on communication performance. In this work, we provide the community with a comprehensive dataset that fills all these gaps. We have collected detailed connectivity information in an outdoor LoRa network composed of 21 nodes for more than four months. Our dataset does not only focus on network-level performance (e.g., the average number of correctly-exchanged packets), but sheds light on link-level information such as the received signal strength, signal-to-noise ratio, and the number of available neighbours over time. We further collect environmental information from an online weather site, as well as the on-board temperature of each node in the network, which varies considerably across the deployed locations. We collect all this information while perpetually changing physical layer settings such as the spreading factor and the RF channel. A preliminary analysis of our dataset, which is available in Zenodo[1], reveals that temperature has a significant correlation with the link quality and connectivity in the outdoor LoRa network, confirming the findings of earlier studies.

CCS CONCEPTS

• **Networks** → *Network measurement*; *Network experimentation*; *Network performance analysis*; *Wide area networks*.

KEYWORDS

ChirpBox, Connectivity, Internet of Things, Link quality, LoRa, Physical layer settings, SX1276, Temperature, Weather.

---

[1]Dataset and scripts are available at: https://doi.org/10.5281/zenodo.5594944.

1 INTRODUCTION

LoRa is currently one of the most widespread and well-known radio technologies for building low-power wide area networks (LPWAN). Thanks to its long communication range, the remarkably low device cost, and the growing ecosystem, an increasing number of LoRa-based systems is being deployed all over the world, especially in the context of outdoor Internet of Things (IoT) applications such as intelligent transportation [2], smart metering [8], precision agriculture [11, 14], and wild animal tracking [15].

Compared to indoor scenarios, outdoor LoRa nodes are required to work under more extreme environments and withstand very diverse weather conditions and temperature fluctuations. Such environmental factors are hard to model and have a strong impact on LoRa's communication performance [4, 9], as they can degrade connectivity, link quality, and – in turn – affect key system properties such as reliability, end-to-end latency, and achievable lifetime.

For this reason, the existence of a dataset specifically focusing on the environmental impact on LoRa's connectivity and link quality over a long timespan is of particular interest to the community. In particular, such a dataset can serve as a reference to developers of LPWAN-based systems, helping them to understand how the environment can affect LoRa links or even to predict potential issues before actually deploying their final product(s).

Limitations of existing datasets. The research community has already published a few LoRa datasets focusing on the connectivity of real-world LoRa systems. However, many of these datasets do not record any environmental information [1, 3]. The dataset published by Cardell-Oliver et al. [6] is a notable exception, in which weather information fetched from a nearby station is recorded and included in the traces. Also the temperature of five transmitters is reported, but not that of remaining nodes, which is fundamental to model the impact of temperature on the connectivity and quality of individual links [5]. In fact, temperature strongly depends on the deployment location: a node placed under the shadow of some vegetation and

a node deployed on top of a roof may experience a temperature difference by up to 35 °C, as later illustrated in Fig. 2.

Another key limitation of existing LoRa datasets is that they are not meant to *systematically benchmark* the connectivity and quality across all links in the network as was done, e.g., using the TelosB nodes deployed in the FlockLab testbed [10]. Although information such as transmission/reception timestamp, received signal strength (RSS), and signal-to-noise ratio (SNR) of individual packets is often reported [1, 3, 6], the actual packet transmissions are typically infrequent (e.g., only every 10 minutes) and non homogeneous (e.g., packets have different payload lengths, and are only exchanged between specific nodes). Because of this, it is hard to derive a comprehensive picture about the number of neighbours of each node and about short-term variations in the link quality across different links. Furthermore, packets are often transmitted with specific physical layer settings that are statically selected at compile time.

Our contributions. We collect a comprehensive dataset about the connectivity and link quality in an outdoor LoRa network using ChirpBox [13], an infrastructure-less LoRa testbed consisting of 21 nodes deployed over 28 hectares in the city of Shanghai, China. The dataset captures over 400 thousand packets exchanged over 4 months (May – September 2021) using 18 combinations of spreading factors (SFs) and frequency channels. Every node periodically transmits packets to all surrounding nodes using different PHY settings combinations and logs connectivity and link quality statistics as well as its on-board temperature. Weather information is also reported based on the readings of an online weather site [12].

To the best of our knowledge, this is the first public long-term dataset capturing the connectivity and link quality of an outdoor LoRa network with detailed environmental information. In addition to the dataset itself, we provide evaluation scripts for data analysis and visualization, in order to facilitate data exploration and re-use. According to our preliminary analysis, the dataset shows clearly a negative correlation between temperature and link quality, confirming the findings of earlier studies such as [4] and [7].

The rest of this paper is organized as follows: Sect. 2 explains the setup used to collect the dataset and summarizes its features. Sect. 3 describes the scripts used to generate and analyse/visualize the data. Finally, Sect. 4 presents interesting findings from our preliminary data exploration, and Sect. 5 concludes the paper.
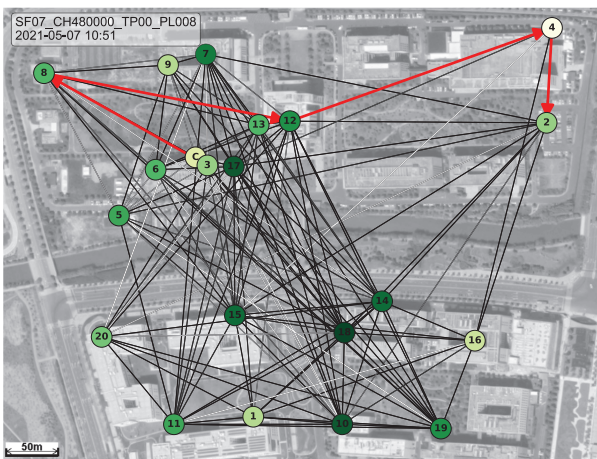

Figure 1: Location of the 21 nodes deployed using ChirpBox.

2 THE DATASET

We describe next how the dataset was created and its structure.

Measurement setup. We use ChirpBox [13] to set up an outdoor LoRa network comprising 21 nodes embedding a Semtech SX1276 radio in the premises of the Shanghai Advanced Research Institute, as illustrated in Fig. 1. To measure detailed connectivity and link quality information, we exploit the *connectivity evaluation service* provided by ChirpBox [13]. Specifically, a control node connected to a personal computer (node **C** in Fig. 1), periodically instructs all the nodes in the network to perform a link quality evaluation by flooding pre-defined control commands using LoRaDisC [13]. Each node is scheduled to send N probe packets (with 8-byte payload) in a round-robin manner, such that only one node sends packets at a given time, while the other nodes count the number of received packets and log both the RSS and SNR. All probe packets are sent with a transmission (TX) power of 0 dBm: we use this value instead of the maximum supported value (14 dBm) to form a multi-hop network. The SF and RF channel to be used when exchanging messages are piggybacked in the control commands flooded by the control node. In this dataset, we iterate through all supported SFs (from 7 to 12) and three RF channels (470, 480, and 490 MHz). When setting $N = 20$, it takes 50 minutes to iterate through all SFs on a given RF channel. After completing the connectivity evaluation for a given SF, the on-board temperature of each ChirpBox node is logged using the built-in sensor available in the SX1276 transceiver. Once all SFs are analyzed, also the weather information from an online weather site [12] is collected and stored: this information includes temperature, humidity, wind speed, and air pressure.

Dataset structure. Table 1 summarizes the structure of the collected dataset, which contains three types of attributes. The *basic* attributes refer to the measurement configuration: `time` refers to the beginning of a connectivity evaluation, whereas the remaining four attributes (`sf`, `channel`, `tx_power`, and `payload_len`) refer to the nodes' configuration. The *connectivity and link quality* attributes are statistics about the packet reception rate (PRR), RSS, SNR, and the nodes that can be successfully reached. Finally, the *environmental* attributes refer to the temperature measured by each node and to the weather information fetched from the online site.

3 DATA COLLECTION & ANALYSIS SCRIPTS

Besides making our dataset publicly available, we also want to (i) make it easy for others to replicate similar experiments, and (ii) facilitate data exploration and re-use. To this end, we provide the community also with scripts for data collection and analysis.

Data collection scripts. Whilst the hardware design and software used to set-up ChirpBox is already open-source[2], it is alone insufficient to replicate the same experiments that led to the creation of a connectivity dataset. We have therefore prepared and published additional scripts calling ChirpBox's connectivity evaluation service and retrieving the related statistics as well as environmental information. These scripts[3] instruct the control node to periodically iterate through different PHY settings configurations (e.g., SF and RF channel) and trigger the transmission of probe packets from all

---

[2]https://github.com/sari-wesg/ChirpBox
[3]https://github.com/sari-wesg/ChirpBox/tree/master/ChirpBox_manager/Tools/

Attribute	Description
Basic	
time	Local timestamp at the beginning of a connectivity evaluation
sf	Spreading factor used
channel	RF channel used (kHz)
tx_power	Transmission power in dBm
payload_len	Length of a probe packet in bytes
Connectivity and link quality	
node_degree_list	The list of neighbours for each node
node_link_matrix	Packet reception rate matrix
max/avg/min_snr	Maximum, average, and minimum SNR for all the received packets among all the links (dBm)
max/avg/min_rssi	Maximum, average, and minimum RSS values of all the received packets among all the links (dBm)
max/avg/min_degree	Maximum, average, and minimum number of neighbours per node
Environmental	
node_temperature_list	List of all temperature readings from each node in the network (°C)
weather_temperature	Temperature (°C) fetched from the weather site
wind_speed	Wind speed (m/s) fetched from the weather site
pressure	Atmospheric pressure (hPa) fetched from the weather site
humidity	Relative humidity (%) fetched from the weather site

Table 1: Attributes available in the collected dataset.

nodes in the network. After that, LoRaDisC is used to collect all the connectivity and link quality statistics from each node, as well as the sensed temperature stored on each node's internal memory. The script further shows how to use the weather API [12] to retrieve and process weather information.

Data analysis and visualization. We also provide scripts[4] to parse and visualize the data in an intuitive way. Specifically, data_analysis.py allows to derive network-level statistics (e.g., in terms of average number of correctly-exchanged packets), link-level statistics (e.g., in terms of SNR, RSS, and PRR), and node-level statistics (e.g., in terms of number of neighbours and temperature evolution over time). The script dataset.ipynb, instead, gives some examples on how to use data_analysis.py, how to visualize the various attributes in the dataset, as well as how to display the link quality on a map. Fig. 1 shows an example of the connectivity in the network when using SF=7 and 480 MHz as RF channel. The color shade of a node represents the degree of the node (i.e., the number of the neighbours of the node – the darker, the higher). The color shade of a link (i.e., the line between two nodes) shows the reliability (i.e., the corresponding PRR – the darker, the higher). Furthermore, the red lines also identify the longest network diameter: in this specific configuration, there are four hops from node **C** to node **2**.

[4]https://doi.org/10.5281/zenodo.5594944

4 PRELIMINARY DATASET ANALYSIS

Using the scripts provided in Sect. 3, we performed a preliminary analysis of our dataset and derived the following observations.

High temperature gradient across the deployed nodes. Fig. 2 shows the temperature recorded from the weather station (black dotted line) and the maximum and minimum temperature recorded across all the nodes in the network (red and blue solid line, respectively) over 72 hours. While the temperature recorded from the weather station remains rather constant between 20 and 30 °C, the on-board temperature of the nodes varies by as much as 47 °C. This reflects the different location of the nodes, which are deployed on top of roofs as well as below some vegetation on the ground.

High correlation between temperature and RSS. Fig. 3 shows the average RSS for a specific link in the network. Regardless of the employed SF, the RSS values drop by roughly 3 dB when temperature increases by 20 °C, in line with the results presented in [4].

The impact of temperature on connectivity and link quality is greater than that caused by other weather conditions. Figs. 4 and 5 illustrate the impact of various weather conditions (i.e.,
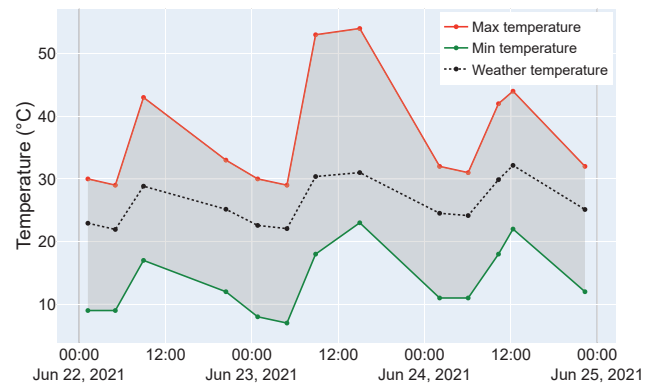
Figure 2: Temperature recorded by the weather station, and maximum/minimum temperature measured across all the nodes in the network over a timespan of three days.
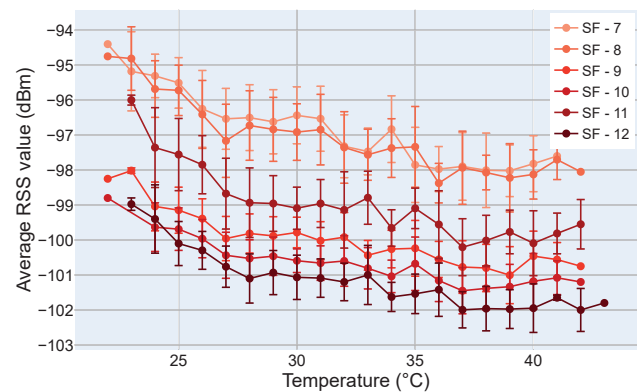
Figure 3: Relationship of the nodes' temperature on the RSS for link 19→10. The data refers to the timespan between June 25 and July 25, 2021. The RF channel used is 480 MHz.
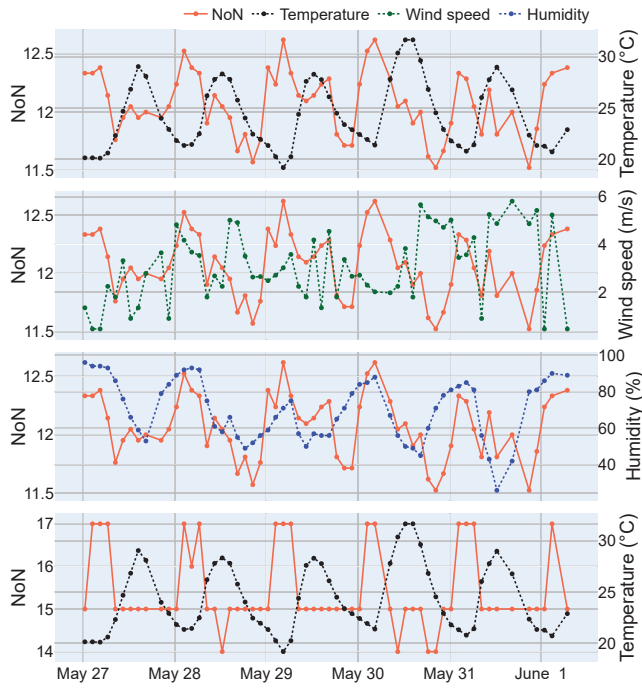
Figure 4: Impact of various weather conditions on the average number of neighbours (NoN) in the network (top three plots) and on the NoN of node 13 (bottom plot) over 6 days. We employed RF channel 480 MHz and SF=7.

temperature, wind speed, and humidity recorded by the weather station) on network connectivity (in terms of number of neighbours) and link quality (in terms of PRR), respectively. Temperature exhibits the closest correlation with the changes observed in both connectivity and link quality, which follow a diurnal variation. These findings are consistent with those of previous studies [4, 7].

5 CONCLUSIONS

We release our dataset and all related scripts to help developers of LoRa-based systems in studying and modelling the impact of the environment on network performance. We firmly believe that the availability of ChirpBox, together with these additions, will foster experimentation with LoRa and facilitate the development of dependable networking protocols in the years to come.

REFERENCES

[1] Michiel Aernouts, Rafael Berkvens, Koen Van Vlaenderen, and Maarten Weyn. 2018. Sigfox and LoRaWAN Datasets for Fingerprint Localization in Large Urban and Rural Areas. *Data* 3, 2 (April 2018).
[2] Wael Ayoub, Abed Ellatif Samhat, Mohamad Mroue, Hussein Joumaa, Fabienne Nouvel, and Jean-Christophe Prévotet. 2020. Technology Selection for IoT-Based Smart Transportation Systems. In *Vehicular Ad-hoc Networks for Smart Cities*. Springer, 19–29.
[3] Laksh Bhatia, Michael Breza, Ramona Marfievici, and Julie McCann. 2020. LoED: The LoRaWAN at the edge dataset: dataset. In *Proceedings of the 3rd Intl. Workshop on Data Acquisition To Analysis (DATA)*. ACM, 7–8.
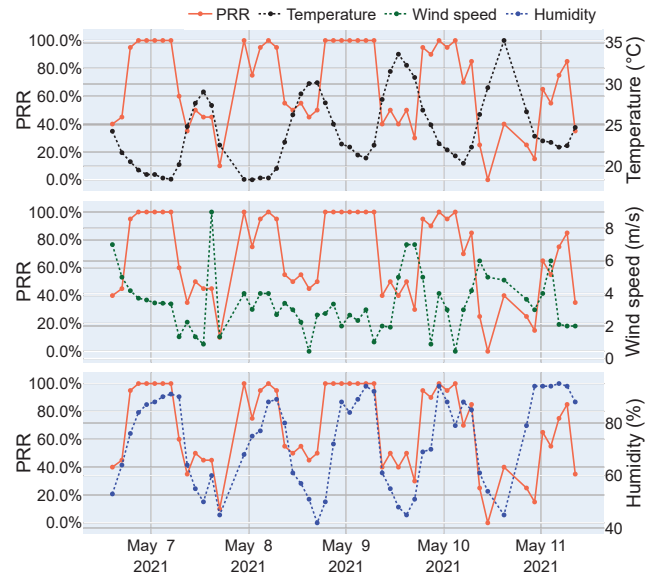
Figure 5: Impact of various weather conditions on the PRR of link 7→5. We employed RF channel 480 MHz and SF=7.

[4] Carlo Alberto Boano, Marco Cattani, and Kay Römer. 2018. Impact of Temperature Variations on the Reliability of LoRa: An Experimental Evaluation. In *Proceedings of the 7th Intl. Conference on Sensor Networks (SENSORNETS)*. SciTePress, 39–50.
[5] Carlo Alberto Boano, Hjalmar Wennerström, Marco Antonio Zúñiga, James Brown, Chamath Keppitiyagama, Felix Jonathan Oppermann, Utz Roedig, Lars-Åke Nordén, Thiemo Voigt, and Kay Römer. 2013. Hot Packets: A Systematic Evaluation of the Effect of Temperature on Low Power Wireless Transceivers. In *Proceedings of the 5th Extreme Conference on Communication (ExtremeCom)*. ACM, 7–12.
[6] Rachel Cardell-Oliver, Christof Hübner, Matthias Leopold, and Jason Beringer. 2019. Dataset: LoRa Underground Farm Sensor Network. In *Proceedings of the 2nd Intl. Workshop on Data Acquisition To Analysis (DATA)*. ACM, 26–28.
[7] Marco Cattani, Carlo Alberto Boano, and Kay Römer. 2017. An Experimental Evaluation of the Reliability of LoRa Long-Range Low-Power Wireless Communication. *Journal of Sensor and Actuator Networks (JSAN)* 6, 2 (June 2017).
[8] Marco Cattani, Carlo Alberto Boano, David Steffelbauer, Stefan Kaltenbacher, Markus Günther, Kay Römer, Daniela Fuchs-Hanusch, and Martin Horn. 2017. Adige: An Efficient Smart Water Network based on Long-Range Wireless Technology. In *Proceedings of the 3rd Intl. Workshop on Cyber-Physical Systems for Smart Water Networks (CySWATER)*. ACM, 3–6.
[9] Oana Iova, Amy L. Murphy, Gian Pietro Picco, Lorenzo Ghiro, Davide Molteni, Federico Ossi, and Francesca Cagnacci. 2017. LoRa from the City to the Mountains: Exploration of Hardware and Environmental Factors. In *Proceedings of the 2nd Intl. Workshop on New Wireless Communication Paradigms for the Internet of Things (MadCom)*. Junction Publishing, 317–322.
[10] Romain Jacob, Reto Da Forno, Roman Trüb, Andreas Biri, and Lothar Thiele. 2019. Dataset: Wireless Link Quality Estimation on FlockLab – and Beyond. In *Proceedings of the 2nd Intl. Workshop on Data Acquisition To Analysis (DATA)*. ACM, 57–60.
[11] M.L. Liya and D. Arjun. 2020. A Survey of LPWAN Technology in Agricultural Field. In *Proceedings of the 4th Intl. Conference on IoT in Social, Mobile, Analytics and Cloud (I-SMAC)*. IEEE, 313–317.
[12] OpenWeatherMap. [n.d.]. Weather API. [Online] http://openweathermap.org/api/ – Last accessed: 2021-09-15.
[13] Pei Tian, Xiaoyuan Ma, Carlo Alberto Boano, Fengxu Yang, Ye Liu, Xin Tian, Dan Li, and Jianming Wei. 2021. ChirpBox: An Infrastructure-Less LoRa Testbed. In *Proceedings of the 18th Intl. Conference on Embedded Wireless Systems and Networks (EWSN)*. Junction Publishing, 115–126.
[14] Wenju Zhao, Shengwei Lin, Jiwen Han, Rongtao Xu, and Lu Hou. 2017. Design and Implementation of Smart Irrigation System based on LoRa. In *Proceedings of the IEEE Globecom Workshops*. IEEE, 1–6.
[15] Nicholas Zinas, Sotirios Kontogiannis, George Kokkonis, Stavros Valsamidis, and Ioannis Kazanidis. 2017. Proposed Open Source Architecture for Long Range Monitoring. The Case Study of Cattle Tracking at Pogoniani. In *Proceedings of the 21st Pan-Hellenic Conference on Informatics (PCI)*. ACM, 1–6.